# The Handbook of Psycholinguistics

# Blackwell Handbooks in Linguistics

This outstanding multi-volume series covers all the major subdisciplines within linguistics today to offer a comprehensive survey of linguistics as a whole.
To see the full list of titles available in the series please visit www.wiley.com/go/linguistics-handbooks

*The Handbook of Contemporary Syntactic Theory*
Edited by Mark Baltin and Chris Collins

*The Handbook of Historical Linguistics*
Edited by Brian D. Joseph and Richard D. Janda

*The Handbook of Second Language Acquisition*
Edited by Catherine J. Doughty and Michael H. Long

*The Handbook of Applied Linguistics*
Edited by Alan Davies and Catherine Elder

*The Handbook of Pragmatics*
Edited by Laurence R. Horn and Gregory Ward

*The Handbook of Speech Perception*
Edited by David B. Pisoni and Robert E. Remez

*The Handbook of the History of English*
Edited by Ans van Kemenade and Bettelou Los

*The Handbook of English Linguistics*
Edited by Bas Aarts and April McMahon

*The Handbook of World Englishes*
Edited by Braj B. Kachru; Yamuna Kachru, and Cecil L. Nelson

*The Handbook of Educational Linguistics*
Edited by Bernard Spolsky and Francis M. Hult

*The Handbook of Clinical Linguistics*
Edited by Martin J. Ball, Michael R. Perkins, Nicole Müller, and Sara Howard

*The Handbook of Pidgin and Creole Studies*
Edited by Silvia Kouwenberg and John Victor Singler

*The Handbook of Language Teaching*
Edited by Michael H. Long and Catherine J. Doughty

*The Handbook of Phonetic Sciences, Second Edition*
Edited by William J. Hardcastle and John Laver

*The Handbook of Language and Speech Disorders*
Edited by Jack S. Damico, Nicole Müller, and Martin J. Ball

*The Handbook of Language Contact*
Edited by Raymond Hickey

*The Handbook of Computational Linguistics and Natural Language Processing*
Edited by Alexander Clark, Chris Fox, and Shalom Lappin

*The Handbook of Language and Globalization*
Edited by Nikolas Coupland

*The Handbook of Hispanic Sociolinguistics*
Edited by Manuel Díaz-Campos

*The Handbook of Language Socialization*
Edited by Alessandro Duranti, Elinor Ochs, and Bambi B. Schieffelin

*The Handbook of Phonological Theory, Second Edition*
Edited by John A. Goldsmith, Jason Riggle, and Alan C. L. Yu

*The Handbook of Intercultural Discourse and Communication*
Edited by Christina Bratt Paulston, Scott F. Kiesling, and Elizabeth S. Rangel

*The Handbook of Hispanic Linguistics*
Edited by José Ignacio Hualde, Antxon Olarrea, and Erin O'Rourke

*The Handbook of Historical Sociolinguistics*
Edited by Juan M. Hernández-Campoy and J. Camilo Conde-Silvestre

*The Handbook of Conversation Analysis*
Edited by Jack Sidnell and Tanya Stivers

*The Handbook of English for Specific Purposes*
Edited by Brian Paltridge and Sue Starfield

*The Handbook of Bilingualism and Multilingualism, Second Edition*
Edited by Tej K. Bhatia and William C. Ritchie

*The Handbook of Language Variation and Change*
Edited by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes

*The Handbook of Spanish Second Language Acquisition*
Edited by Kimberly L. Geeslin

*The Handbook of Chinese Linguistics*
Edited by C.-T. James Huang, Y.-H. Audrey Li, and Andrew Simpson

*The Handbook of Language, Gender, and Sexuality, Second Edition*
Edited by Susan Ehrlich, Miriam Meyerhoff, and Janet Holmes

*The Handbook of Language Emergence*
Edited by Brian MacWhinney and William O'Grady

*The Handbook of Bilingual and Multilingual Education*
Edited by Wayne E. Wright, Sovicheth Boun, and Ofelia Garcia

*The Handbook of Discourse Analysis, Second Edition*
Edited by Deborah Tannen, Heidi E. Hamilton, and Deborah Schiffrin

*The Handbook of English Pronounciation*
Edited by Marnie Reed and John M. Levis

*The Handbook of Classroom Discourse and Interaction*
Edited by Numa Markee

*The Handbook of Korean Linguistics*
Edited by Lucien Brown and Jaehoon Yeon

*The Handbook of Speech Production*
Edited by Melissa A. Redford

*The Handbook of Narrative Analysis*
Edited by Anna De Fina & Alexandra Georgakopoulou

*The Handbook of Contemporary Semantic Theory, Second Edition*
Edited by Shalom Lappin

*The Handbook of Portuguese Linguistics*
Edited by W. Leo Wetzels, Joao Costa, and Sergio Menuzzi

*The Handbook of Translation and Cognition*
Edited by John W. Schwieter and Aline Ferreira

*The Handbook of Linguistics, Second Edition*
Edited by Mark Aronoff and Janie Rees-Miller

*The Handbook of Dialectology*
Edited by Charles Boberg, John Nerbonne, and Dominic Watt

*The Handbook of Technology and Second Language Teaching and Learning*
Edited by Carol A. Chapelle and Shannon Sauro

*The Handbook of Psycholinguistics*
Edited by Eva M. Fernández and Helen Smith Cairns

# The Handbook of Psycholinguistics

Edited by

*Eva M. Fernández and*
*Helen Smith Cairns*

WILEY Blackwell

This edition first published 2018


*Registered Office*
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

*Editorial Office*
9600 Garsington Road, Oxford, OX4 2DQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.



*Library of Congress Cataloging-in-Publication Data*

Names: Fernández, Eva M., editor. | Cairns, Helen Smith, editor.
Title: The handbook of psycholinguistics / edited by Eva M. Fernández, Helen Smith Cairns.
Description: Hoboken, NJ : John Wiley & Sons Inc., [2017] | Includes bibliographical references and index. | Description based on print version record and CIP data provided by publisher; resource not viewed.
Identifiers: LCCN 2017003586 (print) | LCCN 2017012571 (ebook) | ISBN 9781118829547 (Adobe PDF) | ISBN 9781118829554 (ePub) | ISBN 9781118829509 (cloth)
Subjects: LCSH: Psycholinguistics–Handbooks, manuals, etc.
Classification: LCC P37.3 (ebook) | LCC P37.3 . H36 2017 (print) | DDC 401/.9–dc23
LC record available at https://lccn.loc.gov/2017003586

Cover Image: © OttoKrause/GettyImages
Cover Design: Wiley

Set in 10/12pt Palatino by SPi Global, Pondicherry, India

10 9 8 7 6 5 4 3 2 1

For
William and Maya Jagels
and
Sadie Haltom Cairns

# Contents

# Notes on Contributors

**Teresa Bajo** is Full Professor at the Department of Experimental Psychology of Granada University and head of the Memory and Language Research Group in the Mind, Brain and Behaviour Research Center. Her research is dedicated to the study of inhibitory control in language selection and in memory, cognitive processes in language translation, and interpreting.

**Andrew Barss** is a faculty member in the departments of Linguistics and the Program in Cognitive Science at the University of Arizona. He has conducted research on the syntax and semantics of anaphora, WH-movement, quantifier scope, and focus constructions.

**Catherine T. Best** is Chair in Psycholinguistic Research at MARCS Institute, Western Sydney University, Australia. Her research and theory have focused on cross-language speech perception, and on the role of regional accent variation in native speech perception and spoken word recognition. She is widely known for her Perceptual Assimilation Model (PAM) and its more recent extension to second language (L2) learners, PAM-L2.

**Ocke-Schwen Bohn** is the professor and chair of English Linguistics at Aarhus University, Denmark. He received his Ph.D. from Kiel University (Germany) and spent his time as a post-doc working with James Flege at the University of Alabama in Birmingham. With funding from German and Danish research agencies, and in collaboration with American, Canadian, and Australian colleagues, Bohn's research focuses on the causes and characteristics of foreign accented speech, speech perception (in infants, cross-language, and second language acquisition), and bilingual memory. He has published widely in journals such as *Applied Psycholinguistics*, *Journal of the Acoustical Society of America*, *Journal of the International Phonetic Association*, *Journal of Phonetics*, *Memory*, *Speech Communication*, and *Studies in Second Language Acquisition*.

**David Braze** is a linguist and senior scientist at Haskins Laboratories. He studies the cognitive structures and processes that support the human ability to fluidly assemble compositional meaning from more-or-less novel strings of words.

Dr. Braze's investigations of reading comprehension explore how its cognitive bases (ability to comprehend speech, word knowledge, decoding skill, memory, executive function, and so on) are related to the ability to construct meaning from print, and how that may vary from one person to the next due to differences in biology or experience. Central to his research are questions of how lexical, grammatical, semantic and pragmatic processes interact with one another to yield the apprehension meaning from language, whether perceived by ear or by eye.

**Trevor A. Brothers** is a doctoral candidate in the Department of Psychology at UC Davis. His research focuses on anticipatory processes in sentence and discourse comprehension.

**Esteban Buz** is a post-doctoral researcher and lecturer at Princeton University. He received his Ph.D. in the Brain and Cognitive Sciences at the University of Rochester in 2016. His graduate work was partly funded by a pre-doctoral National Research Service Award from the National Institutes of Health (F31HD083020), a Provost Fellowship from the University of Rochester, and a National Institutes of Health training grant awarded to the Center for Language Sciences at the University of Rochester (T32DC000035). His research focuses on language production and specifically how and why speakers' vary their speech in and across different situations. Of special focus in this researcher is whether speech variability is partly for achieving and maintaining robust communication.

**Krista Byers-Heinlein** (B.A., McGill University; M.A., Ph.D., University of British Columbia) is Associate Professor in the Department of Psychology at Concordia University, in Montreal, Canada. She holds the Concordia University Research Chair in Bilingualism and directs the Concordia Infant Research Lab. Her research investigates monolingual and bilingual language development in the early years.

**Helen Smith Cairns** is Professor Emerita, City University of New York, in the Department of Linguistics and Communication Disorders, Queens College, and the doctoral programs in Linguistics and in Speech, Hearing, and Language Sciences, at the CUNY Graduate Center. After receiving her Ph.D. in Psycholinguistics and Cognitive Processes at the University of Texas at Austin in 1970, she joined the Queens College faculty in 1971, where she taught until her retirement in 2003, having served as Department Chair and Dean of Graduate Studies and Research. She has pursued research in sentence processing and in first language acquisition, writing or editing six books (one co-authored with her husband, Chuck) and numerous articles and chapters. She is most proud of the students she has mentored over the years.

**Laurie Beth Feldman** is Professor of Psychology, at The University at Albany, State University of New York and Senior Scientist at Haskins Laboratories in New Haven, CT. She completed her undergraduate degree at Wellesley College and graduate degrees at the University of Connecticut. The unifying theme to her program of research is the question of how a language user concurrently manages two linguistic codes. These interactions include two languages, two writing systems for one language, native and accented speech, speech and text, and emoticon and text. Much of this work has depended on collaboration with people in China, Germany, Israel, The Netherlands, Serbia, and Turkey and was supported by funds from NAS, NSF, NICHD, IARPA, and Fulbright. She is a Fellow of APS, and a member of the AAAS and the Psychonomic Society.

**Fernanda Ferreira** is Professor of Psychology and Member of the Graduate Group in Linguistics at the University of California, Davis. She obtained her Ph.D. in Cognitive Psychology in 1988 from the University of Massachusetts, Amherst, and prior to moving to UC Davis in 2015, she held faculty positions at Michigan State University and the University of Edinburgh. She has published over 100 papers and her research has been funded by the National Sciece Foundation and the National Institutes of Health in the United States, and the Economic and Social Research Council in the United Kingdom. She served as Editor in Chief of the *Journal of Experimental Psychology: General*, and she is currently an Associate Editor of *Cognitive Psychology* and of *Collabra*, an Open Access journal recently launched by University of California Press. She is a Fellow of the American Psychological Society and the Royal Society of Edinburgh, and is currently an elected member of the Psychonomic Society's Governing Board.

**Eva M. Fernández** received her Ph.D. in Linguistics in 2000. She is Professor at Queens College, City University of New York (CUNY) in the Department of Linguistics and Communication Disorders, and doctoral faculty at the CUNY Graduate Center in the Programs in: Linguistics; Hispanic and Luso-Brazilian Literatures and Languages; and Speech-Language-Hearing Sciences. Her research focuses on cross-linguistic language processing in monolingual and bilingual populations; she has additional research interests in measures of student success in higher education. Her work has been supported by grants from the National Science Foundation, the United States Department of Education, and the Association of American Colleges & Universities. She currently serves as Assistant Provost at Queens College.

**Julie Franck** is Maître d'Enseignement et de Recherche at the University of Geneva. Her work combines theoretical insights from linguistic theory and psycholinguistic experimental methods with the aim of identifying general principles of syntactic production and comprehension in adults and children. Her interest in exploring the relations between grammar and processing has developed through several topics including agreement, the acquisition of word order and subordination, the learning of hierarchical structure in artificial grammars, and the processing of islands.

**Angela D. Friederici** studied German, Romance languages, linguistics, and psychology in Bonn, Germany and Lausanne, Switzerland. She obtained her doctorate from Bonn University, Germany and conducted postdoctoral research at the

Massachusetts Institute of Technology. After being appointed full professor of Cognitive Psychology at the Free University Berlin, Germany, she was founding director of the Max Planck Institute for Human Cognitive and Brain Sciences, where she is head of the Department of Neuropsychology. Her research is focused on the neuroscience of language. Angela D. Friederici is Vice President of the Max Planck Society, and honorary professor at the Universities of Leipzig, Potsdam and at the Charité Berlin.

**Chiara Gambi** is a post-doctoral researcher at the University of Edinburgh. She received her Ph.D. in Psychology from the same university in 2013 and has subsequently been a post-doc with the Psycholinguistics Group at Saarland University before joining the Developmental Lab at the University of Edinburgh in 2014. She has published in Cortex, Journal of Experimental Psychology: Learning, Memory, and Cognition, and in Language, Cognition, and Neuroscience, Her research focuses on the relationship between comprehension and production, and particularly between the latter and prediction processes.

**Merrill F. Garrett**, Ph. D. University of Illinois, 1965. Professor Emeritus, University of Arizona, 2008-date. Professor of Psychology and Linguistics, University of Arizona, 1987–2008. Assist., Assoc., Full Professor of Psychology, Massachusetts Institute of Technology, 1968–87. Research area: language processing, with contributions in language development, language comprehension and language production. He was Director of Cognitive Science at Arizona 1987–2007, and was instrumental in its development and related programs in Cognitive Neuroscience. Fulbright fellow (Australia, 1974); Sloan Fellow in Cognitive Science (UMass, 1979); Fellow, Japan Society for Promotion of Science (Tokyo, 1997); Leibnitz Professor in Cognitive Science, University of Leipzig (1998); Fellow, Institute for Advanced Study, Berlin, Germany: 1999–2000, Fellow, American Association for the Advancement of Science (elected 2002).

**Roberta Michnick Golinkoff**, Unidel H. Rodney Sharp Chair in the School of Education at the University of Delaware, and also in the Departments of Psychological and Brain Science and Linguistics and Cognitive Science, directs the Child's Play, Learning, and Development Lab whose focus is language acquisition, spatial development, and the benefits of playful learning. Her research has resulted in numerous professional articles and 14 books, including books for parents and practitioners. Her latest book is entitled, *Becoming Brilliant: What Science Tells Us About Raising Successful Children* (APA Press). Dr. Golinkoff has received various awards including the John Simon Guggenheim Fellowship, the James McKeen Cattell Sabbatical award, and with Dr. Hirsh-Pasek, the James McKeen Cattell Award from the Association for Psychological Science and the Society for Research in Child Development Distinguished Scientific Contributions Award. She is also a Francis Alison Scholar, the highest honor given to faculty at her University. Dr. Golinkoff has appeared on many television and radio shows and speaks at conferences around the world about children's development.

**Tao Gong** is a computational modeler and research scientist at Haskins Laboratories. His research investigates the origins and evolution of language, the unique ways in which biology and culture interact during language learning, and the roles of language and literacy related skills on reading behaviors. By means of agent-based and mathematical modeling, as well as behavioral, artificial language learning, and eye-movement experiments, Dr. Gong's studies of language evolution and learning explore how domain-general cognitive capacities (sequential learning, pattern extraction, memory, and so on) give rise to specific language processing mechanisms underlying particular language structures. His work addresses questions concerning how individual learning, biological evolution, and socio-cultural transmission interact with one another to yield culturally specific language and linguistic behaviors.

**Eve Higby** is a post-doctoral researcher at the University of California, Riverside. She received her Ph.D. in Speech-Language-Hearing Sciences from the City University of New York Graduate Center, with a focus on neurolinguistics and cognitive science. Her research focuses on brain and language changes in bilingualism and aging and the interaction between language and other cognitive faculties. Her publications include topics such as multilingualism and the brain, cross-linguistic influence in bilingualism, brain and language in aging, and second language acquisition. She is currently working on the second edition of the *Language and the Brain* book by Cambridge University Press with Loraine K. Obler, Kris Gjerlow, and Ioulia Kovelman.

**Kathy Hirsh-Pasek** is the Stanley and Debra Lefkowitz Faculty Fellow in the Department of Psychology at Temple University and a Senior Fellow at the Brookings Institution. Her research examines the development of early language and literacy as well as the role of play in learning. With her long-term collaborator, Roberta Golinkoff, she is author of 14 books and hundreds of publications. She is the recipient of three American Psychological Association Awards: the Bronfenbrenner Award, the Distinguished Service to Psychological Science Award, and the Distinguished Lecturer Award. She is the President Elect of the International Society for Infant Studies and served as an Associate Editor of Child Development. She is on the Steering Committee of the Latin American School for Education, Cognitive, and Neural Science and Vroom, the Boston Children's Museum, and Jumpstart. Her 2016 book, *Becoming Brilliant* (with Golinkoff) was on *The New York Times* best sellers list in education and parenting. Dr. Hirsh-Pasek is also routinely quoted on national media such as *The New York Times* and *National Public Radio*.

**Liv J. Hoversten** is a doctoral candidate in the Department of Psychology at UC Davis. Her research focuses on language competition and language control in bilingual speakers.

**Nina Hyams** is Professor of Linguistics at the University of California, Los Angeles. Her primary area of research is childhood language development, especially the acquisition of syntax and semantics. She is author of the book *Language Acquisition and the Theory of Parameters* (D. Reidel Publishers), a milestone in language acquisition research and she co-authors (with Robert Rodman) a popular introductory textbook, *An Introduction to Language* (Cengage, Wadsworth Learning). She has published numerous articles on the development of syntax, morphology and semantics in children and she frequently lectures on these topics throughout the United States, Europe and Japan. She received her Ph.D from the Graduate Center of the City University of New York in 1983.

**T. Florian Jaeger** holds an M.A. in Linguistics and Computer Science from Humboldt University to Berlin (2001) and a Ph.D. in Linguistics with a designation in the Cognitive Sciences from Stanford University (2006). Other education includes stints at MIT, UC Berkeley, Edinburgh, and at post-doc in psychology at UCSD. In 2007, he joined the University of Rochester, where he now is a Professor in the Departments of Brain and Cognitive Sciences, Computer Science, and Linguistics. He is the Director of the Center for Language Sciences, an Alfred P. Sloan Fellow in the Neurosciences, and NSF CAREER recipient in Robust Intelligence. Research in his Human Language Processing lab focuses on the development of computational frameworks for language production and understanding. This includes research on the consequences of noise and variability (e.g., between talkers) for communication and the study of inference under uncertainty and learning in production and comprehension. This research is funded by the National Science Foundation and the National Institutes for Health.

**Judy Kegl** is Professor of Linguistics and Director of the Signed Language Research Laboratory at the University of Southern Maine, where she teaches both linguistics and interpreting. Her primary research areas are the linguistics of American Sign Language (focusing on syntax), neurolinguistics of both signed and spoken languages, and predicate argument structure in English and ASL. She is best known for her work documenting the emergence of a signed language in Nicaragua, which has been ongoing for the past 25 years. She argues that it is a creation of language from language-like, but non language input that, via children's acquisition process, has yielded a full-fledged human language.

**Elaine C. Klein** is Professor of Linguistics Emerita at the CUNY Graduate Center. Professor Klein specializes in second language development and, recently, in studies of and interventions for immigrant students with special language and literacy needs. With Dr. Gita Martohardjono, Dr. Klein studied a population of adolescent emergent bilinguals who had been under-served and under-researched: Students with interrupted or inconsistent formal education (SIFE). Currently, as PI for Bridges for Academic Success, Dr. Klein and her team, funded by the New York State Education Department, have developed a set of research-based curricula for SIFE, and professional development for their teachers, now being used in school districts around New York State.

**Judith F. Kroll** is Distinguished Professor of Psychology at the University of California, Riverside and the former director of the Center for Language Science at Pennsylvania State University. The research that she and her students conduct concerns the way that bilinguals juggle the presence of two languages in one mind and brain. Their work, supported by grants from NSF and NIH, shows that bilingualism provides a tool for revealing the interplay between language and cognition that is otherwise obscure in speakers of one language alone. She was one of the founding editors of the journal *Bilingualism: Language and Cognition* (Cambridge University Press).

**Patricia K. Kuhl** is the Bezos Family Foundation Endowed Chair for Early Childhood Learning, Co-Director of the UW Institute for Learning & Brain Sciences, Director of the NSF-funded Science of Learning Center, and Professor of Speech and Hearing Sciences. She is internationally recognized for her research on early language and brain development, and studies that show how young children learn. Dr. Kuhl's work has played a major role in demonstrating how early exposure to language alters the brain. It has implications for critical periods in development, for bilingual education and reading readiness, for developmental disabilities involving language, and for research on computer understanding of speech.

**Dani Levine** is a Ph.D. candidate in the Department of Psychology at Temple University. She received her B.A. in Psychology and Biology with a concentration in Neuroscience from Williams College. Her research interests include language acquisition, cognitive development, and early childhood education. Dani is currently studying the processes young children use to parse events into meaningful units and is examining the link between event segmentation and word learning. She also worked on a project, funded by the Institute of Education Sciences, in which she helped to develop a new way to measure language growth in preschoolers.

**Casey Lew-Williams** (B.A., University of California, Berkeley; Ph.D., Stanford University) is an Assistant Professor in the Department of Psychology at Princeton University. He co-directs the Princeton Baby Lab, and his research focuses on understanding how experience shapes language learning in infants, children, and adults.

**Sarah Roseberry Lytle** is the Director of the Outreach and Education division at the Institute for Learning & Brain Sciences (I-LABS) at the University of Washington. The I-LABS Outreach and Education team communicates the latest science of child development to those who can act on it, including parents, educators, policymakers, and opinion leaders. Dr. Lytle was previously a Postdoctoral Fellow at I-LABS. Before coming to the Institute, she earned a B.A. in Psychology and Spanish from the University of Notre Dame and a Ph.D. in Developmental Psychology at Temple University. Her research focused on the role of social cues in infants' and toddlers' language learning and how social interactions might help toddlers learn from screen media. Dr. Lytle was 2014–2016 Zero To Three Fellow.

**Fengyang Ma** is Visiting Assistant Professor in the School of Education at the University of Cincinnati. Her research interests center on the investigation of bilingualism from psycholinguistic and neurolinguistic perspectives. She received her Ph.D. from Tsinghua University in 2012 and then from 2012 to 2015 was a postdoctoral visiting scholar at Pennsylvania State University, working in the Center for Language Science and the Social, Life, and Engineering Sciences Imaging Center.

**Gita Martohardjono** is Professor of Linguistics and the Executive Officer of the Ph.D. Program in Linguistics at the Graduate Center, where she also co-directs RISLUS (Research Institute for the Study of Language in Urban Society). Her research interests center on psycholinguistic investigations of syntactic development in second language learning; the development and use of the first-learned language in heritage language speakers; and the development of academic literacy in child and adolescent bilinguals. Her most recent work examines cross-linguistic influence in first- and second-generation bilinguals through the use of ERP and eye-tracking methodologies. She is currently leading a team of linguists in developing native language tasks in 15 languages for the New York State Education Department (NYSED).

**Dana McDaniel** is a Professor of Linguistics and the Chair of the Department of Linguistics at the University of Southern Maine. Together with Professor Wayne Cowart, she developed the USM undergraduate Linguistics program from the ground up. She directs the Child Language Lab, which emphasizes children's syntax and language production. Her current focus is the potential role of the language production system in the evolution of syntax, cross-linguistic variation, and children's acquisition of syntax. Her work in these areas has been supported by the National Science Foundation.

**Cecile McKee** is a Professor of Linguistics at the University of Arizona. She has affiliate appointments in Psychology, Cognitive Science, and Second Language Acquisition and Teaching. She directs the Developmental Psycholinguistics Lab, where current research emphasizes children's production system and children's informal science learning. With collaborator Sam Supalla, McKee has also studied deaf children's literacy development in research that was supported by the James S. McDonnell Foundation. McKee is also active in outreach, especially via festivals and museums. She also served from 2006 to 2015 on University of Arizona institutional review boards. She worked as a program director at the National Science Foundation in 2001–2003 and 2011.

**Jürgen M. Meisel** is Emeritus Professor of Romance Linguistics at the University of Hamburg (Germany) and Adjunct Professor in the Department of Linguistics and Distinguished Fellow in the Language Research Centre at the University of Calgary. In 2004, he was awarded an honorary doctorate by the Faculty of Humanities, Lund University (Sweden). He directed the Research Center on Multilingualism at the University of Hamburg from its foundation in 1999 through 2006, and he is founding editor of the journal Bilingualism: Language and Cognition. His main research areas are bilingual first language acquisition, naturalistic second language acquisition by children and adults, Romance syntax, and historical linguistics.

**Petar Milin** is Senior Lecturer in Data Science at The University of Sheffield, United Kingdom. In the late 2000s he was among most active researchers in developing what is known as information-theoretic approach to language processing. He continued his work as one of the principal co-developers of the Naive Discrimination Learning model, which is simultaneously a computational model and research framework to investigate language acquisition and processing. He collaborates widely and on various topics with colleagues from the United States, Canada, UK, Belgium, Germany, and Italy. He is a member of the Psychonomic Society, and a co-founder of the International Quantitative Morphology Association.

**Luis Morales** is currently Professor at Universidad Loyola Andalucía (Spain), although the chapter was written during his academic affiliation at the University of Granada (Spain). His research is about lexical access in monolingual and bilingual speakers and the cognitive control processes involved in language selection. He focuses at the level of grammatical gender through production and comprehension tasks.

**Janet L. Nicol** is a Professor of Linguistics, Psychology, and Cognitive Science at the University of Arizona. She is an active faculty member in the Graduate Interdisciplinary Programs in Cognitive Science and Second Language Acquisition and Teaching. She has conducted extensive research on the comprehension of agreeing elements and discontinuous dependencies, and the production of agreement, testing various populations (including children, people with aphasia, and monolingual and bilingual young adults), with a range of methodologies (including self-paced reading, eye-tracking, ERP, and elicited production). Recent research has also focused on the learning of vocabulary and grammar in a second language and the factors that optimize such learning, including the manner of presentation of materials and learner-inherent capacities, such as working memory and executive function. Other research interests include how to optimize the learning of any kind of material, particularly in the context of online teaching. Her edited volume, *One Mind, Two Languages: Bilingual Language Processing*, was published in 2001.

**Loraine K. Obler** is Distinguished Professor of Speech-Language-Hearing Sciences, with a joint appointment in Linguistics, at the City University of New York Graduate Center. She also heads the Steering Committee of the MA

Program in Multilingual and Multicultural Speech-Language Pathology at Hadassah Academic College Jerusalem. Her publications in neurolinguistics focus on aphasiology (e.g., cross-language studies of agrammatism) and changes in comprehension and lexical retrieval associated with advancing age. Her bilingualism publications also include The Bilingual Brain: Neuropsychological and Neurolinguistic Aspects of Bilingualism (co-authored with Martin Albert) and Bilingualism Across the Lifespan (co-edited with Kenneth Hyltenstam), as well as studies on poor and excellent second-language learners. As well she is currently in the process of revising Language and the Brain (with Kris Gjerlow) with co-authors Eve Higby and Ioulia Kovelman.

**Robyn Orfitelli** is a Lecturer of Linguistics at the University of Sheffield. Her primary research areas are syntax and the childhood language development of syntax. She has published articles in a variety of areas, including the development of syntax in first and second language acquisition, experimental methodology, and prosody, and has a forthcoming monograph, *Argument Intervention in the Acquisition of A-movement* (Springer). She has presented conference papers and lectures on these topics in North America, Europe, and Australia. She received her Ph.D. from the University of California, Los Angeles in 2012.

**Daniela Paolieri** is Professor of Memory and Augmentative Communication Systems at Granada University. In her research, she has focused on grammatical gender processing in language production and comprehension, addressing this issue both in monolingual and bilingual contexts.

**Jennifer S. Pardo** is Professor of Psychology and Director of the Speech Communication Laboratory at Montclair State University. She received her Ph.D. in Cognitive Psychology from Yale University in 2000, and has held academic positions at Barnard College, Wesleyan University, and The New School for Social Research. Her research centers on the production and perception of spoken language with an emphasis on understanding variation and convergence in phonetic form, and has been supported by grants from the National Science Foundation and the National Institutes of Health. Some recent publications of this work have appeared in *Journal of Memory & Language, Frontiers in Psychology*, and *Attention, Perception, & Psychophysics*.

**Martin J. Pickering** is Professor of the Psychology of Language and Communication at the University of Edinburgh. He is the author of more than 130 journal articles and numerous other publications in language production, comprehension, dialogue, reading, and bilingualism. His articles have appeared in *Psychological Bulletin*; *Trends in Cognitive Sciences*; *Journal of Neuroscience*; *Cerebral Cortex*; *Psychological Science*; *Cognitive Psychology*; *Journal of Memory and Language*; *Cognition*; and *Journal of Experimental Psychology: Learning, Memory, and Cognition*. In particular, he published "Toward a mechanistic psychology of dialogue" (*Behavioral and Brain Sciences*, 2004) and "An integrated theory of language production and comprehension" (*Behavioral and Brain Sciences*, 2013), both with Simon Garrod. He is a Fellow of the Royal Society of Edinburgh and a former editor of the *Journal of Memory and Language*.

**David B. Pisoni** holds the positon of Distinguished Professor of Psychological and Brain Sciences and Chancellor's Professor of Cognitive Science at Indiana University in Bloomington. He has carried out seminal research on human speech perception, spoken word recognition, language processing, and perceptual development in infants and children. For the past 24 years, Professor Pisoni has worked with researchers at the IU School of Medicine on clinical problems associated with hearing impairment in deaf children who use cochlear implants. His work has had important clinical applications for understanding the enormous variability in speech and language outcomes following implantation. Throughout his career, Professor Pisoni has made significant contributions in basic, applied and clinical research in areas of speech perception, production, synthesis, and spoken language processing.

**Elizabeth Pratt** received her Ph.D. in Linguistics from the City University of New York, where she is currently engaged with linguistic and educational research through the Center for Advanced Study in Education (CASE), and the Research Institute for the Study of Language in Urban Society (RISLUS). Her primary interests are in the unique integration of linguistic theory and research within a general framework of cognitive processing. Her recent research has concentrated on structural and prosodic effects on agreement processing and comprehension, with the goal of incorporating these linguistic insights into a comprehensive model of memory, executive control, and attention. Aligned with this research program is the parallel development and testing of pedagogical strategies and interventions for speakers and learners from diverse language backgrounds and profiles.

**Petra B. Schumacher** is Professor of Empirical Linguistics at the Department of German Language and Literature I at the University of Cologne, Germany. She received her Ph.D. in Linguistics from Yale University. Her research focuses on processes at the interface of syntax, semantics and pragmatics, including reference resolution, information structure, and experimental pragmatics. She is primarily interested in the time-course of language processing and conducts research with different populations and languages.

**Michael A. Skeide** studied Linguistics at Heidelberg University, Germany, and Neurobiology at Harvard University. For his work on the functional and structural plasticity underlying the developmental emergence of higher-order language processing, he received a Ph.D. in Psychology from the University of Leipzig, Germany. He is now a postdoctoral researcher at the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, where he studies the neural basis of normal and impaired language and literacy development.

**Eva Smolka** is a psycho- and neurolinguist. She studied cognitive psychology and English linguistics at Bar-Ilan University and the University of Haifa, Israel, and at the Philipps-University Marburg, Germany. In her research, she explores the storage and processing of complex word formations in lexical memory. She applies psycho- and neurolinguistic methods to study word formation in healthy adults, language acquisition in children and bilinguals, as well as language loss in patients with language impairments such as aphasia. Dr. Smolka is currently at the University of Konstanz, Germany, and is collaborating with many researchers from Europe, Israel, Canada, and the United States.

**Kristina Strother-Garcia** is a student in the Department of Linguistics & Cognitive Science at the University of Delaware. Her research interests include computational/mathematical phonology, syllable theory, abstract representations of linguistic structures, and grammatical inference.

**Matthew J. Traxler** is a professor of psychology and administrator at the University of California, Davis. His research focuses on sentence and discourse comprehension in healthy adults, deaf adults, and patients with schizophrenia.

**Kamil Ud Deen** was born and raised in Nairobi, Kenya, and came to the United States as a young man. He received his Ph.D. from UCLA, and is currently as associate professor in Linguistics at the University of Hawaii at Manoa. He has conducted research on his native Swahili, as well as on a wide variety of languages, including Korean, Japanese, Thai, Tagalog, Serbian, and Vietnamese. His interests lie in cross-linguistic research on the acquisition of morphology and syntax.

**Virginia Valian** is Distinguished Professor of Psychology at Hunter College and is a member of the doctoral faculties of Psychology, Linguistics, and Speech-Language-Hearing Sciences at the CUNY Graduate Center. She directs the Language Acquisition Research Center. Valian works in two areas—the psychology of language and gender equity. In first language acquisition, Valian performs research with the aim of developing a model of acquisition that specifies what is innate, how input is used by the child, and how the child's syntactic knowledge interacts with knowledge in other linguistic and extra-linguistic domains. To approach those questions she uses a variety of methods, including computer-assisted corpus analysis, comprehension experiments, elicited imitation experiments, and elicited production experiments. Valian's second language area is the relation between bilingualism and higher cognitive functions in adults; she thinks there is one though she has yet to find it. To learn more about her work: http://maxweber.hunter.cuny.edu/psych/faculty/valian/valian.htm, http://littlelinguist.hunter.cuny.edu/, and http://www.hunter.cuny.edu/genderequity/.

**Jet M. J. Vonk** is a Ph.D. Candidate in the Neurolinguistics Laboratory of Loraine K. Obler, Ph.D., within the Speech-Language-Hearing Sciences department at the

Graduate Center, City University of New York (CUNY). Additionally, she is a Visiting Scholar at the Memory and Aging Center of the University of California, San Francisco (UCSF) working with Maria Luisa Gorno-Tempini, M.D., Ph.D., and Nina Dronkers, Ph.D. Her research projects focus on lexical-semantic aspects in healthy aging, Alzheimer's disease, and primary progressive aphasia and her broader interests include changes in language and cognition due to age and neurobiological factors in healthy aging and dementia. Multiple of her projects are in collaboration with the Taub Institute for Research on Alzheimer's Disease and the Aging Brain at Columbia University and the Language in the Aging Brain Project at Boston University/Boston VA Healthcare System.

**Ronnie B. Wilbur,** Ph.D, is a theoretical and experimental linguist who holds the rank of full professor in Linguistics and in Speech, Language, and Hearing Sciences at Purdue University. She has conducted research on sign languages for more than three decades, pioneering new methods for analyzing sign languages that parallel those available for spoken languages and developing educational applications for her research findings. She was the founding editor of the field's flagship journal, *Sign Language and Linguistics*. In addition to her funded research (NIH, NSF, Office of Education), guest lectureships and visiting professorships in the United States and Europe, Wilbur served a dozen years as Director of the Purdue Linguistics Program. A member of the Purdue University Book of Great Teachers, she received the 2014 Schleman Gold Medallion Award for accomplishments related to women's issues on campus, as well as a Focus Award for diversity and disability service, and a Seed for Success Award for research funding in excess of $1 million. In 2015 she was awarded the Purdue University Research and Scholarship Distinction Award, Purdue's highest research honor in Humanities and Social Sciences.

# Prologue

EVA M. FERNÁNDEZ AND HELEN SMITH CAIRNS

Queens College and Graduate Center, City University of New York

Psycholinguistics is the field of study that addresses how people process and acquire a central aspect of what it means to be human: language. The body of research surveyed in the chapters that follow addresses this essential faculty of the human species from a number of perspectives, drawing predominantly from the disciplines of linguistics and psychology, from cognitive science, and from neuroscience. Our goal in assembling this collection of contemporary research is to present the state of the field of psycholinguistics early in the twenty-first century, a field with vibrant research trends that, combined, provide a rich picture of how language works in the human mind and how it is acquired.

Psycholinguistics is a relatively new field, with origins in a seminar in 1953 at Indiana University held in conjunction with the Linguistics Institute, resulting in a book edited by Charles Osgood and Thomas Sebeok titled *Psycholinguistics: A Survey of Theory and Research Problems* (Osgood & Sebeok, 1954). Their approach to the study of language focused on three disciplines: linguistics, learning theory, and information theory. It laid out foundational questions regarding the mechanisms and units that underpin hearing and speaking. A reviewer of Osgood and Sebeok's book made the prescient observation that "the joint exploration which it describes is something more than just another interdisciplinary venture" (Newman, 1955: 1097). Indeed, in the mid-twentieth century the then-new field of generative linguistics collided with behaviorist psychology, resulting in a scientific revolution with many of the characteristics observed by Thomas Kuhn in his book *The Structure of Scientific Revolutions* (2012). The history of psycholinguistics has been a story of the influence of linguistic theory on theoretical psychology, and the emergence of psycholinguistics as a dramatically altered but ultimately autonomous and prolific science.

In the early days of psycholinguistics, linguistic theory was actually taken to be a theory of linguistic performance, but the falsification of the Derivational Theory of Complexity (Fodor, Bever, & Garrett, 1974) demonstrated the fallacy of that approach. However, the profound insight (Chomsky, 1959) that language is a mental construct—specifically, that knowledge of language is represented in the individual's mind/brain—changed psychology forever and replaced behaviorist models of language with a cognitive view that relies on mental representations and processes that underlie the linguistic life of humans. The depth and breadth of the chapters in this *Handbook* attest to the depth and breadth of the content of contemporary psycholinguistics, a field that has expanded well beyond its early conceptualization as providing a model of how linguistic competence is deployed in the production and comprehension of sentences. A glance at the table of contents of this *Handbook* reveals how far we have come from that original conceptualization: the research is informed by cognitive and neurocognitive frameworks as well as by information theory, explores sociocultural parameters, incorporates concepts from evolutionary biology, and has direct relevance to education, speech/language pathology, and medicine.

Parts 1 and 2 of the *Handbook* focus respectively on language production and language comprehension, but it is clear that in some ways this is an artificial distinction. There are representations and processes that are common to both speaking and comprehending, and they overlap as they unfold. The field is not only concerned with the production and comprehension of spoken language, but also of signed languages. Sign language has profound ramifications for cognition and offers insights and research avenues unavailable if we restrict the domain of study to spoken languages. Likewise, research concerned with how speakers of more than one language produce and comprehend their languages provides unique ways to explore the architecture of the mechanisms that underpin linguistic performance.

Both comprehension and production at the sentence level and beyond rely on the activation of lexical information, prosodic analysis, and internal parsing principles, as well as principles of linguistic organization subsumed under the formal grammar. Higher-level processes are invoked when speakers and hearers engage in conversations. Linguistic theory, while characterizing individual linguistic competences, describes universal characteristics of human languages that constrain representations at every level of both production and comprehension. Psychological processes involved in the production and comprehension of language (and in some cases multiple languages) go far beyond representations constrained by linguistic theory to encompass powerful processes of linguistic organization and parsing.

The acquisition of language, the subject of Part 3, has similarly undergone extensive revision and expansion since the early days of psycholinguistics. A child's development progresses from an initial state sensitive to universal properties of languages to a state consisting of fully formed representations of the native language. This development takes place in a remarkably brief period of time: by the time a child begins school (typically around 5 or 6 years old), a marvelously sophisticated mental system is in place. That trajectory is informed and constrained by basic principles of linguistic organization, as well as by the child's developing perceptual system, lexical store, and additional cognitive abilities. Powerful internal capacities of pattern recognition, statistical monitoring, and memory contribute to the acquisition of a child's native language. An explosion of research on the acquisition of two or more languages and also of signed languages has enriched what we know about language development. We have always known that a child must be exposed to a language to acquire it, but recent advances in contemporary research have augmented how we understand and describe the characteristics of linguistic input, the feedback available to the child, and the quality of interactions with the child's linguistic environment.

We are extremely fortunate to have recruited 52 leading scientists in contemporary psycholinguistics from 32 institutions in 9 different countries in North America, Australia, and Europe. Their contributions to this *Handbook* describe both the results of contemporary psycholinguistic research and the puzzles that remain for scholars to tackle in the future. To better frame the presentations in each of the chapters, each of the three sections begins with a chapter providing an overview of the contributions in that section, how they connect to one another, and how they relate to psycholinguistics in general. Our contributors have also strived to make the content accessible to readers who may not necessarily be experts in the sub-disciplines featured in each chapter. Our hope is that this volume will be of value to students and senior scholars alike and will make a contribution to the exciting, robust field of psycholinguistics.

The development of this volume has been very positively informed by the rich stimulation in the area of the study of language provided by the City University of New York, and we are grateful to all of our colleagues and students there for their support and insights, particularly Diane Bradley, Janet Fodor, and Irina Sekerina. We also owe a special thanks to the support team at Wiley: Danielle Descoteaux (who helped us envision the volume in its earliest phases), Mark Calley, and Tanya McMullin, as well as Manish Luthra, Vimali Joseph, and the editorial/production group. But we owe the most to our contributors who, individually and collectively, have made this volume an extensive and authoritative review of the state of the field.

## REFERENCES

Chomsky, N. (1959). A review of *B. F. Skinner's Verbal Behavior. Language*, 35(1), 26–58.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. McGraw-Hill.

Kuhn, T. S. (2012). *The Structure of Scientific Revolutions: 50th Anniversary Edition*. Chicago, IL: University of Chicago Press.

Newman, S. (1955). Review of Psycholinguistics: A Survey of Theory and Research Problems. *American Anthropologist*, 57(5), 1096–1098. doi: 10.1525/aa.1955.57.5.02a00530

Osgood, C. E., & Sebeok, T. A. (Eds.) (1954). *Psycholinguistics: A Survey of Theory and Research Problems*. *International Journal of American Linguistics 10*. Baltimore: Waverly Press.

# Part I Production

# 1 Overview

FERNANDA FERREIRA

University of California, Davis

As a person speaks, a great deal of processing activity is taking place behind the scenes. In the situation of some sort of dialogue or conversation, a speaker will grab one of the many ideas activated in working memory and commit it to the processes of linguistic formulation. This thought must then be translated into specific concepts, the words that express those concepts must be retrieved, the words must be organized into a structure that communicates the thought effectively, and the entire plan must be converted to a phonological representation (either in speech or sign) that will allow the utterance eventually to be articulated. At the same time, speakers must take into account the needs of their conversational partners. They also need to keep track of how the communication is unfolding: That is, they must consider the knowledge their interlocutors bring to the conversation, how their common ground is incrementally being built up as the interaction proceeds, and how effective their contributions are given the goals of the exchange. The chapters in this section touch on many of these important ideas, spelling out in detail how syntactic structures are generated, how redundancy and givenness are conveyed, how multiple language systems are coordinated, and how conversations are managed. These chapters make clear the enormous progress that has been made over the last 50 years in uncovering the architecture of the production system as well as the systems with which it interacts.

## Generating a syntactic structure

The syntactic level of representation takes center stage in research on language production. As Bock pointed out decades ago (Bock, 1982), it is at least coherent to ask whether people build a syntactic level of representation to mediate between form and meaning during comprehension, in part because, for many semantically

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

constrained sentences, a "bag of words" approach will be sufficient to establish the underlying event structure specifying who did what to whom. But when it comes to production, you can't fake syntax. If you attempt to speak a language with grammatical agreement, it will be painfully obvious when you err. If you do not understand the rules of word order, the result may be awkward at best and misleading at worst. For these and other reasons, theoretical models of production have always taken the problems of syntactic planning seriously. In speech error models of production like the ones proposed by Garrett and Fromkin (Fromkin, 1971; Garrett, 1975), the first stage of grammatical encoding (in Garrett's model, the functional level) is the entry point into the language system proper, as distinct from the general-purpose conceptual system that supports any type of perceptual-motor encoding. Syntax cannot be ignored in theories of language production, and research has proceeded accordingly.

Franck thus opens up this production section with a careful, analytical discussion of the syntactic system in general, and then moves quickly to what she treats as the model process for syntactic planning, which is agreement. As Franck notes, agreement is genuinely syntactic: It is a formal mechanism for linking words across often widely separated sentence positions. Franck begins with the studies by Bock and colleagues that initiated this research program, which used agreement errors as a source of information about the nature of syntactic representations. As Franck also points out, this early work seemed to reinforce Garrett's fundamental assumptions about the architecture of production, which mapped the modularity of representations onto strictly serial processing. Speakers were assumed to first plan the meaning of what they would say, then generate the appropriate syntactic forms, and then engage in phonological planning.

This early work was critical for fostering a discussion about the consequences of this architecture for people's ability to make syntactic decisions online, and as one hopes to see in a field of inquiry in which concepts are specified in enough detail to be falsifiable, some of the key assumptions did not survive tough empirical scrutiny. Franck summarizes the problematic results, many of which came from her own research on agreement, done in collaboration with her colleagues. These challenges to what Bock and colleagues referred to as the Marking and Morphing model motivated the development of Franck and colleagues' alternative Feature and Controller Selection model, which differs from the Marking and Morphing model in a number of key respects. Perhaps the most central is the separation of the stage responsible for selecting features relevant to agreement from the stage that identifies the relevant controller of agreement. Semantic, syntactic, and morphological features influence both stages, but in different ways. This model also tightly links syntactic structures and memory representations, since the selection of both features and controllers is strongly influenced by the availability of information in working memory. In addition, the model assumes that the more prominent the syntactic position of a word, the greater its accessibility in memory. Franck ends her chapter with a useful roadmap for future research on syntax in production.

But what does this model of agreement imply for the concept of incremental production, for example? In traditional models of language production, a key question concerned the planning units for any level of representation, including for syntax. Do speakers plan an entire clause before beginning to speak, or are planning and execution processes cascaded? In the early days of production research, many papers were published on this subject (e.g., Ford & Holmes, 1978), with evidence suggesting clausal planning units for grammatical encoding. Late twentieth-century models such as Levelt's (1993) moved away from this idea and toward incrementality with the suggestion that planning domains should be as small as possible, and perhaps no larger than a single word. But how do we reconcile this incremental approach with the facts concerning agreement, where, as Frank notes, the controller and the form with which it agrees could in principle be indefinitely separated (and in practice are often separated by several words)? In fact, in recent years, the pendulum has begun to swing back toward the view that planning units for syntax are probably at least phrasal (Allum & Wheeldon, 2007; Bock & Cutting, 1992; Ferreira, 2000; Martin, Crowther, Knight, Tamborello, & Yang, 2010), and that the size of those units are likely not architecturally determined, but instead vary depending on the goals of the speaker (Ferreira & Swets, 2002; Wagner, Jescheniak, & Schriefers, 2010). It would be interesting to know, then, how findings concerning agreement speak to this question of planning units for grammatical encoding in more detail.

## Distributing information

Typically, the same idea can be linguistically conveyed in more than one way, which presents the production system with both an opportunity and a set of processing decisions. This issue of flexibility in production connects to the previous discussion concerning how syntactic structures are generated, because one of the tasks of the production system is to make syntactic choices such as whether to produce an active or passive sentence, or whether to include an optional element such as the complementizer *that* in a sentence. In addition, speakers vary the way they pronounce the same word depending on features such as familiarity as well as predictability. The chapter by Jaeger and Buz focuses on the phenomenon of reduction, and probabilistic reduction in particular, which they link to contextual predictability. The general idea is that the more expected something is, the more reduced will be its pronunciation.

Jaeger and Buz then link the phenomenon of reduction to three general accounts of production: one that emphasizes ease of production, another that emphasizes the facilitation of communication, and one that emphasizes representational issues. The first account they consider assumes that reduction occurs because it facilitates the job of the speaker. The second account links reductions to speakers' attempts to make the task of the listener easier. And the third account attempts to connect phenomena of language change to online language production. Jaeger and Buz conclude by arguing that all three influences must play a role in explaining the robust, cross-linguistically attested tendency on the part of speakers to reduce predictable forms. As they point out at the end of their chapter, an important question that remains to be answered is precisely how these three approaches mesh with one another. Another open question concerns omission of linguistic forms, which is also related to predictability. For example, in a null subject language, the likelihood that a speaker will produce an overt pronoun or leave the position null depends on the predictability of the corresponding referent. Is omission simply the extreme case of reduction, or does the speaker face a binary choice in cases such as these? This would seem to be an important question for future work, as Jaeger and Buz note.

## Bilingualism, multilingualism, and signing

The majority of people live with more than one language system, and so a critical question for theories of language production concerns how these different databases of information are coordinated and managed. Paolieri, Morales, and Bajo's chapter examines this issue in detail. They present the problem as follows: How do speakers choose between two forms from different languages that express the same idea? One class of models of bilingual production assumes selective access, so that the decision to speak in a given language effectively shuts off any other languages the speaker might know. In contrast, in nonselective models, forms across languages interact, potentially leading to competition and interference. Linking back to Frank's interesting chapter, Paolieri *et al*. (this volume) go beyond standard evidence for lexical interference effects to highlight findings concerning grammatical gender interactions across languages as well. The so-called gender congruency effect is observed when words in different languages happen to belong to the same gender class. Negative transfer is even observed during production of a non-gendered language (e.g., English) when spoken by people who also know a gendered language (e.g., Spanish). Syntactic information is also thought to be explicitly marked according to whether forms are shared or not, a conclusion that emerges from research on cross-linguistic syntactic priming.

Of course, no discussion of bilingual language production would be complete without some consideration of the so-called bilingual advantage. Paolieri *et al*. present a balanced and up-to-date analysis of the evidence for and against the theory of bilingual production which postulates the need for inhibitory control, and which further assumes that the frequent exercise of cognitive control sharpens the non-language cognitive system overall. These ideas have recently received a fair bit of pushback in the literature, with some investigators highlighting concerns related to publication bias (de Bruin, Treccani, & Della Sala, 2015), and others claiming not to find any evidence that bilinguals indeed reliably show any cognitive advantages (e.g., Paap & Greenberg, 2013). Paolieri *et al*. do the field a great service by providing a nuanced perspective, suggesting that immersion and language experience play a role in determining how selection operates in an individual, which in turn has implications for the extent to which any bilingual advantage will be observed.

Signing is another domain in which issues relating to production involving multiple languages arise. This is because, as Wilbur points out in his chapter, most people who communicate in sign know sign as a second language. Many of the challenges for people producing sign languages are similar to those that have been identified for spoken languages, in part because the two kinds of languages have many similarities. Wilbur notes that the prosody of sign is based on prosodic constituents that are ordered hierarchically, starting with the smallest unit, the mora, and topping out with the intonational phrase and the phonological utterance. This organization is just like what is observed for spoken languages. One prosodic domain that has been extensively studied is the production of syllables in sign, with research suggesting that although both spoken and signed languages have syllables as prosodic constituents, their internal structures differ due to the differing modes of transmission in the two modalities. In addition, whereas English permits sentence stress to be marked on any constituent within a sentence, American Sign Language (ASL) is similar to spoken languages like Italian in that sign permits only sentence-final stress. Thus, one consideration for ASL speakers generating a syntactic form is to decide how to organize the sentence so this prosodic constraint can be respected while at the same time conveying the intended semantic focus within a grammatical form. Signers also generate speech errors similar to those found in spoken languages, including word substitutions and errors involving phonetic features. Wilbur's chapter ends with a discussion of how speech and sign are coordinated in individuals who attempt to communicate in both modalities simultaneously. Contrary to what might seem intuitive, it appears that the simultaneous production of a sign and a spoken expression is interfering (similar to what is observed for multiple spoken languages, as Paolieri *et al*. argue), leading to disruptions in the production of both types of linguistic forms, as would be expected from any attempt to communicate two spoken languages at the same time. One interesting advantage of sign is that signers have the ability to communicate more than one concept simultaneously—for example, two referents can be conveyed, one with each hand. A fascinating question for psycholinguistic investigation is to determine how this information is represented and executed in sign compared with speech, and to conduct experiments to discover how comprehenders efficiently process such information.

## Linking production and comprehension

Of course, the production system does not operate in a psychological vacuum: It works with other cognitive systems, including those responsible for perception, attention, and memory. Production processes also interact with those responsible for comprehension, and vice versa. The two systems influence each other. The two chapters in this section, one by Pardo and the other by Gambi and Pickering, discuss ideas for capturing these relationships, as well as the empirical evidence concerning the details of these mutual effects. The fundamental conclusion that emerges from both chapters is that the demands of communication helped to shape the structure of language, which in turn influence the online processes that allow speakers to efficiently generate utterances that are communicatively effective.

Pardo focuses on speaker-addressee interactions, noting the large body of research showing how speakers tailor linguistic forms to suit the addressee. She summarizes studies demonstrating coordination, entrainment, alignment, and accommodation between interlocutors. At the same time, divergence is also observed, particularly when the conversational participants are of different status or differ from each other on other traits tightly bound up with social identity. Moreover, individuals differ in their tendency to adapt in this way to their interlocutors. Pardo further makes the case that these effects challenge traditional approaches to psycholinguistics that distinguish competence from performance, and those that treat language as a system that is primarily for the transmission of information. This argument is not new; it appears that many researchers investigating these kinds of topics believe their conclusions and even the entire approach is incompatible with, say, a formal analysis of grammatical encoding. But this claim seems to me to be somewhat exaggerated. Consider, for example, Frank's chapter on the computation of agreement during production. Is any mechanism or process proposed in that chapter inconsistent with the notion that speakers would tailor their utterances so they're appropriate given their addressees? The answer, it seems to me, is no; it's more a matter of whether an important topic—the tailoring of utterances to addressees—receives attention or is neglected. Pardo is certainly right to emphasize the importance of processes promoting alignment between interlocutors, and it is also clearly true that the field had for too long ignored the kinds of questions her chapter brings to the fore. Both kinds of inquiries can co-exist, and indeed must co-exist, if we are to emerge with a complete theory of the language production system.

These ideas are further delineated in the chapter by Gambi and Pickering, which focuses on models linking production and comprehension. One of their original suggestions is for the field to redefine what it means for something to be a production or a comprehension process. As they note, the traditional approach is to assume that whatever happens during production is a production process, and whatever happens during comprehension is a comprehension process. On this view, production permits feedback to the extent that we observe "lower-level" processes influencing those that originate from higher representational levels. For example, if phonological information affects choice of syntactic form during production, that is an example of feedback, and the existence of such effects motivates non-modular models. The same logic holds for comprehension, except that the interactive effects are ones in which higher levels influence lower ones (e.g., a semantic effect on syntactic parsing decisions). Gambi and Pickering's suggestion is to abandon this approach and instead to define production processes as those that map higher-level representations onto lower-level ones, and to define comprehension as processes that do the opposite. On this view, then, the production and comprehension systems interact with each other, but the production and comprehension systems themselves are not interactive. For example, self-monitoring, the ability of speakers to evaluate the quality of their utterance before overtly producing them, is a process that takes place during the production of an utterance but which involves looping the comprehension system in at a specific point during planning. The chapter includes a summary of Pickering and Garrod's self-monitoring theory, which provides a specific example of this approach. Their theory also captures the phenomenon of prediction during comprehension as another example of how the production and comprehension systems work interactively (and also imply that prediction effects are not evidence for interactive comprehension systems). If as a listener I am able to anticipate your next word, it is because I have invoked my production system to model what I would say in that specific context. This proposal is consistent with evidence suggesting that an individual's production skills correlate with that person's ability to predict effectively during comprehension.

But perhaps the most well known contribution these models make, as Gambi and Pickering argue in their chapter, concerns the insights they provide about the fundamental nature of dialogue. As many researchers studying language production have argued, the standard psycholinguistic model that treats production and comprehension as separate systems makes dialogue somewhat of a mystery. Indeed, many researchers who focus exclusively on production have argued that production is hard, but that is not the intuition most of us have when we talk to someone—instead, our sense is that production is pretty easy, and we sometimes marvel at the way our ideas flow out as speech without our awareness of the unfolding processes and without the need for conscious planning. Indeed, Churchland (2013) in her recent book *Touching a Nerve: Our Brains, Our Selves* describes this phenomenon very compellingly, based on personal experience. She notes that not only is production usually quite easy, requiring little conscious planning, it is often precisely when we become conscious of how we are talking that we find ourselves struggling, and in these circumstances we often become disfluent as well as communicatively ineffective. Thus, twenty-first century psycholinguistic theory must explain what makes dialogue easy, at least most of the time. The answer that models like Garrod and Pickering's provide is that it is based on rapid coordination between the production and comprehension systems, which in turn is likely grounded in humans' ability to generate recursive models of other minds and intelligent agents.

## Themes, resolutions, and challenges

As I hope this overview makes clear, the chapters in this section on language production lay out some exciting, important new perspectives. At the same time, notably absent from this section is any chapter discussing the processes that support the generation of prosodic forms during production. The rich interplay among semantic, syntactic, and prosodic sources of information is not addressed in these discussions, which is unfortunate. The fault, however, is not with the editors of the volume but rather the lack of interest in the topic in the field more generally. In my own view as someone who has worked on this issue, the questions and perhaps also the answers are simply not provocative enough for a field that has far too often exaggerated theoretical distinctions as a way of generating controversy. But if there is one basic fact about language production, it's that speakers generate a prosodic form each and every time they utter even a single syllable. They mark syntactic and semantic structure, and they even mark discourse constituency using prosodic features such as pitch and intensity (Tyler, 2013, 2014). This issue, then, should be the target of active investigation.

Another area in which the field has not made enough progress, in my view, is in developing clear, specific theories about exactly how referential forms are chosen. There is research demonstrating that people take into account the needs of their listeners, respond to immediate feedback, and so on, and much of that research is summarized in the chapters in this section. But what exactly are the mechanisms that support these abilities? On this question, we have little information. Another example: Decades ago, Levelt (1982) conducted a series of clever experiments using simple figures consisting of connected colored circles to assess speakers' ability to make macro decisions about how to structure a discourse, and how they keep track of what had been said and what still needs to be communicated. Levelt gave the example of describing a house or an apartment, which requires the speaker to decide where to begin and how to proceed when there is a conceptual choice point (e.g., when a hallway splits off into two wings). The speaker also must keep track of what has been described and what has not. This work showed that speakers attempt to minimize their memory load, beginning with the discourse segment that is shorter and less complex. This strategy enables them to plan the longer, more complex segment during articulation, as we showed in our own work following up Levelt's (Ferreira & Henderson, 1998). We also suggested that both speaker and listener benefit from this strategy because both need to use working memory resources as efficiently as possible. Unfortunately, we still know about as much concerning discourse planning today as we did 30 years ago, suggesting that the topic is under-investigated.

Continuing with this theme, it appears that we have made a great deal of progress in understanding production since the days of the speech error models developed by Garrett and Fromkin in the 1970s. Our experimental methods permit us to isolate specific bits of the production process and determine the factors that influence it, whether the process is computing agreement, managing multiple languages, or coordinating a conversation. The availability of huge corpora makes it possible to conduct large-scale data analyses of very specific phenomena, including things like phonological reduction. But perhaps we are now missing something that those global speech error models gave us, and that was a road map for the entire system. A researcher might focus on segmental speech errors, but given the constant backdrop of the global models, discussions would ultimately come back around to the big questions concerning the architecture of the production system itself. Now, the connections from specific empirical phenomena to global models of production are somewhat less clear. For example, how does alignment promote communication and influence phonological reduction? More importantly, as researchers in this area, what do we now believe about the overall structure of the language production system, from discourse planning all the way down to segment retrieval and articulation? Some might say that the original speech error models no longer hold up given findings from controlled experiments or corpus analyses, which is a reasonable point. However, it would be useful if today's researchers would try to come up with some sort of alternative models that have the same scope and ambition as the ones our pioneers gave us decades ago. Psychology is already known for being a field that sometimes seems too focused on techniques and effects at the expense of theories and mechanisms. Psycholinguistics should be an exception given its rich theoretical history.

At the same time, the study of language production has clearly advanced in many significant ways. Investigations of topics such as multilingualism, signing, and comprehension-production interactions are genuinely novel and exciting. Much more is known also about more traditional issues such as syntactic planning. There is no doubt that there has been a major increase in methodological and statistical rigor. My plea is simply that we not turn our backs on the twentieth century as our field continues to progress, but that we build on previous insights and ideas as we continue to investigate production in the twenty-first century and beyond.

## REFERENCES


Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: the role of grammatical units. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33(4), 791–810. doi: 10.1037/0278-7393.33.4.791

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89.

Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31(1), 99–127. doi: 10.1016/0749-596X(92)90007-K

Churchland, P. S. (2013). *Touching a nerve: Our brains, our selves*. New York, NY: WW Norton & Company.

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, 26(1), 99–107. doi: 10.1177/0956797614557866

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. Wheeldon, L. (Ed.) *Aspects of language production* (pp. 291–330). East Sussex, UK: Psychology Press.

Ferreira, F., & Henderson, J. M. (1998). Linearization strategies during language production. *Memory & Cognition*, 26(1), 88–96. doi: 10.3758/BF03211372

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46(1), 57–84. doi: 10.1006/jmla.2001.2797

Ford, M., & Holmes, V. M. (1978). Planning units and syntax in sentence production. *Cognition*, 6(1), 35–53. doi: 10.1016/0010-0277(78)90008-2

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47(1), 27. doi: 10.2307/412187

Garrett, M. F. (1975). The analysis of sentence production. *Psychology of Learning and Motivation*, 9, 133–177. doi: 10.1016/S0079-7421(08)60270-4

Levelt, W. J. (1982). Linearization in describing spatial networks. In *Processes, beliefs, and questions* (pp. 199–220). Springer Netherlands.

Levelt, W. J. (1993). *Speaking: From intention to articulation* (Vol. 1). MIT press.

Martin, R. C., Crowther, J. E., Knight, M., Tamborello, F. P., & Yang, C. L. (2010). Planning in sentence production: Evidence for the phrase as a default planning scope. *Cognition*, 116(2), 177–192. doi: 10.1016/j.cognition.2010.04.010

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232–258. doi: 10.1016/j.cogpsych.2012.12.002

Tyler, J. (2013). Prosodic correlates of discourse boundaries and hierarchy in discourse production. *Lingua*, 133, 101–126. doi: 10.1016/j.lingua.2013.04.005

Tyler, J. (2014). Prosody and the interpretation of hierarchically ambiguous discourse. *Discourse Processes*, 51(8), 656–687. doi: 10.1080/0163853X.2013.875866

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(2), 423–440. doi: 10.1037/a0018619

# 2 Syntactic Encoding: Novel Insights Into the Relationship Between Grammar and Processing

JULIE FRANCK

University of Geneva

## Introduction

When Nim, the chimpanzee raised at Columbia University, produced one of his longest sentences: *Give orange me give eat orange me eat orange give me eat orange give me you*, he succeeded in conveying his desire to be given an orange and to eat it. However, his sentence could, in principle, have had several other meanings: for example, that he was eating the orange that was given to him by his interlocutor, that he wanted an orange to give to his interlocutor, or that his interlocutor has given him an orange that is now eating him. These alternative interpretations can reasonably be ruled out by our knowledge of the world and the situation (oranges don't eat monkeys). Nevertheless, they all seem equally compatible with the concatenation of words that Nim produced. Syntax is the component of language that allows us to express ideas with as little indeterminacy as possible, by constraining dependencies between words in the sentences. It does so by two major devices: constraints on the ordering of the words and grammatical morphemes linking words together. Nim was able to develop an idea, learn signs to express the various pieces of that idea, and concatenate them to convey it within a (rudimentary) social interaction. This is an amazing achievement, demonstrating that he was able to communicate. But Nim never managed to develop syntactic knowledge, necessary to express his ideas with the precision that human language users routinely exhibit. Systematic failures from the various programs dedicated to teaching syntax to animals stand in sharp contrast with the ease with which humans deal

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

with the syntax of their sentences. Within three years, little humans master most of the syntax of their mother tongue. Children spontaneously create syntactic devices when the input language is impoverished (like pidgins or first generations of newly created sign languages; Bickerton, 1984; Senghas & Coppola, 2001; Singleton & Newport, 2004). Adults produce sentences almost flawlessly, with less than 1 error every 1000 words (Levelt, 1989), and the great majority of these errors preserve the syntactic well-formedness of the sentence.

Despite broad agreement that grammar is what makes human language different from other communication systems (Hauser, Chomsky, & Fitch, 2002) and despite its intriguingly powerful machinery that by far surpasses that of lexical processes (most errors in production are lexical errors), psycholinguistic research on language production has for the most part focused on single words. The reason is no doubt the challenge of grasping the abstract, relational nature of syntactic structures. Psycholinguists interested in syntactic encoding face two major questions. First, how can we characterize the syntactic representations underlying the sentences that speakers build? This question has to do with the shape of these representations, and whether they look like the formal hierarchical structures proposed in syntactic theory. Second, how can we characterize the processes that deal with these representations? This question relates to the identification of the functional components involved in syntactic encoding, and the relationships between them, that is, relations between lexical and syntactic processes, and between syntactic and non-syntactic levels of representation.

This chapter addresses these two major questions through the lens of studies of agreement production. Agreement is a syntactic phenomenon par excellence: it ties words together in virtue of their syntactic status, for the most part independently of their semantic or morpho-phonological content. The majority of natural languages have agreement constraints, which may involve features like number, gender, and case. Through several properties, agreement offers an especially revealing perspective on our two questions. The first property is that agreement is structure-dependent. In many languages the verb agrees in number with its subject and sometimes with its object, but there is not a single language in which the verb agrees with, for example, a noun inside a clausal modifier (e.g., **The goat that ate the radishes are mean*). That is, not any element in the sentence may enter into an agreement dependency. The second property of agreement is that it shows, to some extent, autonomy from the semantic representation of the words that constitute sentences. We are perfectly able to produce correct agreement in a sentence that does not make sense like *The subversions of the boy are genuine* or even in a Jabberwocky sentence with senseless words like *Which rabun did you say that the livols are eating?* The autonomy of agreement with respect to other levels of representation makes it possible to investigate the key question of the modularity of syntactic encoding. Moreover, agreement is at the crossroads of lexical processes (features are retrieved from the lexicon) and syntactic processes (features are transmitted to the targets in virtue of their structural positions), allowing us to explore the relations between lexical and syntactic processes involved in syntactic encoding.

Agreeing units can be contiguous in the sentence, or far apart, in the same clause or in separate clauses; yet, speakers as early as age three (Keeney & Wolfe, 1972) are able to produce multiple agreement dependencies within a single second, without effort, and usually without errors (younger children tend to produce singular verbs as "default," that is, singular verbs in the context of a plural subject, but they do not produce erroneous plural verbs in the context of a singular verb, which shows that they do not produce verbal agreement morphology at random; Clark, 1998). Nevertheless, agreement computation sometimes fails. Following the long tradition of spontaneous speech error research, Bock and Miller (1991) initiated a fruitful line of experimental work on a particular kind of errors called *attraction errors*. Attraction refers to the erroneous agreement with an element that is not the agreement controller, as in **The key to the cabinets are rusty*, where the verb *are* incorrectly exhibits plural number marking as though it were agreeing with the plural modifier, *cabinets*, rather than with the singular head *key*. Identifying the conditions that modulate attraction errors has turned out to be an extremely rich way to explore both the shape of syntactic representations and the processes that build them. In about two decades of experimental research, an extensive body of evidence has accumulated showing influences from semantic, syntactic, morphological, and morphophonological properties of the words and sentences. The prominent *Marking and Morphing* model (M&M) of agreement attraction (Bock, Eberhard, Cutting, Meyer, & Schriefers, 2001, further developed in Eberhard, Cutting, & Bock, 2005) is founded on this empirical work. It employs the core features of Garrett's seminal model of language production with a modular, two-level functional architecture in which semantic information penetrates the first level only, while the second level deals with structural and morphological information (Garrett, 1975, 1980, 1989). As such, it provides a representative illustration of what has become the standard psycholinguistic approach to syntactic encoding, and it allows us to see both the strengths and weaknesses of its assumptions.

The aim of this chapter is to establish a new pathway to the analysis of syntactic representations and processes involved in agreement production. The framework, *Feature and Controller Selection*, takes insights from both syntactic theory and the psychological theory of memory retrieval in order to capture the wide empirical range of attraction effects that have been observed. Like *Marking and Morphing*, the model cuts the pie into two parts, but in a fundamentally different way. The first process, *Feature selection*, retrieves functional units from the long-term memory lexical store. It is responsible for selecting the grammatical features associated with the nouns in the sentence. This process crystallizes several of the semantic, morphological, and morphophonological influences reported in research on agreement errors. It operates under the guidance of multiple, statistically distributed cues, in line with interactive models of production (Dell, 1986; Goldrick, 2006). Interestingly, most of the data points from the attraction literature focus on this fundamentally lexical component. Even though these effects are manifest in the context of an attractor element, I argue here that they arise on top of attraction, independently of it. In that regard, they only represent the cherry on the pie. The pie, that is, the core syntactic component of agreement, is the second process, *Controller selection*. This process is responsible for retrieving the controller in the sentence in order to copy its features onto the target. It is the locus of what I consider to be "attraction proper," conceived of as the incorrect identification of the attractor as the controller, in line with the initial suggestion of Badecker and Kuminiak (2007). I suggest that by positing a process of controller selection, and properly distinguishing it from feature selection, we can unite a range of syntactic effects showing the sensitivity of attraction to major syntactic constructs (intervention, movement, c-command, hierarchical depth), as well as another set of semantic and morphological effects, different from those arising during feature selection, and that appear to lie in the similarity between the attractor and the controller. The syntactic, semantic, and morphological effects arising at this level are argued to reveal the inner workings of a cue-based retrieval process (Lewis & Vasishth, 2005) in which these factors act as cues to retrieve the controller.

In the first section of the chapter, I sketch the standard model of syntactic encoding in psycholinguistics (Garrett, 1975) and the M&M model of agreement developed by Bock and her colleagues, which instantiates one of the most elaborate illustrations of the standard approach. In the second section, I lay out a typology of attraction effects, organized in terms of my proposed *Feature and Controller Selection* account. In the third section, I present new evidence showing a close alignment between syntactic structure and memory retrieval in attraction effects, and suggest that syntactic theory describes the strength of memory representations. The chapter ends with some challenging issues for future research.

## Syntactic encoding in the M&M of agreement

The approach of sentence production initiated by Garrett in the 1970s assumes that the syntax of the sentence is encoded at two separate, successive levels, each of them responsible for a set of lexical and syntactic processes (Garrett, 1975, 1980, 1989). The Functional level, which executes first, ensures the retrieval of words and the construction of a functional, hierarchical structure specifying the words' syntactic role (like subject, object) and the relations between them. The Positional level, which executes second, is responsible for retrieving word forms and inserting them within a frame of grammatical morphemes (like determiners, inflections) arranged in the linear order in which words will be pronounced. The major evidence in favor of this model comes from the observation that speech errors are distributed in two broad classes. The first class involves whole words (e.g., *She sings everything she writes*; Fromkin, 1971). In the great majority of these errors, the words exchanged are of the same grammatical category, they show no phonological similarity, and they may be part of a rather large unit (e.g., the clause). The second class of errors involves units smaller than words. Most of these errors involve phonemes (e.g., *plit spea soup*; Fromkin, 1971). They do not respect the grammatical category of the words; rather, they respect phonological and prosodic constraints (like syllable position or the vocalic versus consonantal status of the sounds). A smaller subset of these sublexical errors, so-called stranding errors, involve the exchange of two lexical roots while the closed class morphology of the words is left in place (e.g., *That's why they sell the cheaps drink*; Garrett, 1989). Like phonological errors, stranding errors appear to follow sound-level constraints; they involve words of different grammatical categories, lexical roots are often phonologically similar, and they usually arise within smaller units (phrases). Garrett suggested that the first class of errors arises during the operation of the Functional level, when words are retrieved and assigned a syntactic function, while the second class arises during the operation of the Positional level, when words are specified both morphologically and phonologically. In Garrett's view, the production system is fundamentally sequential: the Functional level takes as input semantic information from the Message, which provides the content the speaker is wishing to convey. The Positional level takes the output of the Functional level as its input, and it sends its output to the articulatory system responsible for encoding the surface phonetic forms of phrases. The strict seriality of this framework is responsible for its modularity: semantic information only penetrates the Functional level while the Positional level is immune to direct conceptual influences (level n-1 influences level n, not level n + 1), and the Functional level is immune to the morphological and phonological specifications that are only specified subsequently at the Positional level (level n + 1 cannot influence level n).

Nearly four decades of experimental research has not challenged the major claims of Garrett's account: the separation of the system in two functional levels, and their serial order with the resulting modular architecture (although see for example, Dell, 1986, for an alternative, interactive approach). I now consider how these two properties are implemented in the M&M model of subject-verb agreement production (Bock *et al*., 2001; Eberhard *et al*., 2005). The description summarizes the major assumptions of the model and the types of data it explains, and then highlights the challenges the model is confronted with.

## *Two functional levels: Marking and Morphing*

The model involves two functional components. Marking takes place in the syntactic component of the first level of Functional assembly. It is responsible for translating the number notion from the Message into a linguistic feature. At the level of the message, a process of notional number valuation takes place by which notional singulars are distinguished from notional plurals in the speaker's reference model. Although notional number is continuous in that entities can be conceived as more or less single or multiple, Marking receives the output of number valuation and translates it into the selection of a syntactic feature (singular or plural). The site of Marking is not the subject head but the root of the whole subject phrase. At the same level, lexical processes also take place. These are responsible for recruiting nouns with meanings that are consistent with the notional number of the corresponding concepts. For example, *clothing* and *clothes* equally express a notion of multiplicity, and are therefore equivalent options for the number Marking process: their grammatical properties differ, but these grammatical properties only come into play during the next stage, Morphing. Morphing takes place at the level of Structural integration, which binds together lexical forms (morphemes), and structural forms (the hierarchical representation of the sentence). Morphing reconciles the syntactic features selected during Marking and number specifications from the lexicon, which are argued to percolate the tree up to the subject root, where reconciliation takes place. Morphing also ensures that the feature selected by the reconciliation process will be transmitted to the agreement target. The model adopts two other assumptions. One assumption is that single count nouns are unspecified, or weakly specified, for number. As a result, only plural nouns have the possibility to percolate the tree and enter into the reconciliation process. The other assumption is that if an inconsistency is encountered between number marking and morpheme specifications (e.g., in collectives, like *army*), the morpheme specifications prevail.

## *Serial order: Marking before Morphing*

The model adopts the standard assumption of seriality, such that Marking takes place before Morphing. This architecture has two consequences for the information flow in the system. The first consequence is that Morphing is insensitive to semantic information from the message: semantic information only penetrates the first stage of Marking. The second consequence is that Marking is insensitive to morphological information, which is specified after its output has been sent to Morphing.

This architecture accounts for three major sets of facts: some semantic effects on agreement, the asymmetry between singular and plural attractors, and the profile of differences and similarities between verb and pronoun agreement. The semantic effects considered for the model have their locus in the Marking component. Notionally plural but grammatically singular subjects (e.g., a collective noun like *army* or a distributed subject like *The label on the bottles*) trigger the selection of a plural marking. However, they carry a singular morpheme specification. These two features have to be reconciled during Morphing, and the stronger power of morphological specification is such that the plural marking feature will be overridden most of the time by the more powerful singular morphological specification. Still, on some occasions, the plural feature will win, giving rise to plural agreement on the verb. Attraction results from the contamination from the attractor's feature of the subject node during the process of reconciling number marking and number specification. Morpheme specifications anywhere in the structure have the potential to percolate, but their influence depends on the structural proximity of the attractor to the subject's maximal node, that is, the locus of agreement control. In this way, the model correctly predicts that the head noun's feature will usually be the controller of agreement, given its privileged structural position, closest to the maximal node, involving the smaller percolation path. The finding that most attraction effects, at least in English, arise from a plural attractor while virtually no attraction arises from singular attractors results from the lack of (or weak) morpheme specification of singulars: only plural, morphologically specified features have the potential to percolate and therefore influence the reconciliation process. The model accounts for the finding that pronouns are more sensitive than verbs to the notional number of the subject (Bock, Eberhard, & Cutting, 2004) by the fact that pronouns receive their number from the semantics, through Marking, whereas verbs' number is assumed to be semantically empty, such that they can only receive their agreement feature through Morphing. The finding, in contrast, that pronouns and verbs are equally sensitive to the presence of a plural attractor, and equally insensitive to the notional plurality of the attractor, is explained by the fact that attraction takes place at a stage that is common to both pronoun and verb agreement, which is sensitive to morpheme specifications but insensitive to notional representations: the stage of Morphing.

Although M&M has the merit of accounting for a wide range of data points reported in the literature, the model is challenged by two major issues. The first issue concerns the locus of semantic influences. In M&M, semantics only affects agreement by way of its influence on the selection of the agreement feature for the whole subject phrase via Marking. This property of the model fails to account for the finding that semantic information that has no consequence for the whole subject noun phrase also influences agreement. For example, more attraction is found when the subject head noun and the attractor noun semantically overlap (Barker, Nicol, & Garrett, 2001) or when the attractor is a plausible semantic subject (Hupet, Fayol & Schelstraete, 1998; Thornton & MacDonald, 2003). The model does not capture either effects found in gender agreement like the finding that heads with semantic gender are more resistant to attraction than those with grammatical gender (Vigliocco & Franck, 1999, 2001). This shows that notional gender information that is tied to the head itself, but critically not to the subject phrase, also affects agreement. Hence, semantic influences are found outside of the realm of the Marking process, which can therefore not be considered as the sole locus of semantic influences on agreement. In the alternative model proposed in the next section, I suggest decomposing semantic effects in three types: effects due to the notional valuation of the subject phrase (in line with Marking), effects due to the feature stability of the nouns (features with semantic correlates are more stable), and effects due to the semantic similarity between the attractor and the head. Whereas the first two effects are argued to have their loci in one component of agreement production (Feature selection), the latter has its locus in another component (Controller selection).

The second issue concerns the underspecification of the structural conditions over which Morphing takes place (Franck, 2011). In M&M, attraction is a function of the structural distance between the attractor and the subject node. Despite the critical role that structural distance is assumed to play, the theory does not provide a description of what the structure looks like, let alone a tool to measure structural distance. In contrast to Garrett's early model which represented hierarchical structure at the first level of encoding (Garrett, 1989), that is, at the level of Functional assembly in M&M, current models assume that hierarchical structure is built at the second level of encoding, that is, during Structural integration in M&M (e.g., Bock & Levelt, 1994; Bock & Ferreira, 2014; Eberhard *et al.*, 2005). In these models, syntactic units are directly assigned to their surface hierarchical position. This position is based on evidence by Bock and colleagues (Bock, Loebell, & Morey, 1992) who argued that the hierarchical structure of a passive sentence is constructed with the patient in the subject position right away, without transiting through a deep hierarchical structure in which it occupies the position of complement of the verb, as assumed in movement-based linguistic theory (Chomsky, 1981). On this view, the speaker builds a single hierarchical structure, at the same stage as the linear structure, while the representation of the first stage is reduced to the "flagging" of the units for a particular syntactic function, assigned on the basis of the message. The consequence for the analysis of attraction patterns is that if attraction occurs at the second stage of Structural integration, it is predicted to occur in the same way for structures that have identical surface hierarchical structures, and differently for structures that have different surface hierarchical structures. Empirical evidence casts doubt on that prediction. For example, significant attraction is found from a plural subject modifier in the interrogative structure (e.g., **Are the helicopter for the flights safe?;* Vigliocco & Nicol, 1998), whereas no attraction is found in the superficially identical structure in Italian involving free inversion (e.g., *Telefonera l'amica dei vicini, Will-phone-Sg the friend of the neighbors*, Franck *et al.*, 2006). Similarly, the moved object of the target verb triggers attraction (*patients* in *Jean parle aux patientes que le médicament guérissent, John speaks to the patients that the medicine cure-Pl*), while the object of the main verb situated in the same surface position does not (*patients* in *Jean dit aux patientes que le médicament guérit, John tells the patients that the medicine cures-Sg*, Franck *et al.*, 2010; Franck, Colonna, & Rizzi, 2015). That is, two structures that are superficially identical but have different underlying structures show different attraction profiles. Moreover, two structures that are superficially different but have identical underlying hierarchical structure show similar attraction profiles. For example, the English interrogative structure generates similar attraction to the corresponding declarative (**The helicopter for the flights are safe*), despite their different surface structures. These data suggest that it is not the properties of surface hierarchical structures that account for attraction profiles but rather properties of their underlying hierarchical organization.

In the next section, I sketch the alternative model of Feature and Controller selection that makes use of the fine constructs from linguistic theory to describe underlying hierarchical structures and capture these syntax-based attraction patterns. The model aims at capturing a wide range of attraction effects and opens new windows to the understanding of syntactic encoding and its relation to both syntactic theory and the theory of memory processes.

## Syntactic encoding in the Feature and Controller selection model of agreement

The model involves two functional processes: Feature selection, which retrieves nominal features, and Controller selection, which selects the agreement controller that will transmit its features onto the target. In this model, attraction does not arise because of the incorrect percolation of a feature into the tree, but because the attractor has incorrectly been selected as controller and as a result its features have incorrectly been transmitted on the agreement target. The model is an extension of the Selection and Copy model proposed in our previous work (Franck, Vigliocco, Antón-Méndez, Collina, & Frauenfelder, 2008; Franck, 2011). I suggest that two classes of factors influence Feature selection: factors that modulate the lexical stability of the head and attractor features (depending on the strength of their association to semantic, morphological and morphophonological correlates) and factors that modulate the notional representation of the subject (lying in properties of the head or its relation to the attractor). Following Badecker and Kuminiak (2007), Controller selection is argued to be a retrieval mechanism operating on the basis of cues. It is modulated by the similarity between the attractor and the controller at the semantic, morphological and syntactic levels.

In the following section, the various factors that have been shown to modulate attraction are organized into classes and subclasses, in the manner of a typological classification. The highest classes are those defined by the two functional processes, Feature Selection and Controller Selection. The subclasses are defined by the nature of the factors they involve. Note that I only consider here studies manipulating agreement in the context of an attractor word. A growing set of empirical evidence showing grammatical modulations of agreement production in the context of sentences with no (or no clear) attractor should ultimately be incorporated into the picture, and the question of their relation to attraction should be discussed (like constructions involving conjunctions, disjunctions, pseudo-partitives, or quantified noun phrases, e.g., Haskell & MacDonald, 2005; Haskell, Thornton & MacDonald, 2010; Mirković & MacDonald, 2013; Marušič, Nevins & Badecker, 2001; Marušič, Nevins & Saksida, 2007; Smith, Franck & Tabor, 2016).

## Effects on Feature selection

Feature selection is the process responsible for selecting the grammatical features of the nouns in a sentence. It is fundamentally a process of lexical retrieval by which nominal features are selected from the functional lexicon. The process shows the property of interactivity widely reported in the literature on lexical retrieval of content words, in that it is influenced by semantic and form information (e.g., Dell, 1986; Goldrick, 2006). Two types of factors affect feature selection. The first factor is *feature stability*. If a grammatical feature is regularly associated with converging semantic and/or form (morphological or morphophonological) information, it is more stable and has more chance to be selected. As a result, converging correlates of the head's feature back up its grammatical feature and thus reduce the risk of an agreement error. In contrast, if these correlates diverge from the grammatical feature, that is, point to the opposite direction, the feature is less stable and more susceptible to a selection error. Feature stability similarly influences the selection of the attractor's feature; however, its effect on agreement is diluted given that it only shows up if the attractor is incorrectly selected as controller. The second factor influencing feature selection is the *notional representation of the subject phrase* at the message level. The conceptual representation of the numerosity of the phrase depends on factors like the distributivity of the subject, the semantic integration between the head and attractor nouns or the spatial distribution of units, which all have the potential to influence the number feature that will eventually be selected.

The key difference with M&M is that whereas M&M attributes semantic and formal effects to two separate functional components of agreement, respectively Marking and Morphing, the semantic and form effects grouped in this first category all arise at the level of the same functional component of Feature selection. Moreover, whereas in M&M form effects are intrinsically linked to attraction, the current model assumes that the semantic and form effects affecting Feature selection are independent of attraction proper (by "attraction proper," I mean the erroneous selection of the agreement controller), although their influence is sometimes only detectable in the context of an attractor noun, either because the attractor directly modulates the notional representation of the subject phrase or because it indirectly boosts error rates, allowing for these factors to show up. In other words, the effects of the factors listed here arise on top of attraction.

## Effects of feature stability

### *Semantic stability*

Semantic correlates boost agreement errors when they diverge from the grammatical feature. For example, more plural verb agreement is found with grammatically singular collective heads denoting a plural entity (e.g., *The cast in the weekend performances*) than with notionally singular nouns (*The actor in the weekend performances*, e.g., Bock, Nicol, & Cutting, 1999; Bock *et al.*, 2004; Haskell & MacDonald, 2003). Similar effects are found for gender agreement with epicene nouns that have a fixed grammatical gender but can refer either to a feminine or to a masculine entity. Speakers produce more erroneous masculine agreement on the predicative adjective when a grammatically feminine epicene head (e.g., *La victime*, The victim-F) refers to a man than when it refers to a woman (Vigliocco & Franck, 2001). In contrast, semantic correlates reduce agreement errors when they provide converging information to the grammatical feature. For example, head nouns with semantic gender (e.g., *La jument*, The mare-F referring to a female horse) give rise to fewer agreement errors than those with a purely grammatical gender feature (e.g., *La méduse*, The jellyfish-F; Vigliocco & Franck, 1999). Heads with regular plurals (like *bubbles*) generate fewer erroneous singular verbs than invariant plurals (like *suds*) (Middleton & Bock, 2004). Regular plurals are judged conceptually more plural than invariant plurals, suggesting that the presence of a clear semantic correlate of plurality backs up the grammatical feature (Haskell & MacDonald, 2003). Finally, the observation that gender attraction tends to be weaker than number attraction (e.g., Eberhard *et al.*, 2005; Lorimor *et al.*, 2008) may also be related to the fact that grammatical gender lacks semantic correlates. When semantic correlates are manipulated on the attractor nouns, the same factors turn out to have a much weaker effect (e.g., Bock *et al.*, 2004; Deutsch & Dank, 2009; Haskell & MacDonald, 2003); the reason is that for these effects to show up, the attractor first needs to be incorrectly selected as controller, which only arises in a small portion of the sentences produced.

## *Morphophonological stability*

Several studies across various languages (Italian, Spanish, French, Dutch, German) have shown that the strength of the association between the controller's feature and its morphophonological realization modulates agreement production. Heads with nominal endings or determiners that carry converging morphophonological cues (e.g., nouns ending in *–o* in Italian, which are usually masculine) are less prone to attraction than heads lacking these cues (nouns ending in *–e* in Italian, which can be masculine or feminine, see also Hartsuiker, Schriefers, Bock, & Kikstra, 2003; Vigliocco *et al.*, 1995). Heads carrying morphophonological information diverging from the grammatical feature are particularly sensitive to attraction (e.g., a masculine noun ending in *–a* in Spanish, although most nouns ending in *–a* are feminine, Franck *et al.*, 2008). Again, when manipulated on the attractor noun, the same factors typically show either a weaker effect on agreement, or no effect at all (e.g., Bock *et al.*, 2001; Bock & Eberhard, 2003; Hartsuiker, Anton-Mendez, & van Zee, 2001; Hartsuiker *et al.*, 2003; Meyer & Bock, 1999; Vigliocco *et al.*, 1995).

## *Morphological stability*

One major consistent finding of the number attraction literature, cross-linguistically, is that attraction is often stronger in sentences with a singular head and a plural attractor than in sentences with a plural head and a singular attractor. Similarly, attraction from a gender mismatching noun appears to be stronger in sentences with neutral heads, followed by masculine heads and finally by feminine heads (e.g., Badecker & Kuminiak, 2007; Malko & Slioussar, 2013). These asymmetries have been classically interpreted as stemming from the morphological markedness of plural and possibly also feminine attractors. According to M&M, singular nouns carrying no feature do not have the potential to percolate and therefore be erroneously transferred to the verb. That is, the asymmetry is explained by the properties of the attractor noun. However, these asymmetries may also lie in the markedness of the head, since it systematically co-varies with that of the attractor in conditions where attraction can arise (that is, when the two nouns have mismatching features). In other words, the asymmetry may stem from the fact that plural (or feminine) heads, being marked, are more likely to be selected. One argument in favor of that account comes from the finding that semantic and morphophonological correlates of number or gender (described in the next sections), which do not systematically co-vary on the head and attractor nouns as is the case of markedness, show much clearer effects when manipulated on the head than when manipulated on the attractor. Another argument comes from the fact that languages have grammatical structures like pseudo-partitives, superficially similar to prepositional phrase modifiers, but which nevertheless take plural verbs (e.g., *A bunch of people *is/are demonstrating*). Haskell, Thornton, and MacDonald (2010) estimated that, in English, complex subjects with a singular first noun and a plural second noun take plural agreement in more than 20% of the cases, while subjects with a plural first noun and a singular second noun show less than 3% singular agreement. These authors reported experimental evidence that the grammatical plural in pseudo-partitive constructions primes attraction errors in constructions with PP modifiers (e.g., *A cluster of reporters were…* primes **The pencil in the gift bags were…*). In sum, it seems plausible that markedness effects do not arise from the markedness of the attractor (and thus have nothing to do with attraction per se), but rather from the markedness of the head: a marked feature on the head would be stronger/more stable, increasing the chances that it is correctly selected.[1]

## Effects of the notional representation of the subject phrase

### *Distributivity*

The presence of a plural local noun in some structures sometimes forces the interpretation of the subject as distributed (e.g., *The label on the bottles*). Speakers produce more plural verbs in sentences containing distributive subjects, as compared to sentences with non-distributive ones (e.g., *The key to the cabinets*) (e.g., Eberhard, 1999; Foote & Bock, 2012; Vigliocco, Butterworth & Garrett, 1996; Vigliocco, Butterworth & Semenza, 1995; Vigliocco, Hartsuiker, Jarema & Kolk, 1996). Distributivity can also be a function of the preposition; in *The gang on the motorcycles*, *on* promotes a distributive reading in which each member of the gang seems to be understood as capable of independent action, whereas in *The gang near the motorcycles*, *near* promotes a more collective reading, where the gang members are viewed as a unit. Again, more plural verbs were found with distributive subjects (Humphreys & Bock, 2005).

### *Semantic integration*

This factor refers to how closely the head noun and the attractor are linked in the semantic representation of the sentence. For example, in *The drawing of the flowers* a particular, integrated relation of the two referents (drawing and flowers) is implied, whereas in *The drawing with the flowers* the relation is a very generic, non-integrated relation of juxtaposition. The initial set of experiments on this factor showed that speakers tend to produce more plurals with semantically integrated subjects (Solomon & Pearlmutter, 2004), supposedly because in integrated subjects, the head and attractor nouns are more likely to be encoded together (but no timing measure was provided to back-up that claim). However, subsequent studies made contradictory claims, suggesting that non-integrated subjects were actually more likely to be interpreted as referring to distinct entities, i.e., they are more individuated, and therefore more likely to give rise to more plural agreement than integrated subjects (e.g., Brehm & Bock, 2013; Veenstra, Acheson, Bock, & Meyer, 2013). A clearer theoretical approach of the semantics of these structures is clearly needed to shed light on the reason for these inconsistencies. Nevertheless, these observations show that the way a speaker represents the numerosity of the subject has an effect on the feature that will be selected on the controller.

### *Spatial distribution*

Visual cognition research has shown that visual arrays occupying more space are perceived as containing more items. When the head noun of a quantified phrase (e.g., *Each alligator with humungous claws*) is illustrated with the constituent elements (alligators) spread far apart from one another, speakers tend to produce more plural agreement than when the same sentence is illustrated with a more condensed spatial distribution (Brehm, 2015).[2]

## Effects on Controller selection

Controller selection is the process responsible for selecting the controller whose features will be copied onto the agreement target. Badecker & Kuminiak (2007) suggested that attraction reveals the incorrect selection of the attractor as controller, via a cue-based retrieval process triggered by the verb (a similar, though different, proposal in comprehension has been proposed in Wagers *et al*., 2009 and subsequent studies[3]). Here, I adopt this hypothesis, and suggest that various effects reported in the literature actually show the signature of a cue-based process: similarity-based interference. In this view, attraction errors are similarity-based interference errors; they arise because there is an element in memory bearing some similarity to the controller, such that it is selected for agreement computation. Experimental work suggests that an element triggers stronger attraction if it is similar to the head semantically (in terms of *animacy*, *semantic overlap*, and *thematic roles*) or morphologically (in terms of *case marking*). I will suggest here that some of the syntactic modulations of attraction reported in the literature may also be interpreted as syntactic similarity effects: attractors in a syntactic position typically occupied by agreement controllers (*c-command* and *hierarchical height*) trigger more attraction than those in a position that is not occupied by controllers.

## Effects of semantic similarity

### *Animacy*

Attraction is stronger when the head and the attractor have the same animacy feature (e.g., *The blackboard behind the desks*) than when they differ in animacy (e.g., *The blackboard behind the teachers*; Barker *et al*., 2001).

### *Semantic overlap*

Attractor nouns with high overlap of semantic features with the head, that is, semantically similar to it (e.g., *The canoe by the sailboats*) trigger more attraction than those with lower overlap (e.g., *The canoe by the cabins*; Barker *et al*., 2001).

### *Thematic roles*

Attraction is stronger when the attractor noun is a plausible thematic agent for the verb (e.g., *The album by the classical composers was praised*) than when it is not (e.g., *The album by the classical composers was played*) (see also Hupet, Fayol, & Schelstraete, 1998). Along the same lines, the rate of plural agreement found in pseudo-partitive constructions (e.g., *A subset of problems are resolved*) increases with the relative topicality of the attractor with respect to the head (Smith, Franck, & Tabor, 2016). These findings show that an attractor that is a good topic is more likely to be selected as controller.

## Effects of case marking similarity

Probably the most prominent factor affecting attraction is the case ambiguity of the noun phrases: virtually all studies reporting attraction involve controller and attractor nouns that lack morphological marking of syntactic roles, either because the language lacks case markers (in English and many of the languages tested), or because case markers are present but ambiguous, which happens when the attractor has nominative case like the head (Badecker & Kuminiak, 2007; Hartsuiker *et al*., 2001, 2003). Attraction is virtually nonexistent when the head and attractor are distinctly case-marked (Badecker & Kuminiak, 2007; Lorimor *et al*., 2008; Malko & Slioussar, 2013; Marusic *et al*., 2013). These findings suggest that attractors that are more controller-like in terms of case marking trigger more attraction than those that are less similar to controllers.

## Effects of syntactic similarity

### *C-command*

C-command refers to a particular configuration of two nodes in the hierarchical structure: X c-commands Y if it has a sister node that dominates Y. C-command plays a crucial role in agreement in that agreement only takes place with a c-commanding head (Chomsky, 2000). Thus, c-command is a property of controllers. Experimental evidence shows that attractors occupying a position of c-command trigger more attraction than those occupying a position of precedence, which is not a position occupied by controllers. The plural accusative clitic *les* in French, which is in a position c-commanding the verb, triggers more attraction (e.g., **Le professeur les lisent*, *The professor them-Pl read-Pl) than the plural dative clitic *leur*, which precedes the verb (e.g., **Le professeur leur plaisent*, *The professor to-them-Pl please-Pl). Moreover, error rates with dative clitics are similar to those with prepositional phrase modifiers (e.g., **Le professeur des élèves lisent*, *The professor of the students read), which also occupy a position of precedence to the verb (Franck *et al*., 2010). Similarly, in sentences with moved complex objects that contain a head and a prepositional phrase modifier, the c-commanding head triggers more attraction (*patientes* in **Quelles patientes du médecin dis-tu que l'avocat defendent?* *Which patients of the doctor do you say that the lawyer defend?) than the modifier that precedes the verb (*patients* in **Le médecin de quelles patientes dis-tu que l'avocat defendent?* *The doctor of which patients do you say that the lawyer defend?) (Franck *et al*., 2015).

### *Hierarchical height*

Agreement controllers typically occupy a high position in the hierarchical structure. Studies have shown that when the subject contains two embedded prepositional phrase modifiers, attraction is stronger with the modifier situated high in the tree (*programs* in **The computer with the programs of the experiment are broken*) than with the one situated low (*experiments* in **The computer with the program of the experiments are broken*) (Franck, Vigliocco, & Nicol, 2002; Gillespie & Pearlmutter, 2011). The two modifiers occupy a position of precedence with respect to the verb, however, one may entertain the possibility that hierarchical height is a proxy to c-command, since c-commanding elements are higher than preceding elements.

## Syntactic structure and memory in attraction: Evidence for a close alignment

The cue-based memory retrieval process underlying Controller selection is assumed to take place on the hierarchical representation of the sentence, such that it is tightly constrained by it. The conception of hierarchical structure adopted here critically differs from syntactic encoding models like M&M in that multiple hierarchical representations are assumed to be encoded successively as elements progressively move to reach their final, surface position (a detailed description of the linguistic formalism as well as illustrations of the hierarchical representation of the structures manipulated can be found in Franck, Frauenfelder, & Rizzi, 2007; Franck, Lassi, Frauenfelder, & Rizzi, 2006; Franck, Soare, Frauenfelder, & Rizzi, 2010). Experimental support to this conception comes from the findings, discussed earlier in the critical analysis of M&M, that two structures that are superficially different but have identical underlying hierarchical structure show similar attraction profiles, whereas two structures that are superficially identical but have different underlying structures show different attraction profiles. Additional evidence comes from the various reports ever since Bock and Miller (1991) of attraction from a moved object (e.g., Franck *et al*., 2006, 2010; Santesteban, Pickering, & Branigan, 2013; Staub, 2009, 2010).

Such an effect is at first glance unexpected since the subject and the verb are linearly contiguous in the structures tested. Nevertheless, formal syntax has argued that the object leaves a trace not only in its canonical position but also in an intermediate position through which it transits to satisfy syntactic constraints (Kayne, 1989; see Warren & Gibson, 2005 for psycholinguistic evidence for intermediate traces in comprehension). Critically, that intermediate position intervenes between the subject and the verb in the hierarchical structure, and is therefore expected to be visible to the Controller selection process.[4]

In order to test more directly the hypothesis that the memory retrieval process underlying Controller selection operates on hierarchical structure, Matt Wagers and I designed a study that tested the prediction that attractors situated in syntactic positions triggering more attraction are easier to retrieve from memory (Franck & Wagers, 2015). To do this, we combined a grammaticality judgment experiment with a probe recognition experiment using the response-signal speed-accuracy trade-off procedure (SAT). We used a grammaticality judgment task because our previous research indicates that this task replicates the syntactic modulation of attraction found in sentence production (Franck *et al*., 2015). I will assume here that both tasks tap into the process of Controller selection. We used the SAT paradigm because it investigates the fine-grained time-course of processing and enables separate measures of retrieval speed and retrieval accuracy (e.g., McElree & Dosher, 1989). In such a probe recognition experiment, participants are trained to respond to a signal presented at varying time points after the onset of the recognition probe (spanning the full time course of retrieval between about 100 ms to 3000 ms), indicating whether the probe was in the list. Distribution of accuracy as a function of retrieval time typically shows an initial phase of chance level performance (participants did not have enough time to select the correct answer), followed by a phase of increasing accuracy, followed by an asymptotic period. The asymptote provides a measure of the overall probability of retrieval (accuracy), which is a joint function of the overall quality of the memory representation and cues at retrieval. Retrieval speed is measured by the intercept of the function, indicating when information first becomes available, and the rate of rise, indicating the rate at which accuracy grows from chance to asymptote. These two parameters provide key indicators of the dynamics of retrieval, independently of the quality of memory representations. In contrast to previous SAT studies of sentence processing that all employ a sentence-acceptability judgment task (e.g., McElree *et al*., 2003; Van Dyke & McElree, 2006), we used such a probe recognition task in order to get a direct measure of retrieval parameters. Participants first read the sentence presented word by word in a Rapid Serial Visual Presentation manner. At the end of the sentence, they were asked to judge whether a probe word was in the sentence or not. Probe words were subjects, attractors, and words that were not in the sentence. SAT parameters were then linked to the attraction rates obtained in the grammaticality judgment experiment on the same items, in order to examine the alignment of the two measures.

We also incorporated another novelty into the design: the materials involved French Jabberwocky in which nouns were replaced by pseudo-nouns but grammatical morphemes and verbs were preserved. This made it so that participants had no difficulty judging the grammaticality of agreement, while semantic similarity influences were out of the way. Two types of structures were manipulated, both involving two attractors. The first structure involved object relatives with complex objects similar to those tested in Franck *et al*. (2015). These structures contain an attractor c-commanding the verb (*dafran* in *Which dafrans of the brapou do you say that the bostron defends?*) and one preceding it (*dafran* in *The brapou of which dafrans do you say that the bostron defends?*). The second structure involved two subject modifiers like those tested in Franck *et al*. (2002). Hierarchical height was manipulated by contrasting a modifier situated higher (*dafrans* in *The bostron of the dafrans of the drapou sleeps)* and one situated lower (*dafrans* in *The bostron of the drapou of the dafrans sleeps*). Here, both attractors precede the verb.

Results from the grammaticality judgment experiment showed that Jabberwocky elicits attraction. Importantly, the c-command versus precedence contrast in complex objects found in natural language was replicated: more attraction was found with the c-commanding object head than with the object modifier intervening by precedence. We found as much attraction from the low as from the high attractor in double subject modifiers. This might be due to the fact that Jabberwocky promotes a more strictly syntactic computation of agreement, in which all that counts is the distinction between c-command and precedence, while finer distinctions among precedence relations have no role to play in the syntax (its effect may be in the semantics). But the more important finding is that results from the probe recognition experiment showed a close alignment with grammaticality judgments. First, subjects are more accessible (higher asymptote) and retrieved faster (faster dynamics) than the two attractors. Interestingly, the higher accessibility of subjects is found independently of their linear position: it is found equally in the complex object condition, where the subject is linearly just before the probe word, and in the subject modifier condition where the subject is situated linearly far from the probe word. This contrasts with list memorization where accessibility is a function of distance. The finding that subjects are retrieved faster than attractors also contrasts with list memorization where all units (apart from the most recent one) are retrieved at the same speed (McElree, 2006). This suggests that subjects are maintained in the focus of attention, even when separated from their verb by PP modifiers (in line with Wagers & McElree, in press), capturing the fact that in most of the cases, they are correctly retrieved as the controller for agreement computation. Second, the accessibility of the two attractors aligns with their potential to trigger attraction: the c-commanding attractor in complex objects was significantly more accessible (higher asymptote) than the preceding one whereas no difference was found between the two attractors of the subject modifier structure, in line with the finding that a c-commanding element triggers more attraction but PP modifiers trigger similar attraction in the grammaticality judgment task.

These novel results suggest that the memory retrieval processes underlying sentence processing are constrained by the grammar: subjects are especially prominent in memory, followed (by far) by elements c-commanding the verb, and then by those situated in a position of precedence. They bring direct support to the hypothesis that attraction is a function of the attractor's visibility to the memory retrieval process, and that this higher visibility itself depends on the attractor's syntactic position in the hierarchically structured sentence, in keeping with the critical distinction between c-command and precedence. In sum, I tentatively suggest that syntactic theory describes the strength of memory representations.

## Summary and future challenges

Capitalizing on the initial proposition by Badecker and Kuminiak (2007) that attraction results from the incorrect selection of the controller, I have proposed a functional model of agreement that involves two components, Feature selection and Controller selection, the latter being considered as the locus of attraction proper. I developed a typology of attraction effects that splits these effects into two classes according to their functional locus in the model: those that arise during Feature selection, and those that arise during Controller selection. The subclasses of the typology are structured according to the nature of the factors that compose them: semantic, morphophonological, and morphological factors that influence the stability of the feature, and notional factors that influence the construal of the subject phrase, all affecting Feature selection; semantic similarity, case marking similarity and syntactic similarity between the head and the attractor affecting Controller selection. This way of cutting the pie is radically different from that proposed in M&M where effects are split according to the nature of the variables that underlie them: semantic variables affect the first stage of Marking, whereas syntactic and morphological variables affect the second stage of Morphing. In the current proposal, different types of semantic factors and different types of morphological factors are argued to influence both processes, although in different ways.

The framework I have proposed here opens two important avenues for future research. The first avenue concerns the fine characterization of syntactic representations as mental objects. The psycholinguistics of sentence production has managed to set aside this core question for 40 years of research, contributing to a drifting-apart between psycholinguistics and linguistics ever since the disillusionment following the failure of the derivational theory of complexity (Fodor, Bever & Garrett, 1972). As clearly expressed in a recent state of the art review of the sentence production literature (Bock & Ferreira, 2014), psycholinguistic accounts of language are set apart from linguistic approaches in that they are concerned with "the situatedness of sentence production in the circumstances of communication. Speakers have to do a whole lot more than create grammatically acceptable sentences. They have to create acceptable sentences that make sense. This means that they have to convey particular notions to particular people in particular circumstances in a particular language" (p. 42). And indeed, the question of the form of syntactic representations is absent from most sentence production studies. Even the broad research program on syntactic priming showing that speakers tend to re-use a particular syntactic structure (see Chapters 6, 7, 14) remains fundamentally agnostic about the shape of these structures. This line of work is restricted to a few syntactic structures and rests on a superficial description of their properties. That description is sufficient when the prime and target have identical structures, since syntactic identity can easily be assessed without deeper analysis. Nevertheless, a few studies have pushed the question further in exploring the possibility that priming generalizes to other structures (Bock & Loebell, 1990; Bock, 1989; Griffin & Weinstein-Tull, 2003). This approach seems extremely promising in that it reveals an even more abstract level of representation where syntactic similarity, the underlying basis for generalization, has to be characterized. Drawing the map of similarities and differences between structures cannot bypass a fine analysis of their properties.

The second avenue is the study of the relation between sentence production processes and memory mechanisms, which has seldom been raised (in contrast to research in sentence comprehension). In the model of agreement proposed, both Feature selection and Controller selection are memory retrieval processes. I argued that Feature selection, which amounts to retrieving nominal features from the long-term memory store of function words, operates under various types of constraints, in line with the broad literature showing interactivity in single word production. I suggested that the process of Controller selection, responsible for retrieving the controller from the memory representation of the currently built sentence, shows the hallmark of cue-based retrieval: similarity-based interference. Empirical evidence shows that a higher overlap between memorized units endowed with semantic and syntactic controller-specific features creates interference that occasionally manifests in the form of an attraction error. In this view, attraction is the result of the incorrect selection of the controller during agreement computation. I ended with the report of experimental data obtained with a new design allowing us to explore more directly the link between attraction and memory. The data show that syntax-based variations of attraction strength closely align with variations in memory retrieval measures: sentential subjects are more accessible and retrieved faster than attractors, and attractors that generate more attraction are more accessible than those that generate less attraction.

Some major questions remain open. One question is whether an error in Controller selection arises because an erroneous syntactic tree has been built, or whether it is independent of the overall structure building process, as suggested here. It has often been informally observed that the production of an attraction error does not entail that the speaker has reached the wrong interpretation of the sentence in which the attractor is the subject, and experimental evidence seems to support that claim (Lau *et al.*, 2008). Nevertheless, it seems premature to date to firmly conclude in that direction, and more direct empirical tests need to be designed. Another question concerns the precise identification of the aspects of the memory retrieval process that are affected by the syntactic position of the attractor. The results of my work with Matt Wagers suggest that subjects remain in the focus of attention, however, the memory mechanisms underlying attractor's access are less clear. One possibility is that syntactic position affects the decay rates of the controller and attractors in the sentence: the c-commanding object may be reactivated when reaching the verb, boosting its level of activation (McElree, 2000). In order for the model to account for the data, decay rates would then need to be a function of syntactic structure (e.g., Lewis & Vasishth, 2005; McElree, 2000). Another possibility is that syntactic position affects retrieval cues: c-commanding positions and more generally high positions in the tree are typical positions of subjects, such that attractors occupying these positions would be more prone to interfere in Controller selection. Translating syntactic position into retrieval cues is challenging given the fundamentally relational nature of positional information, but proposals have been made along those lines (e.g., Kush, 2013; Wagers & McElree, in press). I pinpointed the relevance of further examining the role of similarity among different types of structures in syntactic priming. The same logic could be applied to the study of attraction, in identifying what syntactic properties (beyond morphological similarity) make an attractor controller-like. Such a view links grammatical factors to surface properties that probabilistically correlate with them, opening the possibility that these correlates play a role in driving the production system as well (e.g., Bever & Poeppel, 2010). It also introduces the intriguing possibility that the fine syntactic constructions characterized by linguistic theory are represented as continuous mental objects, organized along a structural proximity metric (Tabor, Cho, & Szkudlarek, 2013). These questions constitute an interesting program of future research.

The finding that attraction patterns may stem from variations in memory retrieval is not the end of our journey. What remains to be understood is why memory is organized the way it is, which means identifying the cognitive constraints that shape natural language grammars. Nevertheless, the line of research sketched here paves the way for a new relationship between syntactic theory and cognitive psychology, and raises the hope that it will (re)open the debate on the possibility that the theory of competence is also a theory of performance.

## Acknowledgments

The author is grateful to Whit Tabor for the energizing and thought-provoking discussions on the ideas developed in that paper.

## NOTES

1 Some languages like Slovenian show the opposite profile with singular verbs being required with quantified plural heads referring to more than five units (in which case the sentence amounts to saying *five people has arrived*, Marušič *et al*., 2011). Phenomena of closest conjunct agreement also give rise to cases where agreement takes place with the linearly closest conjunct, even if this conjunct is not marked (e.g., Marušič *et al*. 2011; Haskell & MacDonald, 2005). More work is needed to estimate the role of statistical distributions in attraction asymmetries.

2 Recent evidence from structures without attractors suggests that it is the degree of individuation of the head and not the number of units that it involves that counts (Mirković & MacDonald, 2013). In Serbian, quantified noun phrases take singular verbs. However, more plural verbs are found with a quantifier like several, which is judged as more individuated, than with many. Similar differences are found between agentive and existential verbs. The former are claimed to promote a more individuated interpretation, and indeed more plural verbs erroneously occur with the former.

3 The retrieval approach proposed in these studies of attraction in sentence comprehension differs from that of Badecker and Kuminiak (2007) and the one advocated here in assuming that retrieval is selectively triggered when an agreement error is encountered. This conclusion was reached on the basis of the finding that a number mismatch between the controller and the head only affects verb processing in sentences containing an agreement error. The proposal here is that both the generative component of the production process and the predictive component of the comprehension process involve retrieving the controller to compute agreement (even though comprehension involves some specificities, see Franck, Colonna & Rizzi, 2015 for a discussion).

4 In two relevant studies involving on-line response time measures, Staub (2009, 2010) argued in favor of two distinct causes underlying attraction in object relatives and in prepositional phrase modifiers (in line with Bock & Miller, 1991). Evidence comes from the distribution of response times in the production of the verb in these two structures: whereas prepositional modifier attraction shows a small but systematic effect across trials, object attraction shows an irregular, strong effect on only a subset of trials. This finding remains unexplained in the current framework, which suggests that a single mechanism underlies both types of attraction.

## REFERENCES


Alcocer, P., & Phillips, C. (2012). Using relational syntactic constraints in content-addressable memory architectures for sentence parsing. *Ms. University of Maryland, College Park*.

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language*, 56(1), 65–85. doi: 10.1016/j.jml.2006.08.004

Barker, J., Nicol, J., & Garrett, M. (2001). Semantic factors in the production of number agreement. *Journal of Psycholinguistic Research*, 30(1), 91–114. doi: 10.1023/A:1005208308278

Bever, T. G., & Poeppel, D. (2010). Analysis by synthesis: A (re)emerging program of research for language and vision. *Biolinguistics*, 4(2), 174–200.

Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences*, 7(02), 173–188. doi: 10.1017/S0140525X00044149

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition*, 31(2), 163–186. doi: 10.1016/0010-0277(89)90022-X

Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31(1), 99–127. doi: 10.1016/0749-596X(92)90007-K

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, 8(1), 57–99. doi: 10.1080/01690969308406949

Bock, K., Eberhard, K. M., & Cutting, J. C. (2004). Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language*, 51(2), 251–278. doi: 10.1016/j.jml.2004.04.005

Bock, K., Eberhard, K. M., Cutting, J. C., Meyer, A. S., & Schriefers, H. (2001). Some attractions of verb agreement. *Cognitive Psychology*, 43(2), 83–128. doi: 10.1006/cogp.2001.0753

Bock, K., & Ferreira, V. (2014). Syntactically speaking. In M. Goldrick, V. Ferreira and M. Miozzo (Eds), *Oxford handbook of psycholinguistics*, pp. 21–46, Oxford: OUP.

Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics*, pp. 945–984, San Diego: Academic Press.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, 35(1), 1–39. doi: 10.1016/0010-0277(90)90035-I

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: bridging the syntactic cleft. *Psychol Rev*, 99(1), 150. doi: 10.1037/0033-295X.99.1.150

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23(1), 45–93. doi: 10.1016/0010-0285(91)90003-7

Bock, K., Nicol, J., & Cutting, J. C. (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language*, 40(3), 330–346. doi: 10.1006/jmla.1998.2616

Brehm, L., & Bock, K. (2013). What counts in grammatical number agreement? *Cognition*, 128(2), 149–169. doi: 10.1016/j.cognition.2013.03.009

Brehm, L. (2015). *Visual grouping affects number agreement production*. Paper presented at the 2015 CUNY Sentence Processing Conference, Los Angeles, CA, USA.

Brems, L. (2003). Measure Noun constructions: An instance of semantically-driven grammaticalization. *International Journal of Corpus Linguistics*, 8(2), 283–312. doi: 10.1075/ijcl.8.2.05bre

Chomsky, N. (1981). *Lectures on government and binding*. Foris: Dordrecht.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step – Essays in minimalist syntax in honor of Howard Lasnik* (pp. 89–155). Cambridge, MA: MIT Press.

Clifton, C., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di Linguistica*, 11, 11–39.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321. doi: 10.1037/0033-295X.93.3.283

Deutsch, A., & Dank, M. (2009). Conflicting cues and competition between notional and grammatical factors in producing number and gender agreement: Evidence from Hebrew. *Journal of Memory and Language*, 60(1), 112–143. doi: 10.1016/j.jml.2008.07.001

Eberhard, K. M. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language*, 36(2), 147–164. doi: 10.1006/jmla.1996.2484

Eberhard, K. M. (1999). The accessibility of conceptual number to the processes of subject-verb agreement in English. *Journal of Memory and Language*, 41(4), 560–578. doi: 10.1006/jmla.1999.2662

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making Syntax of Sense: Number Agreement in Sentence Production. *Psychol Rev*, 112(3), 531–559. doi: 10.1037/0033-295X.112.3.531

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The Psychology of Language*. New York: McGraw-Hill.

Foote, R., & Bock, K. (2012). The role of morphology in subject-verb number agreement: A comparison of Mexican and Dominican Spanish. *Language and Cognitive Processes*, 27(3), 429–461. doi: 10.1080/01690965.2010.550166

Franck, J. (2011). Reaching agreement as a core syntactic process. *Natural Language & Linguistic Theory*, 29(4), 1071–1086. doi: 10.1007/s11049-011-9153-1

Franck, J., Colonna, S., & Rizzi, L. (2015). Task-dependency and structure-dependency in number interference effects in sentence comprehension. *Frontiers in Psychology*, 6, 349. doi: 10.3389/fpsyg.2015.00807, doi: 10.3389/fpsyg.2015.00349

Franck, J., Frauenfelder, U. H., & Rizzi, L. (2007). A syntactic analysis of interference in subject-verb agreement. *MIT Working Papers in Linguistics*, 53, 173–190.

Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition*, 101(1), 173–216. doi: 10.1016/j.cognition.2005.10.003

Franck, J., Soare, G., Frauenfelder, U.H., & Rizzi, L. (2010). Object interference in subject-verb agreement: The role of intermediate traces of movement. *Journal of Memory and Language*, 62(2), 166–182. doi: 10.1016/j.jml.2009.11.001

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. H. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, 23(3), 329–374. doi: 10.1080/01690960701467993

Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, 17(4), 371–404. doi: 10.1080/01690960143000254

Franck, J., & Wagers, M. (2015). Hierarchical structure and memory retrieval mechanisms in attraction: An SAT study. CUNY sentence processing, Los Angeles, March 19–21.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 27–52. doi: 10.2307/412187

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 133–177). New York: Academic Press.

Garrett, M. F. (1980). The limits of accommodation: Arguments for independent processing levels in sentence production. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen and hand*, 263–271. New York: Academic Press.

Garrett, M. F. (1989). Processes in language production. *Linguistics: The Cambridge Survey: Volume 3, Language: Psychological and Biological Aspects*, 3, 69.

Gillespie, M., & Pearlmutter, N. (2011). Hierarchy and scope of planning in subject-verb agreement production. *Cognition*, 118(3), 377–397. doi: 10.1016/j.cognition.2010.10.008

Goldrick, M. (2006). Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. *Language and Cognitive Processes*, 21(7-8), 817–855. doi: 10.1080/01690960600824112

Griffin, Z. M., & Weinstein-Tull, J. (2003). Conceptual structure modulates structural priming in the production of complex sentences. *Journal of Memory and Language*, 49(4), 537–555. doi: 10.1016/j.jml.2003.08.002

Guasti, M. T., & Rizzi, L. (2002). Agreement and tense as distinct syntactic positions: Evidence from acquisition. *Functional structure in DP and IP: The cartography of syntactic structures*, 1, 167–194.

Hartsuiker, R. J., Anton-Mendez, I., & van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, 45(4), 546–572. doi: 10.1006/jmla.2000.2787

Hartsuiker, R. J., Schriefers, H. J., Bock, K., & Kikstra, G. M. (2003). Morphophonological influences on the construction of subject-verb agreement. *Memory & Cognition*, 31(8), 1316–1326. doi: 10.3758/BF03195814

Haskell, T. R., & MacDonald, M. C. (2003). Conflicting cues and competition in subject-verb agreement. *Journal of Memory and Language*, 48(4), 760–778. doi: 10.1016/S0749-596X(03)00010-X

Haskell, T. R., Thornton, R., & MacDonald, M. C. (2010). Experience and

grammatical agreement: Statistical learning shapes number agreement production. *Cognition*, 114(2), 151–164. doi: 10.1016/j.cognition.2009.08.017
Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579. doi: 10.1126/science.298.5598.1569
Humphreys, K. R., & Bock, K. (2005). Notional number agreement in English. *Psychonomic Bulletin & Review*, 12(4), 689–695. doi: 10.3758/BF03196759
Hupet, M., Fayol, M., & Schelstraete, A.-M. (1998). Effects of semantic variables on the subject-verb agreement processes in writing. *British Journal of Psychology*, 89, 59–75. doi: 10.1111/j.2044-8295.1998.tb02673.x
Kayne, R. (1989). Facets of Romance past participle agreement. *Dialect variation and the theory of grammar*, 85–103.
Keeney T. J., & Wolfe, J. (1972). The acquisition of agreement in English. *Journal of Verbal Learning and Verbal Behavior*, 11, 698–705.
Kush, D. (2013). *Respecting relations: Memory access and antecedent retrieval in incremental sentence processing*. Doctoral dissertation, University of Maryland.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
Lewis, R. & Vasishth, S. (2005). An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. *Cognitive Science*, 29, 375–419. doi: 10.1207/s15516709cog0000_25
Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends Cogn Sci*, 10(10), 447–454. doi: 10.1016/j.tics.2006.08.007
Lorimor, H., Bock, K., Zalkind, E., Sheyman, A., & Beard, R. (2008). Agreement and attraction in Russian. *Language and Cognitive Processes*, 23(6), 769–799. doi: 10.1080/01690960701774182
Malko, A., & Slioussar, N. (2013). Attraction Errors in Gender Agreement: Evidence from Russian. Malko Slioussar. FASL. Michigan Slavic Publications.
Marušič, F., Nevins, A., & Badecker, B. (2011). The grammars of conjunction agreement in Slovenian. *Ms. University of Nova Gorica, University College London, University of Arizona*.
Marušič, F., Nevins, A., & Saksida, A. (2007). *Last-conjunct agreement in Slovenian*. Paper presented at the Formal Approaches to Slavic Linguistics.
McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, 29(2), 111–123. doi: 10.1023/A:1005184709695
McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation*. San Diego: Academic.
McElree, B. & Dosher, B.A. (1989). Serial position and set size in short-term memory: Time course of recognition. *Journal of Experimental Psychology: General*, 18, 346–373. doi: 10.1037/0096-3445.118.4.346
McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, 48(1), 67–91. doi: 10.1016/S0749-596X(02)00515-6
Meyer, A. S., & Bock, K. (1999). Representations and processes in the production of pronouns: Some perspectives from Dutch. *Journal of Memory and Language*, 41(2), 281–301. doi: 10.1006/jmla.1999.2649
Mirković, J., & MacDonald, M. C. (2013). When singular and plural are both grammatical: Semantic and morphophonological effects in agreement. *Journal of Memory and Language*, 69(3), 277–298. doi: 10.1016/j.jml.2013.05.001
Rizzi, L. (1990). *Relativized minimality*: The MIT Press.
Santesteban, M., Pickering, M. J., & Branigan, H. P. (2013). The effects of word order on subject-verb and object-verb agreement: Evidence from Basque. *Journal of Memory and Language*, 68(2), 160–179.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, 12(4), 323–328. doi: 10.1111/1467-9280.00359

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology*, 49(4), 370–407. doi: 10.1016/j.cogpsych.2004.05.001

Smith, G., Franck, J., & Tabor, W. (2016). A theory of agreement attraction based on a continuous semantic representation space. Empirical Methods and Grammatical Theory: The Issue of Gradience, Konstanz, 24-26 February.

Solomon, E. S., & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, 49(1), 1–46. doi: 10.1016/j.cogpsych.2003.10.001

Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, 60, 308–327. doi: 10.1016/j.jml.2008.11.002

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition*, 114(3), 447–454. doi: 10.1016/j.cognition.2009.11.003

Stowe, L. A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, 1(3), 227–245. doi: 10.1080/01690968608407062

Tabor, W., Cho, P. W., & Szkudlarek, E. (2013). Fractal Analysis Illuminates the Form of Connectionist Structural Gradualness. *Topics in Cognitive Science*, 5(3), 634–667. doi: 10.1111/tops.12036

Thornton, R., & MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language*, 48, 740–759. doi: 10.1016/S0749-596X(03)00003-2

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, 55(2), 157–166. doi: 10.1016/j.jml.2006.03.007

Veenstra, A., Acheson, D. J., Bock, J., & Meyer, A. S. (2013). Effects of semantic integration on the production of subject-verb agreement: Evidence from Dutch. Language, Cognition and Neuroscience, 29(3), 355–380.

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, 34(2), 186–215. doi: 10.1006/jmla.1995.1009

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual factors. *Cognition*, 51, 261–298. doi: 10.1016/S0010-0277(96)00713-5

Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language*, 40(4), 455–478. doi: 10.1006/jmla.1998.2624

Vigliocco, G., & Franck, J. (2001). When sex affects syntax: Contextual influences in sentence production. *Journal of Memory and Language*, 45(3), 368–390. doi: 10.1006/jmla.2000.2774

Vigliocco, G., Hartsuiker, R. J., Jarema, G., & Kolk, H.J. (1996). How many labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes*, 11, 407–421. doi: 10.1080/016909696387169

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: is proximity concord syntactic or linear? *Cognition*, 68(1), B13–B29. doi: 10.1016/S0010-0277(98)00041-9

Wagers, M., & McElree, B. (in press). Working memory and language processing: Theory, data, and directions for future. *Biolinguistics*. doi: 10.1017/cbo9780511980435.014

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61(2), 206–237. doi: 10.1016/j.jml.2009.04.002

Warren, T., & Gibson, E. (2005). Effects of NP type in reading cleft sentences in English. *Language and Cognitive Processes*, 20(6), 751–767. doi: 10.1080/01690960500051055

# 3 Signal Reduction and Linguistic Encoding

T. FLORIAN JAEGER AND ESTEBAN BUZ

University of Rochester

## Introduction

Human languages provide speakers with a remarkable degree of flexibility in how to linguistically encode near-meaning equivalent messages. This chapter focuses on what is arguably the most pervasive type of flexibility: flexibility in the amount or quality of the signal that encodes the speaker's message. Figure 3.1 illustrates this for English (inspired by Friedman, 2013). For example, an element may be mentioned or omitted (e.g., the optional complementizer *that*, or the argument *the World Cup*), or the articulatory realization of an element may be more or less detailed (e.g., producing a more centralized vowel or shortening the duration of a word). Such flexibility has been of central interest in psycholinguistic research: speakers' preferences to encode a message with a more or less reduced signal serve as a window into the architecture underlying the language production system.

Although speakers typically do not become of aware of this flexibility while talking, the choice between more or less reduced linguistic forms or signals is ubiquitous within and across languages. Alternations like those illustrated in Figure 3.1 exist across many, if not all, languages. Languages differ, however, in the specific alternations that they afford. For example, many languages allow omission of grammatical subjects in certain contexts (e.g., Italian, Japanese, Russian, and Yucatec Maya), whereas this omission is considered ungrammatical—or restricted to colloquial registers—in other languages (e.g., English). Other examples of reduction include optional mention of case-marking (e.g., in Japanese, Korean, and Turkish) or optional head-marking morphology (e.g., in many languages of the Balkan sprachbund), neither of which are available in English.

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

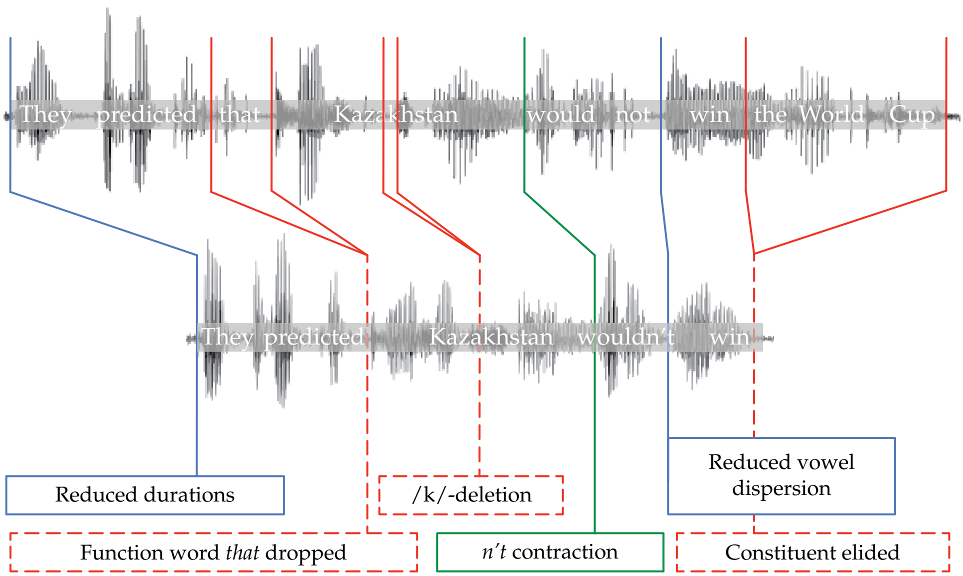



**Figure 3.1** Illustration of a few types of implicit decisions speakers make during linguistic encoding that affect the degree of signal reduction. Reduction can be caused by decisions at multiple levels of linguistic encoding, including sentence, lexical, and phonological planning, as well as articulation. In the appropriate context, the upper and lower utterances encode the same message; yet the lower utterance contains shorter linguistic forms and is realized with a much reduced speech signal, compared to the upper utterance. Inspired by Friedman (2013).

Reduction constitutes the empirical focus of this chapter. Specifically, we focus on *probabilistic reduction*: a large body of work has found that speakers tend to produce shorter linguistic forms and more reduced signals for contextually predictable parts of the message. To the best of our knowledge, a systematic review of research on reduction across different levels of linguistic representations has so far been lacking. We thus begin with a summary of this literature and the questions it raises for future research.

The second part of this chapter reviews competing theories and accounts of the empirical findings discussed in the first part. Although we focus on probabilistic reduction, the discussion bears on more general architectural questions. In particular, we discuss competing views of the link between production and comprehension, as well as the link between online processing and biases implicitly encoded in linguistic representations. We distinguish between three broad classes of accounts. One hypothesis holds that flexibility in encoding a message allows speakers to navigate the attentional and memory demands of language production. This type of explanation is sometimes referred to as "production-internal" (Arnold, 2008), "production-based" (Gahl, Yao, & Johnson, 2012), "production-oriented" (Lindblom, 1990a), or "production-centered" (Watson, Arnold, & Tanenhaus, 2010).

This has been contrasted with the idea that production is affected by communicative considerations. According to the latter view, the mechanisms underlying linguistic encoding are—directly or indirectly—affected by comprehension (e.g., Brennan & Clark, 1996, Clark & Fox Tree, 2002; Jaeger, 2006; Lindblom, 1990a). This alternative idea is variously referred to as, for example, "listener-oriented" (Arnold, 2008), "comprehension-facilitation" (Arnold, Kahn, & Pancani, 2012), "intelligibility-based" (Gahl *et al.*, 2012), or "audience design" (Clark & Murphy, 1982; Galati & Brennan, 2010). Here we use the labels *production ease* and *communicative* accounts to refer to these two views.

Independent of what (mixture of) pressures ultimately drive speakers' preferences, there are questions about whether these pressures operate on-line, directly affecting speakers' preferences during incremental linguistic encoding, or off-line, changing linguistic *representations* and thus only indirectly affecting incremental encoding. Explanations that focus on the latter possibility constitute a third type of account, which we will refer to as *representational* accounts (e.g., Pierrehumbert, 2001; Wedel, 2006). For each of these accounts, we review specific proposals and isolate some challenges we consider particularly pressing for future research. The picture that emerges from this discussion is one in which probabilistic reduction is not driven by any single factor, but rather the result of multiple mechanisms.

*A few terminological clarifications* Throughout this chapter, we refer to such differences as *reduction* (and to *reduced* variants), without meaning to imply a directionality of this process: for many phenomena we discuss, it is an open question whether they are better understood as reduction or enhancement. For example, although it might seem more intuitive to think of the complementizer *that* in Figure 3.1 as being optionally omitted, there are also arguments as to why it is better thought of as being optionally mentioned.[1] In conversational American English, for example, the complementizer *that* is absent in about 83% of all complement clauses (Jaeger, 2010, p. 29). Even when the most frequent complement clause embedding verbs are excluded, omission is more frequent (53%) than mention of complementizer *that* (Jaeger, 2010, Table 1). This makes it difficult to determine whether this alternation is better understood as optional mention or optional omission. Similarly, word durations may undergo reduction (i.e., shortening) or enhancement (i.e., lengthening).

We also distinguish between message components, linguistic forms, and their realization in the linguistic signal. Message components are parts of the message speakers wish to convey (e.g., a specific lexical meaning). Linguistic forms are instances of linguistic categories, such as phonological segments, words, and syntactic structures. These forms are not directly observable. Rather, they underlie the observable linguistic signal. The linguistic signal can be acoustic (in the case of speech) or visual (in the case of gestures, sign language, or writing). We sometimes refer to more or less reduced *forms* to highlight that reduction goes beyond gradient manipulation of the *signal* and includes cases where language provides speakers with several more or less reduced linguistic forms (e.g., mentioning or omitting *the world cup* in Figure 3.1).

## Probabilistic reduction: Contextual predictability and signal reduction

As shown in Figure 3.1, reduction can take place at different levels of linguistic encoding. Reduction at many of these levels has been found to be correlated with contextual predictability, so that more probable (and less informative) message components tend to be realized with reduced signal.[2] We begin with a summary of work on phonetic and phonological reduction. Then we summarize work at successively higher levels of linguistic encoding, including morphological contraction, the omission of optional function words, and the realization of referring expressions. We close this section with an overview of open empirical questions.

### *Phonetic and phonological reduction and omission*

A large number of studies have investigated the articulatory or acoustic reduction of phonemes, syllables, and words. This research has found that contextually predictable instances of words tend to be produced with shorter duration (e.g., Aylett & Turk, 2004; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Tily *et al.*, 2009) and less articulatory detail (e.g., Aylett & Turk, 2006; Gahl *et al.*, 2012; Son & Santen, 2005). Such probabilistic phonetic reduction has been observed in conversational speech corpora (e.g., Arnon & Cohen Priva, 2014; Aylett & Turk, 2004; Bell *et al.*, 2003; Gahl *et al.*, 2012; Pluymaekers, Ernestus, & Baayen, 2005a) and in the lab, including read speech (e.g., Arnon & Cohen Priva, 2013; Gahl & Garnsey, 2004; Kurumada, 2011) and unscripted speech (e.g., Watson *et al.*, 2010).

For example, contextually predictable instances of words tend to have more reduced vowels (e.g., Aylett & Turk, 2006; but see null results in Bürki, Ernestus, Gendrot, Fougeron, & Frauenfelder, 2011; Clopper & Pierrehumbert, 2008; Gahl *et al.*, 2012; Scarborough, 2010) and consonants (Rose, 2017: Ch. 3; Torreira & Ernestus, 2009, 2012). Aylett and Turk investigated predictability based reduction in a corpus of citation speech. They measured syllable durations and first and second formant values of vowels within those syllables and binned syllables into high and low predictability based on unigram, bigram, and trigram probabilities. They found that syllables with high predictability were shorter in duration and vowels within those syllables showed more centralization. Contextually predictable words are also more likely to undergo phonological weakening or deletion (Bell *et al.*, 2009, 2003). As an example, many varieties of English favor the reduction of complex codas in some phonological environments. A specific case of this is *t/d*-deletion, where a *t* or *d* that is present in citation form is not produced. Such *t/d*-deletion is more common in predictable words (Gahl, Jurafsky, & Roland, 2004; Jurafsky, Bell, Gregory, & Raymond, 2001; see also Bybee & Hopper, 2001). Other research has further found that a segment's informativity[3] about the word affects the segment's realization even after the word's predictability is taken into account (e.g., van Son & Pols, 2002; van Son & van Santen, 2005).

Similar reduction effects have also been observed as a function of previous mention of a word (Bard *et al*., 2000; Bell *et al*., 2009; Pluymaekers, Ernestus, & Baayen, 2005b; Watson *et al*., 2010). Since the statistics of human language are such that previous mention generally increases the probability of being mentioned again (e.g., Rosenfeld, 1996, Section 2.3), these effects could at least in part be mediated through effects of previous mention on a word's contextual predictability (for evidence, see J. R. Heller & Pierrehumbert, 2011; for discussion, see Kahn & Arnold, 2012).

While phonological weakening or deletion has been studied extensively, less is known about phonological insertion. One example comes from optional epenthesis. In epenthesis, speakers insert a reduced vowel into a consonant cluster (e.g., fil*e*m). Epenthesis enhances the signal and reduces syllable complexity compared to what would be expected under a faithful realization of the citation form. In a corpus study on conversational Dutch, Tily and Kuperman (2012) found that speakers were less likely to insert the schwa into words that were contextually predictable.

How big are the effects of contextual predictability? Bell *et al*. (2003) find that the top and bottom 5% most predictable instances of English function words (such as *the*, *I*, etc.) differ in their duration by about 20-30 ms, out of a mean duration of about 100 ms. For content words, the most predictable instances of words are sometimes more than 100 ms shorter than their least predictable instances (Demberg, Sayeed, Gorinski, & Engonopoulos, 2012, p. 364). These effect sizes mean that predictability effects tend to be somewhat smaller than, though sometimes comparable to, durational lengthening associated with differences in linguistic structure or meaning (such as contrastive prosodic accents, Berkovits, 1994; phrase final lengthening, Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991). At the same time, these effect sizes imply that at least some probabilistic reduction is clearly perceptible (cf. Beaver, Clark, Flemming, Jaeger, & Wolters, 2007, who report detection of 6 ms durational differences).[4] Indeed, although this is an area that has received surprisingly little attention, there is evidence that the phonetic reduction associated with contextual predictability does affect intelligibility (Bard & Anderson, 1983, 1994; see also Buz, 2016, Ch. 4 for related discussion).

In summary, there is ample evidence that a word's contextual predictability tends to be correlated with its reduction (for further references, see Ernestus & Warner, 2011; Ernestus, 2014). Here we have focused on evidence from English. This reflects the status of the field, with the majority of existing research on phonetic reduction coming from English and typologically related languages (e.g., Dutch: Kuperman, Pluymaekers, Ernestus, & Baayen, 2007; Pluymaekers *et al*., 2005a; van Son & van Santen, 2005; French: Bürki *et al*., 2011; Pellegrino, Coupé, & Marsico, 2011; Torreira & Ernestus, 2009; Italian: Pellegrino *et al*., 2011; Spanish: Torreira & Ernestus, 2012), with only a handful of comparable studies on other languages (e.g., Cantonese: Zhao & Jurafsky, 2009; Japanese: Kurumada, 2011; Vietnamese, among others: Pellegrino *et al*., 2011). By taking advantage of language-specific properties, future studies on languages other than English hold great promise for the study of probabilistic reduction (for the critical importance of cross-linguistic evaluations of psycholinguistic theory, see also Jaeger & Norcliffe, 2009; Norcliffe, Harris, & Jaeger, 2015). For example, Zhao and Jurafsky (2009) investigated the effects of frequency on the realization of lexical tone in Cantonese. They found that frequency influences pitch production: low frequency words are produced with tone contours that are more distinct from each other. Paralleling phonetic and phonological reduction in English, Cantonese speakers thus tend to produce more reduced—or less distinguishable—signals for contextually more expected—and thus less informative—message components.

## *Morphological contraction and omission*

Effects resembling probabilistic phonetic reduction have been observed in speakers' preferences between near-meaning equivalent morphological forms. For example, Frank and Jaeger (2008) investigate morphological contraction in American English conversational speech. Specifically, they focus on NOT (e.g., *isn't* versus *is not*), auxiliary BE (e.g., *he's* versus *he is*), and auxiliary HAVE (e.g., *I've done that* versus *I have done that*). They find that the rate of morphological contraction increases with the predictability of the meaning of the contractible element (e.g., NEGATION in *isn't* versus *is not*). This effect holds while controlling for potentially confounding factors such as speech rate, the type of host word preceding the contractible element, and the complexity of the material following the contractible element (see also Frank & Jaeger, 2008). These effects are confirmed by other studies on morphological contraction in conversational English (Bresnan & Spencer, 2016; Bybee & Scheibman, 1999).

More recent research has investigated alternations in which a bound morpheme can be either mentioned or omitted under near-meaning equivalence (Kurumada & Jaeger, 2015; Norcliffe & Jaeger, 2014). For example, Kurumada and Jaeger (2015) investigate optional case-marking in Japanese. Like in other case-marking languages, Japanese has case-marking morphology on the arguments of the verb that encode the grammatical function assignment. For example, the direct object of a transitive verb is marked with the suffix *-o*. Case-marking is important in understanding Japanese sentences, since Japanese has flexible word order, allowing both subject-before-object and object-before-subject ordering in transitive sentences. Unlike languages in which case-marking is obligatory (e.g., German), informal spoken Japanese allows speakers to omit the case marker without loss of near-meaning equivalence (see also Fry, 2001). In fact, case omission is frequent in informal Japanese (e.g., up to 51% of object case markers are omitted, Fry, 2001). In a spoken recall study, Kurumada and Jaeger find that speakers are more likely to omit the direct object case marker *-o* when the sentence makes the intended grammatical function assignment contextually predictable (e.g., *grandma* is more likely to be case marked in *The doctor sued the grandma* than in *The doctor treated the grandma*). Related corpus-based research has found that the rate of case-marking depends on how typical an argument is for the grammatical function it carries in the sentence (e.g., for Japanese: Fry, 2001; Korean: H. Lee, 2006; for further evidence from artificial miniature language learning, see Fedzechkina, Jaeger, & Newport, 2012, Fedzechkina, Newport, & Jaeger, 2016). For example, in conversational Korean, which also has optional case-marking, definite subjects are less likely to be case marked than indefinite subjects, whereas definite objects are more likely to be case marked than indefinite objects (H. Lee, 2006, Table 4). Since subjects are more likely to be definite than objects are, these findings suggest that case is more likely to be omitted when the meaning it encodes is predictable from context (for discussion, see Kurumada & Jaeger, 2015).

Another example comes from a recent study on optional head-marking in Yucatec Maya (Norcliffe, 2009; Norcliffe & Jaeger, 2014). In head-marking languages, grammatical function assignment and other information is encoded through bound morphology or clitics attached to the verb (rather than the arguments, as in the case of case-marking). In Yucatec some of this morphology is optional in certain environments. Norcliffe and Jaeger (2014) provide evidence that this optional morphology follows similar patterns as described for case-marking in Japanese above.

In sum, existing cross-linguistic evidence suggests that speakers' preferences in morphological reduction environments (i.e., contraction and omission) are affected by contextual predictability in ways that are at least qualitatively similar to phonetic reduction. However, compared to phonetic and phonological reduction, relatively little is known about the pressures driving optional morphological contraction and omission. Research on morphological production in morphologically rich languages seems a particularly promising venue for future work.

## *Omission of optional function words*

Probabilistic reduction has also been documented for morphologically free function words. For example, English allows the omission of complementizer *that*, as in sentences like *She certainly knew (that) this was a required test* (Elsness, 1984; Huddleston & Pullum, 2002). This phenomenon is sometimes referred to as optional complementizer *that*-mention or *that*-omission. Speakers are more likely to produce the optional complementizer *that*, when the complement clause is less predictable given the matrix verb (e.g., *knew* in the example above). This effect has been observed in conversational speech (Jaeger, 2010) as well as production experiments (e.g., in written sentence completion, Garnsey, Pearlmutter, Myers, & Lotocky, 1997, Table 5; spoken or written recall, Ferreira, 2008; Jaeger & Grimshaw, 2013).

Optional function word omission is also observed in certain types of relative clauses. For example, in Standard American English, both finite non-subject-extracted non-pied-piped relative clauses (e.g., *That's the way (that) it is done*) and passive subject-extracted relative clauses (e.g., *These are the type of people (who are) not taken seriously*) allow similar omissions. For these environments, too, speakers have been found to be more likely to omit the optional function words the more predictable the constituent they introduce is in context (Jaeger, 2010, 2011; Wasow, Jaeger, & Orr, 2011; see also Melnick, 2011; Wiechmann, 2015).

In sum, speakers' preference to mention or omit optional function words seems to exhibit sensitivity to contextual predictability in ways that resemble phonetic reduction. However, beyond *that*-omission, the sensitivity of optional function word omission to contextual predictability has remained under-explored. Alternations similar to optional complementizer *that* exist in other languages (e.g., in Danish), though omission is sometimes accompanied by constituent order changes (e.g., in German). English, too, contains a number of additional environments that support optional omission of function words, such as the omission of *to* after verbs like *help* (Rohdenburg, 2004) or in the DO-BE construction (e.g., *all I want to do is (to) go to work*, Flickinger & Wasow, 2013). Additional examples are observed in non-standard varieties of American English, such as optional copula omission in African American Vernacular English (e.g., *You done yet?*; Bender, 2000, p. 85) or relativizer omission in subject-extracted relative clauses in, for example, the English of the British Isles (e.g., *And there were a wee alarm clock sat on the window*; Tagliamonte & Smith, 2005, p. 87).

It is thus an open question whether the effects of contextual predictability observed in research on *that*-omission in Standard American English will generalize to these similar phenomena and across languages. Preliminary evidence comes from ongoing research on the DO-BE construction (Wasow, Levy, Melnick, Juzek, & Zhu, 2015). Wasow and colleagues find that speakers are more likely to omit *to* in the DO-BE construction in lexical contexts that frequently co-occur with the DO-BE construction.

## *Reduction and omission of referring expressions*

Another domain in which languages typically provide multiple near meaning-equivalent forms with more or less reduced signals is referring expressions. For example, in many contexts speakers can choose between a pronoun (e.g., *he*), name (e.g., *John*), or a full lexical noun phrase (e.g., *a colleague of mine*) to refer to the same referent.[5]

It has long been hypothesized that the choice between these different ways of encoding a reference depends on the referents "accessibility" in context (e.g., Ariel, 1999; Givón, 1983). This includes several factors that make referents more predictable (Arnold, 1998, 2010). For example, previous mention of a referent makes it more likely that it will be referred to in subsequent utterances. Previous mention also makes it more likely that a more reduced form is chosen (Bard *et al*., 2000 and references therein). Moreover, the probability that a previously mentioned referent is referred to again decreases with increasing distance from its last mention. Similarly, the preference for a pronoun over a longer referring expression decreases with increasing distance from the last mention of a referent (Arnold, 1998; Arnold, Bennetto, & Diehl, 2009; as summarized in Arnold, 2010, p. 190).

Recent work has more directly assessed the effect of contextual predictability on the realization of referring expressions, paralleling research on probabilistic phonetic reduction. Tily and Piantadosi (2009) employed a type of Shannon guessing game (Shannon, 1951) to obtain estimates of the contextual predictability of over 2,000 references in a newspaper corpus. In their version of the Shannon guessing game, raters saw story fragments up to the next referring expression. Their task was to guess which of the previously introduced referents (or possibly a new referent) the next expression would refer to. Almost 500 raters provided a total of over 70,000 guesses. This made it possible to calculate estimates of the contextual predictability of the actual references made in the corpus. Tily and Piantadosi found that writers had indeed been more likely to use longer linguistic forms (e.g., names rather than pronouns) when the intended reference was less expected given the preceding context. This effect held beyond the effects of previous mention and other previously documented effects (for related results, see also Rohde & Kehler, 2014).

Mahowald, Fedorenko, Piantadosi, and Gibson (2013) investigated speakers' preference between full and reduced lexical forms with the same meaning, such as *mathematics* and *math*. Mahowald and colleagues found that speakers' preference for the shorter form increases with the contextual predictability of the concept encoded by either form. In a corpus study, the average informativity (measured as Shannon information) of long forms was significantly higher than for short forms suggesting that short forms tend to be used in contexts where they conveyed less information. In a sentence completion study, Mahowald and colleagues further found that participants chose the short form for sentences with supportive contexts (e.g., *Susan loves the apes at the zoo, and she even has a favorite …*) as compared to non-supportive contexts (e.g., *During a game of charades, Susan was too embarrassed to act like a …*). This preference closely mirrors the preference observed for contractible auxiliaries and negation (Bresnan & Spencer, 2016; Bybee & Scheibman, 1999; Frank & Jaeger, 2008).

A similar preference to produce reduced linguistic signals for contextually more predictable referents is also observed for optional argument omission (Kravtchenko, 2014; Resnik, 1996). In certain lexical environments, speakers of English can decide to omit an entire argument (e.g., *the semi-finals* in *Germany lost (the semi-finals)*), while maintaining near meaning-equivalence. In his seminal corpus study, Resnik found that verbs that contained more information about the types of arguments they take, thereby making the arguments following them (on average) more predictable, also are associated with a higher rate of argument omission (Experiment 4, Resnik, 1996).

Recent work on optional subject omission in Russian builds on these results (Kravtchenko, 2014). While considered non-standard or ungrammatical in English, many languages allow omission of contextually inferable subjects, sometimes referred to as pro-drop (Dryer, 2013). Using the version of the Shannon guessing game developed by Tily and Piantadosi (2009), Kravtchenko (2014) obtained estimates of the contextual predictability of over 700 subject noun phrases from a Russian corpus. Paralleling the results for the realization of referential expressions in English, Kravtchenko found that Russian subjects are more likely to be omitted when they are contextually predictable.

## *Reduction beyond the level of the clause*

The majority of psycholinguistic research has focused on linguistic encoding at the level of the clause or below. A few more recent studies have begun to investigate reduction beyond the clause. For example, Asr and Demberg (2015) investigated the realization of coherence relations in English (see also Asr & Demberg, 2012). Simplifying somewhat, coherence relations are discourse relations between propositions. Asr and Demberg (2015) focused on the so-called *Chosen Alternative* relation and the coherence marker *instead* in environments in which it is optional (e.g., *They didn't panic during the first round of selling. (Instead,) they sold into the strength, which kept the market orderly*). Asr and Demberg found that *instead* was more likely to be omitted in the presence of a contextual cue to the Chosen Alternative relation (but see Anibel, 2010, for a failure to find such effects for other types of coherence relations).

Another environment in which speakers have the choice between providing more or less linguistic material to encode a near meaning-equivalent message was investigated by Gallo and colleagues (Gallo, 2011; Gallo, Jaeger, & Furth, 2010; Gallo, Jaeger, & Smyth, 2008). For example, Gallo *et al*. (2008) had speakers participate in a version of the Map Task (A. H. Anderson *et al*., 1991; see Pardo, this volume, for a description). Speakers instructed another (confederate) participant to replicate on their screen a specific arrangements of objects seen only by the speaker. Gallo and colleagues coded whether speakers used one or two sentences to convey the same message. For example, participants could say *Move the triangle to Central Park* or use a more verbose message like *Take the triangle. Now move it to Central Park*. Gallo and colleagues found that speakers were more likely to split the message across two clauses when the object (e.g., *the triangle*) consisted of less predictable words (for similar evidence from Spanish, see Gallo *et al*., 2010). These effects held beyond effects of previous mention, which is known to be correlated with the choice between pronoun versus lexical NPs (cf. Tily & Piantadosi, 2009).

Of the areas summarized here, production planning (including preferences regarding reduction) beyond the clause-level is probably the least understood. Further work is required to see whether the tentative evidence summarized here will confirm that principles similar to those observed in phonological, lexical, and syntactic reduction also operate during planning of larger linguistic chunks.

## Summary and open questions

Language provides speakers with an astonishing degree of flexibility in the linguistic encoding of messages. Many of the options available to speakers differ in the amount of signal produced by the speaker. Across all stages of production summarized here, speakers' preferences between different ways of realizing the same message seem to be affected by a similar bias, reflected in a correlation between contextual predictability and reduction. More specifically, it seems that it is the predictability of a linguistic form or *message component* (roughly, part of the meaning a speaker wishes to convey) that correlates with a preference for shorter linguistic forms at the next lower level and more reduced linguistic signals. For example, the predictability of NEGATION following a lexical context (e.g., *President Clinton did* …) correlates with an increased preference for morphological contraction (i.e., saying *President Clinton didn't* … rather than *President Clinton did not* …, Frank & Jaeger, 2008). Similarly, it seems to be the predictability of a complement clause that correlates with an increased preference to omit the relativizer *that* (Jaeger, 2010) and the predictability of a lemma that correlates with the reduction of its word form (Aylett & Turk, 2004; Jurafsky *et al.*, 2001). In this context, a particularly intriguing piece of evidence comes from research on homophones, such as *time* and *thyme*. While *time* and *thyme* have the same phonological citation form, the actual realization of the two words tends to differ subtly (Gahl, 2008). Speakers tend to produce the more frequent lemma (*time*) with a more reduced speech signal, compared to the less frequent lemma (*thyme*). To the best of our knowledge, comparable work on the effects of contextual predictability on homophone pronunciation has yet to be conducted. Still, this type of effect suggests that it is at least partly the predictability of a message component (in this case the lemma or its meaning) that drives the extent to which its realization in the linguistic signal is reduced (see also Jaeger, 2006, Study 6).

While the inverse correlation between predictability and linguistic signal is now firmly established, many questions remain about the nature of this relation. The perhaps most pressing questions regard the processes underlying probabilistic reduction and, in particular, the relation between production planning and the realization of the linguistic signal. Before we address these questions in the second part of this chapter we briefly summarize outstanding empirical questions about probabilistic reduction. One question that deserves further attention is the relation between reduction at different levels of linguistic encoding (e.g., phonetic vs. syntactic reduction). Simply put, what determines the level of linguistic encoding at which speakers reduce or enhance the signal? This question has received some attention in research on phonetic reduction and phonological deletion (e.g., Bürki, Ernestus, & Frauenfelder, 2010; Bürki *et al.*, 2011; Hanique, Ernestus, & Schuppler, 2013; Torreira & Ernestus, 2011). For example, some cases of omission might be better understood as extreme cases of gradient phonetic reduction, while others are better understood as originating in categorical phonological representations.

Another open question is what types of cues affect probabilistic reduction. The majority of previous research on probabilistic reduction has focused on the immediately surrounding lexical context. For example, for phonetic reduction most research has estimated the word's predictability based on its surrounding trigram context (e.g., Aylett & Turk, 2004; Bell *et al.*, 2009, 2003; Gahl *et al.*, 2012; van Son & Pols, 2003; van Son & van Santen, 2005). Arnon and Cohen Priva (2013) find that the predictability of the final word in a 4gram (e.g., *tea* in *a cup of tea*) is correlated with phonetic reduction, even after bi-, tri-, and unigram frequencies are accounted for (see also Arnon & Cohen Priva, 2014; Demberg *et al.*, 2012). Similarly, most research on reduction at higher levels of linguistics encoding has employed local lexical cues (e.g., Frank & Jaeger, 2008; Jaeger, 2010; Mahowald *et al.*, 2013; Resnik, 1996).

There are, however, also some studies that have found less local or more abstract cues to affect reduction. For example, phonetic reduction has been found to be correlated to the word's predictability given its semantic (Sayeed, Fischer, & Demberg,

2015) or syntactic context (or larger *n*grams, Demberg *et al.*, 2012; Kuperman & Bresnan, 2012; Kurumada, 2011; Tily *et al.*, 2009; but see Rose, 2017: Ch. 4).[6] Less local cues have also been found to affect the omission of optional function words (syntactic context, Jaeger, 2006, Study 5; Levy & Jaeger, 2007; Wasow *et al.*, 2011) as well as the reduction of referring expressions (cloze completions, Kravtchenko, 2014; Tily & Piantadosi, 2009), although some of these studies have not tested whether the same effects could be attributed to more local cues.

A closely related question is whether different types of cues are weighted differently depending on the level of linguistic encoding (e.g., phonological versus morphological contraction). This would arguably be expected under most accounts discussed below. Even accounts of linguistic encoding that assume that information from lower levels can affect earlier stages of production generally assume that these influences are weaker than influences from the current or earlier stages of production (e.g., Dell, 1986; Dell, Chang, & Griffin, 1999; Janssen & Caramazza, 2009). For example, segmental phonological properties generally only weakly affect syntactic preferences (Jaeger, Furth, & Hilliard, 2012a; McDonald, Bock, & Kelly, 1993). *Supra*segmental phonological preferences, on the other hand, have been found to affect syntactic production. For example, speakers prefer to insert optional function words or reorder constituents so as to avoid adjacent stressed syllables (Anttila, Adams, & Speriosu, 2010; Jaeger *et al.*, 2012a; M.-W. Lee & Gibbons, 2007). Similar asymmetries in the factors that drive variation have been observed between the phonetic reduction of segments and their omission (for results and discussion, see Bürki *et al.*, 2011; Hanique *et al.*, 2013). Whether similar asymmetries are reflected in what cues affect probabilistic reduction is a question for future research (for preliminary results, see an unpublished study by Jaeger, Snider, Staum, & Jurafsky, 2006, who compared the phonetic reduction and optional omission of complementizer and relativizer *that*).

Another questions is whether and how speakers integrate multiple cues to the same target (e.g., the same word). For example, does such integration follow similar principles that have been observed in comprehension, where comprehenders seems to be able to integrate multiple sources of information (e.g., Hare, McRae, & Elman, 2004; MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995)? To the best of our knowledge, there is so far no published work that addresses this question. A few studies have compared the effect of predictability (or surprisal) estimates based on different types of cues (e.g., Demberg *et al.*, 2012, p. 364; Sayeed *et al.*, 2015). But these studies have not directly compared the *objective* information contained in these cues to their relative importance in the *subjective* language models that speakers implicitly draw on during linguistic encoding. Preliminary evidence comes from an unpublished study on phonetic reduction in speech (Post & Jaeger, 2010). Post and Jaeger integrated multiple lexical and syntactic cues into a single estimate of a word's predictability. They found that both types of cues contributed to a word's phonetic reduction and that they did so proportionally to their contribution to the word's predictability. If confirmed by future work, results like these would suggest that probabilistic reduction draws on multiple contextually available cues, weighted by their relative informativity (see also Jaeger, 2006, Studies 3 and 4, for related evidence for optional complementizer and relativizer *that*).

## Theoretical positions

Psycholinguistic accounts of probabilistic reduction tend to come in three broad flavors: production ease, communicative, and representational accounts. Production ease accounts attribute variation in speakers' preferences to the demands of incremental linguistic encoding. Below we discuss three related classes of proposals about *how* production ease affects linguistic encoding. Following that, we discuss accounts of linguistic reduction that refer to communicative goals. This includes a discussion of research on *audience design*. We also discuss more recent communicative accounts that either draw on information theoretic considerations (cf. Shannon, 1948) or the concept of rational (J. R. Anderson, 1990) or boundedly rational cognition (e.g., Simon, 1990).

Production ease and communicative accounts share a focus on online processes that affect production as it is unfolding. This contrasts with *representational* accounts, which have focused on changes in the phonetic representations of words over longer periods of time (e.g., the lifetime of a speaker or even generations of speakers). The majority of psycholinguistic work on reduction and omission has interpreted speakers' preferences in alternations as providing a window into the mechanisms underlying language production, thereby more or less explicitly assuming the former (e.g., Arnold *et al.*, 2012; Baese-Berk & Goldrick, 2009; Bard *et al.*, 2000; Ferreira & Dell, 2000; Gahl *et al.*, 2012). In research on speech production, however, phonological and phonetic reduction is often described as the result of changes to phonological representations (e.g., Bybee & Hopper, 2001; Kohler, 1990; Pierrehumbert, 2001, 2002; Wedel, 2006; Zipf, 1929; for additional references, see Ernestus, 2014). Following our discussion of production ease and communicative accounts, we turn to this third type of account of reduction mentioned above, representational accounts. We discuss the relation between such offline accounts and online accounts of reduction.

Before we turn to these different accounts, we begin with an important caveat.

### *Production ease versus communicative goals: Mutually exclusive?*

Although it is helpful for the purpose of exposition to group accounts of reduction into broad classes of competing positions, production ease and communicative accounts are arguably better seen as defining a continuum of perspectives. For example, some communicative accounts do not argue against the idea that the resource demands inherent to linguistic encoding affect speakers' production preferences. Rather, speakers' preferences are assumed to *also* be affected by communicative considerations. Specifically, a long-standing idea holds that language production is subject to competing pressures—on the one hand, speakers want to achieve their communicative goals, on the other, they have limited resources (e.g., planning time, memory capacity) to achieve these goals (Zipf, 1949; see also e.g., Jaeger, 2006, 2013; Lindblom, 1990a, 1990b). These types of accounts thus do not predict that speakers will always be able to (or even intend to) rank their communicative goals above production ease.

Similarly, production ease accounts are typically not intended to completely rule out communicative effects on linguistic encoding. Such a position would be untenable in light of existing evidence. Rather, only generic knowledge about common situations is assumed to affect language production (e.g., such as talking louder in a noisy place, Lombard, 1911; child-, hearing impaired-, or foreigner-directed speech, Kuhl *et al.*, 1997; Stern, Spieker, Barnett, & MacKain, 1983; Uther, Knoll, & Burnham, 2007; Picheny, Durlach, & Braida, 1986). In this view, *specific* knowledge about the current situation and interlocutors, including perspective taking, is assumed to have no effect on language production (cf. Dell & Brown, 1991). Later work has further softened this claim to suggest that listener-specific[7] audience design is possible, but very resource-demanding and thus easily abandoned (e.g., Horton & Keysar, 1996; Wardlow Lane & Ferreira, 2008). Some accounts are further only intended for specific aspects of language production. For example, some proposals hold that lexical encoding can be subject to audience design, whereas decisions during phonological encoding and articulation are unaffected by listener-specific audience design (Arnold, 2008; Bard *et al.*, 2000; but see Buz *et al.*, 2016). Other proposals hold that speakers can avoid ambiguity when it is apparent *prior* to linguistic encoding (e.g, when referring to a small bat, when there is also a large bat on the screen), but not when the ambiguity only becomes apparent after lexical retrieval (e.g., when referring to a baseball bat, when there is also an animal bat on the screen, Ferreira, Slevc, & Rogers, 2005). Similarly, some proposals hold that grammatical encoding is unaffected by listener-specific audience design (Ferreira, 2008; Ferreira & Dell, 2000; MacDonald, 2013; but see Jaeger, 2013).

With these caveats in mind, the remainder of this section sets out to isolate core differences in theoretical perspectives. We do so because it is those differences in focus or perspective that often end up driving researchers' decision to conduct a particular study. During our discussion of communicative accounts, we return to questions about the specificity of audience design, and review the available evidence. We refer to Gambi & Pickering (this volume) for additional discussion of these issues.

## *Production ease*

With these clarifications in mind, we now introduce three types of production ease accounts. These accounts differ in whether they attribute speakers' preferences—including reduction—to the planning of previous, the current, or following material.

*Production ease: Planning of upcoming material affects the realization of current material* One influential proposal is *availability-based production* (Bock, 1987; Ferreira & Dell, 2000). Availability-based production holds that speakers prefer linguistic forms that let them articulate whatever material is fully planned, while maintaining grammaticality (Ferreira, 1996, 2008; Ferreira & Dell, 2000). Language production is, in this sense, greedy, presumably because the attentional and memory resources available for sentence planning are limited. By sending completed linguistic plans to articulation, resources are freed up and articulation continues. There is clear evidence that some such principle is at work during sentence production. Availability-based production is reflected in a variety of strategies that speakers seem to employ in order to avoid suspension of articulation. This includes the insertion of filled pauses (Clark & Fox Tree, 2002; Shriberg, 1996), restarts (Clark & Wasow, 1998), and constituent reordering (for review, see Jaeger & Norcliffe, 2009). For example, speakers tend to produce more easily retrievable word forms and word forms associated with more easily retrievable concepts or constituents earlier in a sentence (Arnold, Losongco, Wasow, & Ginstrom, 2000; Bock, 1982, 1987; Branigan, Pickering, & Tanaka, 2008; Bresnan, Cueni, Nikitina, & Baayen, 2007; Ferreira, 1996; Ferreira & Yoshita, 2003; Kempen & Harbusch, 2004; Rosenbach, 2008; Tanaka, Branigan, McLean, & Pickering, 2011).

Similarly, there is evidence that availability affects reduction. Most of this evidence comes from research on phonetic reduction and research on the omission of optional function words. For example, speakers slow down their speech rate before points of production difficulty (e.g., complex, infrequent, or novel words, Fox Tree & Clark, 1997; Watson, Buxó-Lugo, & Simmons, 2015). Speakers are also more likely to produce optional complementizer or relativizer *that* before difficult-to-retrieve clause onsets (Ferreira & Dell, 2000; Jaeger, 2010; Jaeger & Wasow, 2006; Race & MacDonald, 2003; Roland, Elman, & Ferreira, 2006; Temperley, 2003). These types of findings show that the availability of *upcoming* material can cause reduction or omission preceding that material.

One open question is whether such planning effects can also account for the link between the contextual predictability of a linguistic form and its own realization (rather than the realization of preceding forms). For example, consider the case of phonetic reduction. While the majority of studies has not controlled for the availability of upcoming material (e.g., Aylett & Turk, 2004, 2006; Bell *et al*., 2009, 2003; Gahl *et al*., 2012), there is, in fact, some evidence that the effect of contextual predictability on phonetic reduction is independent of the availability of upcoming material (Kidd & Jaeger, 2008; Post & Jaeger, 2010; Watson *et al*., 2010). Similar evidence exists for morphological reduction (Frank & Jaeger, 2008) and morphological or optional function word omission (Jaeger, 2006, 2010; Kurumada & Jaeger, 2015; Norcliffe & Jaeger, 2014).

Could availability nevertheless account for probabilistic reduction? To some extent, this is a question of *granularity* (i.e., at what level of linguistic representation availability is taken to apply). One recent proposal holds that probabilistic phonetic reduction is driven by the incremental availability of the word's *segments* (Watson *et al*., 2015). Indeed, when a word is contextually predictable, so is—on average—the sequence of its segments. Thus, if the contextual predictability of a segment contributes to its availability, it seems plausible that a segment-level availability account could explain word-level probabilistic phonetic reduction. This proposal does, however, face at least two challenges. The first challenge is empirical: we know of no studies that directly test whether word-level effects of contextual predictability can be reduced to segment-by-segment predictability effects. The second challenge is integrating this view with existing theories of lexical production: standard models of lexical production assume that the units at the interface between phonological encoding and articulation are syllables (Dell, 1986; Levelt, 1989). That is, articulation is assumed to proceed syllable-to-syllable, not segment-to-segment. It is thus unclear how standard models could explain probabilistic phonetic reduction of mono-syllabic words (which is observed, e.g., Bell *et al*., 2003; Gahl *et al*., 2012).

*Production ease: Planning of material affects its realization* An alternative type of production ease account holds that the planning of a linguistic unit (e.g., a word) is directly reflected in its *own* realization (e.g., Arnold *et al*., 2012; Baese-Berk & Goldrick, 2009; Bard *et al*., 2000; Goldrick, Vaughn, & Murphy, 2013; J. R. Heller & Goldrick, 2014). Unlike availability-based production, which links reduction to the planning of upcoming material, these proposals thus link reduction of a linguistic form to the planning of that form. Whereas availability-based production has been investigated at the phonetic, lexical, and syntactic level (see references above), the alternative account has so far only been applied to the link between lexical planning and articulation (i.e., phonetic reduction). Within work on phonetic reduction, the idea that activation levels during phonological or lexical planning predict articulatory reduction is receiving increasing attention (e.g., see also Arnold *et al*., 2012; Ernestus, 2014; Gahl *et al*., 2012; Kahn & Arnold, 2012; Seyfarth, 2014; Watson *et al*., 2015). We thus focus on phonetic reduction.

A specific instance of this proposal is the *competition* account introduced by Baese-Berk & Goldrick (2009). Baese-Berk and Goldrick investigated the articulation of words with minimal pair neighbors (e.g., the articulation of the/p/in *pin*, which has the minimal pair neighbor *bin*, compared to the/p/in *pipe*, which lacks the minimal pair neighbor *bipe*). The existence a minimal pair neighbor is assumed to lead to competition during lexical planning. This, in turn, is taken to mean that the target word will reach higher activation before being selected for articulation, and this increased activation is assumed to be correlated with hyper-articulation of the target word. In the words of Baese-Berk and colleagues: "[t]he higher activation level for words […] will lead to more active phonetic representations and consequently more extreme articulatory realizations" (Baese-Berk & Goldrick, 2009, p. 531).

This competition account makes two predictions that have been investigated. First, it predicts the hyper-articulation of words with minimal pair neighbors. This prediction has received support from experiments on isolated word production (Baese-Berk & Goldrick, 2009; Fox, Reilly, & Blumstein, 2015; Peramunage, Blumstein, Myers, Goldrick, & Baese-Berk, 2011) and the acoustic realization of words in conversational speech (Wedel & Sharp, 2014). For example, Baese-Berk & Goldrick (2009) found that the/p/in in *pin* is hyper-articulated with longer voice onset timing (making it more clearly unvoiced and thus distinguishable from a/b/) compared to the/p/in *pipe*.[8]

Second, if a minimal pair neighbor receives contextual support, competition accounts predict further competition with the target word, thus leading to increased hyper-articulation. A number of studies have addressed this question for onset minimal pair neighbors (e.g., *pin-bin*). This work has consistently found hyper-articulation of the contrasting phonetic features (except where phonetic constraints impede hyper-articulation; see e.g., the lack of hyper-articulation for voiced onsets Goldrick *et al*., 2013). For example, speakers tend to hyper-articulate the /p/ in the word *pin* with longer voice onset timing when *pin* is displayed on the same screen as its neighbor *bin*, compared to when that neighbor is not displayed (e.g., Baese-Berk & Goldrick, 2009; Buz, Jaeger, & Tanenhaus, 2014; Kirov & Wilson, 2012; Seyfarth, Buz, & Jaeger, 2016).

The empirical picture is less clear when the critical contrast is on the vowel (Kirov & Wilson, 2012; Schertz, 2013) or a coda consonant (e.g., *coat-code;* De Jong, 2004; Goldrick *et al*., 2013; Seyfarth *et al*., 2016). For example, Goldrick *et al*. (2013) investigate voicing of plosive codas in word with minimal pair coda neighbors (e.g., *coat*, which has the neighbor *code*) and words without such neighbors (e.g., *rap*). After controlling for phonological confounds between their conditions, Goldrick and colleagues find no evidence for hyper-articulation of the voicing contrasts in the coda in words with minimal pair neighbors as compared to words without such neighbors (but see Seyfarth *et al*., 2016).

The possibility of an asymmetry between onset and rhyme minimal pairs is intriguing in light of studies on the effect on onset versus coda overlap on lexical planning. These works have generally found that onset overlap between adjacent words is associated with planning difficulty, whereas coda overlap is associated with facilitation (or at least, less difficulty, O'Seaghdha & Marin, 2000; Jaeger, Furth, & Hilliard, 2012b; Smith & Wheeldon, 2004; Rapp & Samuel, 2002; Sevald & Dell, 1994; Wheeldon, 2003). If future studies confirm that onset neighbors lead to production difficulty during lexical planning, whereas coda neighbors lead to facilitation, and that these differences are reflected in articulation, this would provide support that competition affects articulation. We consider such studies critical: without a clearer understanding of the conditions under which phonological neighbors inhibit or facilitate lexical planning, evoking these processes as an explanation of articulation (including reduction) risks ad-hoc meaning (for further discussion, see also Buz & Jaeger, 2016; Chen & Mirman, 2015; Gahl, 2015).

Future work on competition accounts will also need to elaborate on the linking function between activation levels during planning and reduction during articulation. Baese-Berk and colleagues link higher activation levels during planning to *hyper*-articulation. At the same time, accounts of probabilistic reduction would seem to assume the opposite: contextual predictability is generally assumed to lead to increased activation; yet, contextually predictable words are reduced compared to less predictable words. That is, accounts attributing the reduction of predictable words directly to activation levels during production planning implicitly or explicitly assume that higher activation leads to *hypo*-articulation (this assumption is made explicit in, e.g., Arnold, 2008, p. 506, referring to Arnold (1998); Arnold *et al*., 2012, p. 506, Bard *et al*., 2000, p. 17). If the competition account is to explain probabilistic reduction effects, future research will need to address these *prima facie* conflicting assumptions. One possibility is to distinguish effects on activation and effects on the activation threshold required for a word to be selected for articulation. Another possibility is that effects of *pre*-activation due to anticipatory effects (e.g., contextual predictability given preceding material) differ from effects of competition that arise later during lexical planning and phonological encoding.

*Production ease: Planning of previous material affects the realization of current material* A third type of production ease account attributes reduction to *previously* encountered production difficulty (Bell *et al.*, 2009). We will refer to this type of account as *compensatory reduction*. The compensatory reduction account shares with availability-based production models that it attributes reduction to processes that coordinate linguistic planning at higher levels (e.g., lexical planning) with the execution of the articulation in order to maintain fluency. For example, Bell and colleagues submit that phonetic reduction could in part be attributed to the coordination of articulation plans with the planning of the smallest prosodic units sent off to articulation (e.g., phonological words, Wheeldon & Lahiri, 1997): if a slow down is experienced during the planning of a prosodic unit, this is assumed to trigger a slow down in the execution of articulation for the following phrase in order to maintain information flow between these two stages of production (Bell *et al.*, 2009, p. 106).

By itself the compensatory reduction account is likely insufficient to explain the full range of phonetic reduction: probabilistic reduction is observed, even when local speech rate on preceding syllables is controlled for, thus suggesting that reduction is not exhaustively caused by compensatory changes in speech rate (e.g., for phonetic reduction, Gahl *et al.*, 2012; Kidd & Jaeger, 2008; for morphological reduction, Frank & Jaeger, 2008; for function word omission, Jaeger, 2006, 2010).

Another challenge that the compensatory reduction account faces is that there is as of yet no independent evidence for compensatory reduction. We know of no studies that have investigated effects of previous complexity on subsequent articulation, while holding constant the complexity of following material.

*Summary* We have summarized three production ease proposals that have been put forward to explain probabilistic reduction. While summarized as alternatives, it is possible that multiple production ease mechanisms jointly explain the observed range of probabilistic reduction phenomena. One challenge that any unified account of reduction in terms of production ease will have to address is how to account for reduction beyond the phonetic level. For example, it is unclear the *omission* of optional elements (e.g., optional arguments, adjuncts, or function words) could be a consequence of these elements being easy to produce (for similar arguments, see also Jaeger, 2006, 2010; Kurumada & Jaeger, 2015; Norcliffe & Jaeger, 2014).

## *Communicative accounts*

Communicative accounts hold that speakers' preferences in reducing or enhancing the linguistic signal are affected by their communicative goals. This includes a bias for robust—or even efficient—message transmission. This view is closely related, but not identical, to questions about *audience design*—the idea that "the speaker designs each utterance for specific listeners" (Clark & Murphy, 1982, p. 287).

Unlike this (strong) audience design hypothesis, communicative accounts are not necessarily committed to the claim that speakers design their utterances for a *specific* type of audience (cf. generic vs. listener-specific audience design, Dell & Brown, 1991). We return to this issue below.

Here we refer to message transmission in the most general sense. For example, we include the transmission of non-literal and social meaning (for related discussion, see also Pardo, this volume). The goal of communication in the broad sense intended here is to cause a change in the interlocutors' state of mind. Even under this broad definition, communication is arguably not the only function of language (cf. Chomsky, 2000; Jenkins, 2000), but it is a common and important function.

A variety of communicative accounts have been proposed. For example, some accounts focus on production effort, whereas others focus on fluency; some accounts focus on accurate message transmission, whereas others focus on the speed of accurate transmission. The scope of this chapter prevents us from discussing these differences. Instead, we focus on the idea—shared more or less explicitly between most communicative accounts—that the understanding of production preferences requires reference to the *goals* of language use (see also Tanenhaus & Brown-Schmidt, 2008), and in particular, the goal to transmit specific intended messages.

Central to communicative accounts as defined here is the idea that production ease and communicative are often in competition. We begin by briefly reviewing this notion, and a major challenge it faces. Then we turn to the notion of audience design. We conclude this section with a few key considerations for future work on audience design.

*Trading off production ease and communicative goals* According to one influential view, speakers' production preferences arise from the competition between, on the one hand, a bias for robust message transmission and, on the other hand, a bias for production ease or effort (an idea that goes back to at least, Zipf, 1949). The former bias favors better signal quality. The latter bias favors shorter and less clearly articulated signals. In such communicative accounts, probabilistic reduction arises because contextual predictability increases the *a priori* accuracy and speed of message transmission, thereby allowing speakers to produce less costly signal. For contextually less predictable message components, on the other hand, speakers are expected to provide better signals in order to facilitate comprehension (e.g., Jaeger, 2013; Kohler, 1990; Kurumada & Jaeger, 2015; Lindblom, 1990a; Piantadosi, Tily, & Gibson, 2011; see also Aylett & Turk, 2004; Gibson, Bergen, & Piantadosi, 2013; Jaeger, 2010; Levy & Jaeger, 2007; Pate & Goldwater, 2015, though not all of these accounts do necessarily commit to the exact trade-off described here).

Three points deserve clarification as they have led to frequent confusion in the literature. First, the goal of facilitating comprehension is not to be confused with altruism: as described here, facilitating comprehension serves the speaker's communicative goals. Second, successful communication as defined here is not identical to the facilitation of comprehension of a particular linguistic unit (such as a phone, word, or syntactic structure). Rather, the successful recognition of a linguistic unit is taken to be relevant only to the extent that it serves the speaker's communicative goals (for a discussion of this for phonology, see Hall, Hume, Jaeger, & Wedel, submitted). Third, as formulated here, communicative accounts share with production ease accounts that production is assumed to be inherently costly. Therefore signal enhancement is not always expected, but rather expected when its benefits outweigh its costs (for further discussion, see Kurumada & Jaeger, 2015; Pate & Goldwater, 2015). As a consequence, the communicative accounts described here do not argue against the idea that production planning affects production preferences (see, e.g., Zipf, 1949; Jaeger, 2006, 2013; Lindblom, 1990a, 1990b). In fact, we are not aware of any account that claims that production ease does not affect linguistic encoding.[9]

*What makes a signal 'better'?* A central challenge that communicative accounts still have to address is what constitute a better signal. Presumably, a better signal is one that makes it more likely that the speaker achieves their communicative goals (e.g., the goal to be understood). A common assumption is that less reduced signals generally facilitate comprehension. In general this assumption seems warranted. For example, reduced durations or reduced or missing segments can make a word harder to comprehend out of context (e.g., Ernestus, Baayen, & Schreuder, 2002; see also van de Ven, Tucker, & Ernestus, 2011). There is also evidence that probabilistic reduction affects the out of context intelligibility of words: words that are more predictable in context tend to be pronounced in a way that makes them less intelligible out of context (e.g., Bard & Anderson, 1983, 1994; for similar evidence from *that* omission, see Race & MacDonald, 2003).

It is important to note, however, that signal enhancement is not *always* expected to facilitate comprehension. According to ideal observer models of comprehension (Kleinschmidt & Jaeger, 2015; Levy, 2008; Norris & McQueen, 2008), the property of the signal that determines how much it facilitates comprehension is its likelihood under the intended message, i.e., how likely the signal is to have arisen under the intended message (as compared to alternative messages). This means that extreme hyper-articulation might sometimes impede comprehension, because it produces percepts that are not expected under *any* message (including the intended one). More generally, signal enhancement should facilitate comprehension only to the extent that it increases the relative probability that this signal is observed under the intended message. Since the distribution of signals for any given message is itself conditioned on the context (the same message is realized differently in different contexts), this implies that, in the right context, reduced signals should be processed at least as easily and accurately as less reduced signals. Most pertinent to the current discussion, reduction is to be *expected* in contexts where the message is predictable. In such context, producing an enhanced signal should not necessarily facilitate comprehension and might even do the opposite or trigger additional inferences (see, e.g., Arnold, Fagnano, & Tanenhaus, 2003; Arnold, Kam, & Tanenhaus, 2007; Kravtchenko & Demberg, 2015; see also the discussion of *distinctiveness* versus *formality* in Rischel, 1992, pp. 387–388). In line with this reasoning, there is evidence that more signal does not facilitate comprehension when it occurs in context where reduction would be expected (e.g., Caballero & Kapatsinski, 2014; Jaeger, 2007; Race & MacDonald, 2003).[10]

A related challenge for future work is empirical. Most studies on this question have tested communicative accounts against *production* data, rather than intelligibility measures. This work is generally based on the assumption that more distinct acoustic signals will likely facilitate comprehension. For the reasons outlined above, this is, however, not always to be expected. Future work will thus have to assess whether variations in pronunciation that are attributed to communicative goals indeed facilitate comprehension. In doing so, it will be important to keep in mind that speakers might be willing to increase their production effort even if this results in only a small increase in the *average* probability of successful communication (for further discussion see Buz, Tanenhaus, & Jaeger, 2016; Buz, 2016).

*The strong audience design hypothesis* Broadly speaking, audience design refers to any aspect of production that serves to adapt productions so as to increase the probability of successful communication. The debates in the literature, however, have often focused on a much stronger variant of this hypothesis. The strong audience design hypothesis holds 1) that speakers integrate *listener-specific* information during linguistic encoding and 2) that they do so *immediately* when the information becomes available.

Evidence regarding the first criterion—listener-specificity—is mixed and seems to depend on the level of linguistic encoding. Listener-specific information can be seen as knowledge that speakers have about their interlocutors' perspective (such as information about which referents in a display are visually accessible to an interlocutor). For phonetic reduction, several studies have failed to find evidence for listener-specific audience design (e.g., Arnold *et al.*, 2012; Bard *et al.*, 2000; Kahn & Arnold, 2015). Other studies, however, have found that phonetic reduction is at least in part sensitive to listener-specific audience design (Galati & Brennan, 2010). The evidence is similarly mixed for the reduction or enhancement of prosodic boundaries. Some studies have found that speakers can strengthen the cues to prosodic phrasing (thereby signaling syntactic structure), thereby facilitating comprehension of utterances with temporary syntactic ambiguities (Price *et al.*, 1991; Schafer, Speer, Warren, & White, 2000; but see Allbritton, McKoon, & Ratcliff, 1996; Snedeker & Trueswell, 2003). However, later studies have found that such strengthening takes place regardless of whether the utterance is actually ambiguous in the current context, thus arguing for listener-generic rather than listener-specific audience design (Kraljic & Samuel, 2005).

The picture seems to differ somewhat for lexical encoding and, specifically, the selection of referential expressions. For example, speakers' preference between pronouns (e.g., *he*), names (e.g., *John*), and full lexical noun phrases (e.g., *my colleague*) have been found to be at least partly affected by audience design considerations, though certainly not exclusively (Arnold, 2008, 2010; Arnold & Griffin, 2007; Bard *et al.*, 2000). Speakers also seem to be capable of taking into consideration interlocutors' perspective and knowledge state (e.g., whether an interlocutor knows the label for a referent), when choosing between different ways of referring to same entity, though again not without fail (Brennan & Hanna, 2009; Ferreira *et al.*, 2005; D. Heller, Gorman, & Tanenhaus, 2012).

For syntactic encoding, research on audience design has mostly focused on the question of whether speakers avoid temporary syntactic ambiguities that are known to lead to processing difficulty ("garden paths"). This work has returned little evidence that grammatical encoding is affected by ambiguity avoidance (see also Arnold, Wasow, Asudeh, & Alrenga, 2004; Ferreira, 2008; Wasow & Arnold, 2003; but see Haywood, Pickering, & Branigan, 2005; Roche, Dale, & Kreuz, 2010). For example, Temperley (2003) reports that writers are more likely to produce an optional relativer *that*, if it helps to avoid temporary ambiguity, but other production experiments (Ferreira & Dell, 2000) and corpus studies on written language (Jaeger, 2011; Roland, Dick, & Elman, 2007) and conversational speech (Jaeger, 2006) have failed to replicate this effect.

The second criterion for strong audience design has received comparatively little attention, although it is arguably critical in understanding failures to exhibit audience design. For example, in order to engage in effective audience design, speakers need to notice (though not necessarily become consciously aware of) that there is a potential comprehension problem. Speakers would also have to find a solution to this problem, such as determining which of several ways of encoding the intended message is most likely to successfully convey the intended meaning. This is likely to be computational costly, as has been recognized in arguments against strong audience design (Bard *et al*., 2000; Ferreira, 2008; Shintel & Keysar, 2009). What has received less appreciation in the literature, however, is that this makes audience design a problem of *learning* how to best communicate in a given situation (but see Buz *et al*., 2016; Galati & Brennan, 2010; Jaeger & Ferreira, 2013): when confronted with a novel situation—as is the case in, for example, most production experiments—speakers need to *infer* which variant is most likely to achieve their communicative goals (Buz *et al*., submitted). This requires inference under uncertainty about what is in common ground, what will cause difficulty for the comprehender, and so on. Thus, even a mathematically ideal speaker (in the sense of ideal observers; J. R. Anderson, 1991) that engages in audience design would not be expected to immediately arrive at an optimal solution to the communicative problem. Rather, speakers would be expected to rely on prior expectations and adapt or learn from the perceived communicative success or failure of previous utterances (see also Brennan & Hanna, 2009; Buz *et al*., 2016). While this prediction has, to the best of our knowledge, not received much attention in previous work, there is some evidence in support of it. For example, several recent studies have found that speakers adapt subsequent productions toward less reduced variants, if previous use of more reduced variants resulted in communicative failure (Buz *et al*., 2016; Roche *et al*., 2010; Schertz, 2013; Stent, Huffman, & Brennan, 2008). Such findings challenge accounts that attribute reduction solely to production ease.

*Future research on audience design* A few considerations deserve particular attention in future research on audience design. First, asking whether speakers are in principle capable of audience design is different from determining the conditions under which speakers actually engage in audience design. For example, some experiment that have failed to find audience design employed informed confederates as interlocutors (Dell & Brown, 1991). When the same experiments were repeated with interlocutors that believably benefitted from audience design, participants indeed engaged in audience design (Lockridge & Brennan, 2002). Interlocutors' believability, crucially, seems to be in large parts based on the absence of unnatural communicative *behavior*. For example, speakers and listeners seem to be exquisitely sensitive to the timing with which linguistic information is delivered, including speech rates, back-channels, etc. When interlocutors act in a way that violates these expectations, audience design effects are often suspended (for an excellent review, see Kuhlen & Brennan, 2013).

Second, in interpreting behavioral evidence for or against audience design, it can be helpful to ask whether audience design "makes sense" in a given situation—that is, whether it is a *rational* or *boundedly rational* behavior (see Buz *et al*., 2016; Jaeger, 2013; Kurumada & Jaeger, 2015). Specifically, it can be helpful to think about the *utility* of audience design in the current context: if, for example, listener-specific audience design increases attentional or memory demands, speakers might implicitly weigh these costs against the expected benefits of audience design. These expected benefits in turn depend on how likely audience design is to increase the probability of achieving one's communicative goals. This has several immediate consequences. One of them is believability, as discussed in the previous paragraph: if interlocutors act in a way suggesting that they will not benefit from audience design, these decrease the utility of audience design. More generally, when there is little incentive for successful communication, this decreases the utility of audience design (see also Tanenhaus, 2013). Even if there is an incentive to be understood, the utility of audience design is low when there is little *need* for it in the current context (e.g., when context provides sufficient information, Ferreira, 2008; Jaeger, 2010; Wasow & Arnold, 2003). Finally, thinking about utility requires consideration of the *available alternatives* to *a priori* audience design. For example, in some situations, it might be more effective for speakers to *repair* miscommunication after it has occurred (a form of a *posteriori* audience design), rather than to try to avoid miscommunication *a priori*.

## *Representational accounts*

Our discussion so far has focused on the processes and architecture involved in online language production. A largely independent line of research in phonetics and speech production has focused on how the phonological representations of words become reduced over time (e.g., Bybee & Scheibman, 1999; Bybee & Hopper, 2001; Bybee, 2006; Johnson, 1997; Pierrehumbert, 2001, 2002; Wedel, 2006), including changes to the phonological system (e.g., phonological mergers, Wedel, Jackson, & Kaplan, 2013; Wedel, Kaplan, & Jackson, 2013; lenition, Cohen Priva, 2012, 2015). The primary goal of this line of research is to understand how phonological representations change over multiple generations, thereby explaining language change.

This research has investigated how the alternating forms available to speakers synchronically arise through historical processes (though most empirical evaluations have rested on synchronic data; for exceptions, see Hay, Pierrehumbert, Walker, & LaShell, 2015; Kanwal, Smith, Culbertson, & Kirby, 2017; Sóskuthy & Hay, 2017). Consider for example, speakers' selection between alternatives like *math* and *mathematics*, discussed above. While psycholinguistic research has identified factors that affect speakers' preference between these alternatives in a given context (Mahowald *et al*., 2013), historical research has focuses on changes in the availability of these forms over time (for a review of this and related work, see Hall, Hume, Jaeger, & Wedel, submitted).

It might, therefore, be tempting to consider the two lines of research as entirely orthogonal. However, representational accounts also offer explanations of reduction during online language production (for discussion, see Baese-Berk & Goldrick, 2009; Ernestus, 2014; Seyfarth, 2014). To illustrate this point, we first provide some background. We start with a brief overview of exemplar-based and related models (Johnson, 1997; Pierrehumbert, 2001). We then explain how these representational assumptions can explain the reduction of frequent words and discuss whether a similar account could be applied to probabilistic reduction beyond frequency effects.

*The nature of phonological representations* Research suggests that phonological representations go beyond abstract knowledge about a language's phonology. Rather, listeners seem to have rich knowledge of the specific acoustic realizations of words previously experienced. Some influential accounts of phonology hold that, each time a word is heard, its perceptual input—or at least some subphonemic representation of that input—is stored as an "episode" (e.g., Goldinger, 1996) or "exemplar" (e.g., Johnson, 1997), along with the context it is experienced in. The phonological representation of a word is then taken to be the cloud of all the exemplars that were recognized as that word in previous experience. Later exemplar-based accounts revise this assumption slightly by introducing a process called *entrenchment*. Entrenchment 'compresses' the cloud of exemplars toward the center of the exemplar cloud (Pierrehumbert, 2001). In these accounts, listeners do not necessarily maintain *all* previously experienced exemplars.

Under any of these accounts, the phonological representation of a word is inherently distributional (the cloud of exemplars), word-specific, and context-sensitive. These accounts have received strong support from research on speech perception (for reviews, see Foulkes & Hay, 2015; Pardo & Remez, 2006; Weatherholtz & Jaeger, 2016). For example, context-sensitivity is evidenced by studies finding that the recognition of speech sounds is improved if it occurs in the same context that it was previously experienced in (e.g., Drager, 2011; Goldinger, 1996; A. Walker & Hay, 2011). There is also evidence for word-specificity (for review, see Pierrehumbert, 2002). For example, the typical realization of words can change over time in ways that go beyond what is expected by their abstract phonology (i.e., their phonological citation form). A striking example of this are differences in the realization of homophones (i.e., words that have the same phonological citation form, such as *time* and *thyme*). Homophones with higher frequency tend to have shorter duration, compared to their less frequent homophone partner (Gahl, 2008). More generally, words with higher frequency and higher *average* contextual predictability tend to have not only fewer phonological segments (Manin, 2006; Piantadosi *et al*., 2011; Zipf, 1949), but also shorter average duration even after segment counts and types are taken into account (Seyfarth, 2014). Similarly, the rate with which a word exhibits final consonant deletion (e.g., *t*/*d*-deletion) increases with the *average* predictability of the deleted segment even after the predictability in the current context is taken into account (Cohen Priva, 2008, 2015). This suggests that at least some phonetic and phonological reduction effects are encoded in the word-specific phonetic representations. As we explain next, exemplar-based and related accounts not only offer an explanation for such word-specific effects, but also can explain why higher usage frequency leads to reduced forms.

*Explaining the reduction of frequent words* Pierrehumbert (2002) extends exemplar-based accounts of speech perception to production and shows that such a model predicts reduced realizations of frequent words under very general assumptions (though some of these assumptions are called into question by a recent diachronic study of sound change, see Hay *et al*., 2015). Specifically, production is assumed to consist of sampling from the cloud of previously stored exemplars, while being biased toward more reduced forms. This bias toward reduced forms is taken to follow from a general bias to minimize effort (see also Zipf, 1949). These two assumptions alone predict that words would get further and further reduced, the more often they are used. This correctly predicts more reduced forms for more frequent words, but also incorrectly predicts that words should quickly become reduced to no form at all.

There thus needs to be a competing bias to prevent arbitrary degrees of reduction—very much like in the communicative accounts discussed above. In the exemplar-based account, however, this bias is assumed to operate during *comprehension*: listeners are taken to store the inputs they receive only if understood with sufficient certainty (see also Lindblom, 1990a; Ohala, 1988). Since more frequent words are more likely to be correctly recognized by chance, even with a deteriorated acoustic signal as compared to less frequent words, the stored exemplars corresponding to a frequent word will—over time—contain increasingly more reduced forms. This shifts the *average* realization of the word toward a more reduced form, but not arbitrarily so (for similar ideas, see Guy, 1996; Labov, 1994, pp. 580–588; Ohala, 1988).

This exemplar-based account thus provides a cognitive plausible explanation of the inverse correlation between usage frequency and phonological form. It shares with the communicative accounts discussed above, that competing communicative biases affect what speakers produce. However, whereas the communicative accounts discussed above place both competing biases within production, the exemplar-based accounts distributes the biases across the production-perception loop.

*Probabilistic reduction beyond the reduction of frequent words* The argument made above for frequency readily extends to the *average* effect of contextual predictability on word-specific phonetics (for discussion, see Seyfarth, 2014). But can an exemplar-based account explain the effects of contextual predictability beyond those average effects (i.e., the effects of the *current* context on that word's realization)? The answer to this question depends on what assumptions are being made about the *granularity* of the exemplars that are stored by listeners. For example, as discussed above, there is evidence from production and comprehension that language users' implicit linguistic knowledge includes knowledge of *n*gram statistics (Arnon & Snider, 2010; Arnon & Cohen Priva, 2013; Bannard & Matthews, 2008; for discussion, see also Arnon & Cohen Priva, 2014; Baayen, Hendrix, & Ramscar, 2013). If producers sample from their previous experience in a way that takes into account the lexical *n*gram context, an exemplar-based model could thus explain the correlation between a word's *n*gram predictability and its reduction. More generally, if whatever cues correlate with a word's phonetic reduction are assumed to be stored along with the perceptual exemplar, it would seem that an exemplar-based model can, in principle, account for probabilistic phonetic reduction.

There is, however, evidence that at least the type of production-perception loop described above, which operates *between* interlocutors (Pierrehumbert, 2001, 2003), is insufficient to explain probabilistic reduction. This evidence comes from studies finding that speakers adjust their productions on the perceived communicative success of their previous utterances. One example of such research are perturbation studies on articulation (e.g., Houde & Jordan, 1998; Tourville, Reilly, & Guenther, 2008; Villacorta, Perkell, & Guenther, 2007). In perturbation studies, speakers' productions are manipulated online and played back to the speaker with a non-detectable delay. For example, a speaker might be producing the word *pen*, but hear herself produce something more like *pin* (Frank, 2011; Tourville *et al.*, 2008). These studies provide evidence that speakers can rapidly adjust their articulation if they perceive their own productions to deviate from their intended production. There is also evidence for similar adaptation based on feedback from interlocutors that, critically, does not involve the target word (Buz *et al.*, 2016; Schertz, 2013; Stent *et al.*, 2008). This includes non-verbal indication from interlocutors that they did not successfully understand the speaker. In recent work we find that speakers increased the hyper-articulation of minimal pair onset neighbors when their interlocutor failed to understand them (Buz *et al.*, 2016, submitted) for evidence of similar adaptation to syntactic production, see Roche *et al.*, 2010). The production-perception loop between interlocutors cannot explain this effect. Instead, it seems that speakers can also learn from their own productions, adapting subsequent productions to be better suited for the current context.

*Summary and open questions* While insufficient as a sole account of phonetic reduction, the production-perception loop and exemplar-based accounts likely form part of an explanation of probabilistic phonetic reduction. It remains to be seen how these accounts can be extended to reduction at other levels of linguistics representation.

Beyond the specifics of exemplar-based accounts, the relation between offline and online accounts of probabilistic reduction and related phenomena has only relatively recently become a target of research (Baese-Berk & Goldrick, 2009; Buz, 2016; Buz *et al.*, 2016; Cohen Priva, 2008; Ernestus, 2014; Seyfarth, 2014). This is an area we consider of particular importance in advancing the understanding of linguistic reduction: representational and online accounts offer qualitatively different, though related and mutually compatible explanations for probabilistic reduction. Studying these different explanations in isolation of each other risks missing the bigger picture.

Finally, speakers' ability to learn from their own productions raises an interesting question: are such adaptations stored, so that they remain available for similar future occasions? Research on speech *perception* has found that listeners can learn expectations about talker-specific pronunciations (Kraljic & Samuel, 2007; Norris, McQueen, & Cutler, 2003) and even generalization across groups of talkers (e.g., Bradlow & Bent, 2008; Baese-Berk, Bradlow, & Wright, 2013; Weatherholtz, 2015; for review, see Weatherholtz & Jaeger, 2016). Once learned, talker- and group-specific expectations do not seem to be lost, but can be maintained over longer periods of time (e.g., Eisner & McQueen, 2006; Goldinger, 1996; Hay, Warren, & Drager, 2006; Kraljic & Samuel, 2006; Niedzielski, 1999; Strand, 1999; for review, see Foulkes & Hay, 2015; Kleinschmidt & Jaeger, 2015). It is thus possible that linguistic encoding is sensitive to this or similar talker-, group-, or situation-specific knowledge. Some support for this hypothesis comes from studies suggesting that speakers can maintain multiple phonetic representations for different dialects or registers (Clopper & Pierrehumbert, 2008). It remains to be seen whether speakers develop and store even more specific (talker-, situation-, or task-specific) representations. If this view receives further support, it offers a way to reconcile the presence of context-specific communicative effects on language production with the apparent limitations of listener-specific audience design: when confronted with a novel situation, speakers first need to learn effective communicative behaviors for that situation. Once learned and reinforced, these behaviors—such as targeted hyper-articulation of a specific phonological segment (Buz *et al.*, 2016; Kirov & Wilson, 2012; Schertz, 2013; Seyfarth *et al.*, 2016)—might be stored and thus more easily available in similar future situations.

## Conclusion

Human languages provide speakers with alternative means of expressing near meaning-equivalent messages. The competing variants of such alternations differ in the amount and quality of linguistic signal that is provided to listeners. Research on different types of reduction has often proceeded in separate lines of research. This applies in particular to quantitative study of reduction, with most studies focusing on one type of reduction (but see Finegan & Biber, 2001). Here we have focused on one generalization that seems to apply to reduction at various levels of linguistic representation, *probabilistic reduction*. We have summarized evidence that the reduction of form and signal is sensitive to contextual predictability and that this generalization seems to hold across many levels of linguistic representation. Specifically, speakers prefer comparatively reduced realizations of contextually more predictable message components. This relation between the predictability or informativity of message components and the amount of form and signal provided in encoding them has long intrigued language researchers.

The mechanisms that underlie such probabilistic reduction are still under debate. There is, however, converging evidence that both production ease and communicative goals mediate reduction (see also Arnold *et al*., 2012, p. 506), and that reduction can become entrenched over time, leading to reduced canonical forms (see also Ernestus, 2014). A better understanding of how these factors interact will inform research on the architecture of the language production system, the effect of language use and linguistic representations on production, and language change.

## Acknowledgments

The authors wish to thank Jennifer Arnold, Ellen Bard, Alan Bell, Vic Ferreira, Susanne Gahl, Matt Goldrick, Chigusa Kurumada, Janet Pierrehumbert, Uriel Cohen Priva, Andy Wedel, an anonymous reviewer, and the editors of the volume, Eva Fernández and Helen Cairns, for helpful feedback on earlier versions of this article. We also thank Meghan Sumner and members of the Human Language Processing Lab for helpful related discussions. This work was partially supported by NSF CAREER grant IIS-1150028 to TFJ and a graduate fellowship (NIH training grant T32 DC000035) as well as an NRSA pre-doc (F31HD083020) to EB. T. Florian Jaeger may be contacted at fjaeger@ur.rochester.edu.

## NOTES

1 In conversational American English, for example, the complementizer is absent in 82.5% of all complement clauses (Jaeger, 2010, p. 29). Even when the most frequent complement clause embedding verbs are excluded, omission is more frequent (53%) than mention of complementizer *that* (Jaeger, 2010, Table 1).

2 The informativity of, for example, a word can be measured as its Shannon information (also called surprisal). The Shannon information of a word is defined as the logarithm of the inverse of the word's probability in context (Shannon, 1948).

3 For example, the information that hearing a [t] after an [s] adds to the recognition of the word *street*.

4 Investigations of phonetic reduction in conversational speech have reported that about 1–2% or less of the variance in word duration is attributable to contextual predictability (Aylett & Turk, 2004; Bell *et al*., 2009). This would suggest that contextual predictability

has only a small effect on word durations (for example, word frequency is reported to account for about 10% of the variance, Aylett & Turk, 2004). This is, however, misleading since word durations are strongly affected by the phonological canonical form of the word. It is therefore not surprising that factors that vary *between* word forms (such as word frequency) account for more variance than those that vary *within* word forms (such as contextual predictability).

5 This is not to say that the distributions of these forms is unconstrained. For example, both pronouns and lexical noun phrases are subject to constraints of Binding Theory—including, potentially, categorical ones. However, similarly strong constraints apply to optional function word omission or morphological reduction (for examples and references, cf. Jaeger, 2006). Our point here is that there *are* contexts in which speakers can choose between the different forms while maintaining near meaning-equivalence.

6 Some of these effects of syntactic context might be due to predictability effects on prosodic phrasing, with less expected prosodic boundaries being realized with longer phrase final lengthening (Gahl & Garnsey, 2004, 2006; Kurumada, 2011).

7 Or, in the terminology of Dell and Brown (1991), *listener-particular* audience design. We use the term listener-specific to highlight parallels to talker-specific expectations during *comprehension* (e.g., in speech perception, Bradlow & Bent, 2008; lexical processing, Creel, Aslin, & Tanenhaus, 2008; and sentence processing, Kamide, 2012; Kraljic & Samuel, 2007).

8 More recent experiments suggest that it is the number of *any* type of neighbors, rather than minimal pair neighbors, that is correlated with hyper-articulation of the/p/(Fox *et al*., 2015; for discussion, see also Peramunage *et al*., 2011). This might suggest that the hyper-articulated voice onset timing observed in Baese-Berk & Goldrick (2009) was part of more general across-the-board hyper-articulation for words with many neighbors (which has independently been observed, see e.g., Buz & Jaeger, 2015; Gahl, 2015; but see Gahl *et al*., 2012; Munson, 2007). Regardless of the specifics though, hyper-articulation of words with phonological neighbors is compatible with competition accounts.

9 The opposite position—that at least certain aspects of production are too 'automatic' to be directly affected by communicative goals—*has been* proposed (Bard & Aylett, 2005; e.g., Bard *et al*., 2000; Ferreira, 2008).

10 To further complicate the picture, it is possible that what constitutes a good signal can—at least to some extent—change dynamically as a result of accommodation and alignment processes (see, e.g., Pardo, 2017).

## REFERENCES

Allbritton, D. W., McKoon, G., & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 714–735. doi: 10.1037/0278-7393.22.3.714

Anderson, A. H., Badger, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., … Weinert, R. (1991). The HCRC Map Task corpus. *Language and Speech*, 34(4), 351–366. doi: 10.1177/002383099103400404

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences*, 14(03), 471–485. doi: 10.1017/S0140525X00070801

Anibel, B. (2010). *Optional Marking of Coherence Relations: Production of Coherence Markers in Sentence Completion Tasks*. Master's thesis, University of Rochester.

Anttila, A., Adams, M., & Speriosu, M. (2010). The role of prosody in the English dative alternation. *Language and Cognitive Processes*, 25(7-9), 946–981. doi: 10.1080/01690960903525481

Ariel, M. (1999). Cognitive universals and linguistic conventions: The case of resumptive pronouns. *Studies in Language*, 23(2), 217–269.

Arnold, J. E. (1998). *Reference form and discourse patterns* (doctoral dissertation). Stanford University.

Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, 23(4), 495–527. doi: 10.1080/01690960801920099

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Linguistics and Language Compass*, 4, 187–203. doi: 10.1111/j.1749-818X.2010.00193.x

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, 56(4), 521–536. doi: 10.1016/j.jml.2006.09.007

Arnold, J. E., Bennetto, L., & Diehl, J. J. (2009). Reference production in young speakers with and without autism. *Cognition*, 110, 131–146. doi: 10.1016/j.cognition.2008.10.016

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of psycholinguistic research*, 32(1), 25–36. doi: 10.1023/A:1021980931292

Arnold, J. E., Kahn, J. M., & Pancani, G. C. (2012). Audience design affects acoustic reduction via production facilitation. *Psychonomic Bulletin & Review*, 19(3), 505–512. doi: 10.3758/s13423-012-0233-y

Arnold, J. E., Kam, C. L. H., & Tanenhaus, M. K. (2007). If You Say Thee uh You Are Describing Something Hard: The On-Line Attribution of Disfluency During Reference Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(5), 914–930. doi: 10.1037/0278-7393.33.5.914

Arnold, J. E., Losongco, A., Wasow, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76, 28–55. doi: 10.1353/lan.2000.0045

Arnold, J. E., Wasow, T., Asudeh, A., & Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language*, 51(1), 55–70. doi: 10.1016/j.jml.2004.03.006

Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word Frequency and constituency on phonetic duration. *Language and Speech*, 56(3), 349–371. doi: 10.1177/0023830913484891

Arnon, I., & Cohen Priva, U. (2014). Time and again: The changing effect of word and multiword frequency on phonetic duration for highly frequent sequences. *The Mental Lexicon*, 9(3), 377–400. doi: 10.1075/ml.9.3.01arn

Arnon, I., & Snider, N. E. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62, 67–82. doi: 10.1016/j.jml.2009.09.005

Asr, F. T., & Demberg, V. (2012). Implicitness of discourse relations. In M. Kay & C. Boitet (Eds.), *COLING* (pp. 2669–2684). Mumbai: Indian Institute of Technology Bombay.

Asr, F. T., & Demberg, V. (2015). Uniform information density at the level of discourse relations: Negation markers and discourse connective omission.

In *Proceedings of the 11th international conference on computational semantics* (pp. 118–128). London, UK: Association for Computational Linguistics.

Aylett, M. P., & Turk, A. E. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47(1), 31–56. doi: 10.1177/00238309040470010201

Aylett, M. P., & Turk, A. E. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *The Journal of the Acoustical Society of America*, 119(5), 3048–3058. doi: 10.1121/1.2188331

Baayen, R. H., Hendrix, P., & Ramscar, M. (2013). Sidestepping the combinatorial explosion: An explanation of n-gram frequency effects based on naive discriminative learning. *Language and Speech*, 56(3), 329–347. doi: 10.1177/0023830913484896

Baese-Berk, M. M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes*, 24(4), 527–554. doi: 10.1080/01690960802299378

Baese-Berk, M. M., Bradlow, A. R., & Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *The Journal of the Acoustical Society of America*, 133(3), EL174–EL180. doi: 10.1121/1.4789864

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The Effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, 19(3), 241–248. doi: 10.1111/j.1467-9280.2008.02075.x

Bard, E. G., & Anderson, A. H. (1983). The unintelligibility of speech to children. *Journal of Child Language*, 10(02), 265–292. doi: 10.1017/S0305000900007777

Bard, E. G., & Anderson, A. H. (1994). The unintelligibility of speech to children: Effects of referent availability. *Journal of Child Language*, 21(03), 623–648. doi: 10.1017/S030500090000948X

Bard, E. G., & Aylett, M. P. (2005). Referential form, word duration, and modeling the listener in spoken dialogue modeling listeners while speaking. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 173–191). Cambridge: MIT Press.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M. P., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42(1), 1–22. doi: 10.1006/jmla.1999.2667

Beaver, D., Clark, B. Z., Flemming, E. S., Jaeger, T. F., & Wolters, M. K. (2007). When semantics meets phonetics: Acoustical studies of second occurrence focus. *Language*, 83(2), 245–276. doi: 10.1353/lan.2007.0053

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. S. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60(1), 92–111. doi: 10.1016/j.jml.2008.06.003

Bell, A., Jurafsky, D. S., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America*, 113(2), 1001. doi: 10.1121/1.1534836

Bender, E. M. (2000). *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Doctoral dissertation, Stanford University.

Berkovits, R. (1994). Durational Effects in Final Lengthening, Gapping, and Contrastive Stress. *Language and Speech*, 37(3), 237–250. doi: 10.1177/002383099403700302

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89(1), 1–47. doi: 10.1037/0033-295X.89.1.1

Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, 26, 119–137. doi: 10.1016/0749-596X(87)90120-3

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106(2), 707–729. doi: 10.1016/j.cognition.2007.04.005

Branigan, H. P., Pickering, M. J., & Tanaka, M. N. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118(2), 172–189. doi: 10.1016/j.lingua.2007.02.003

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482–1493. doi: 10.1037/0278-7393.22.6.1482

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science*, 1(2), 274–291. doi: 10.1111/j.1756-8765.2009.01019.x

Bresnan, J., & Spencer, J. (2016). Frequency effects in spoken syntax: "have" and "be" contraction. In S. Grondelaers & R. van Hout (Eds.), *In new ways of analyzing syntactic variation*. Berlin: De Gruyter Mouton.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Krämer, & J. Zwarts (Eds.), *Cognitive foundations of interpretation* (pp. 69–94). Amsterdam: Edita-The Publishing House of the Royal Netherlands Academy.

Buz, E. (2016). *Speaking with a (slightly) new voice: Investigating speech changes, communicative goals, and learning from past communicative success*. Doctoral dissertation, University of Rochester.

Buz, E., & Jaeger, T. F. (2016). The (in)dependence of articulation and lexical planning during isolated word production. *Language, Cognition and Neuroscience*, 31(3), 404–424. doi: 10.1080/23273798.2015.1105984

Buz, E., Jaeger, T. F., & Tanenhaus, M. K. (2014). Contextual confusability leads to targeted hyperarticulation. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*. Quebec City, Canada: Cognitive Science Society.

Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (submitted). Do speakers make inferences about the causes of miscommunication? Manuscript, University of Rochester.

Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language*, 89, 68–86. doi: 10.1016/j.jml.2015.12.009

Bürki, A., Ernestus, M., & Frauenfelder, U. H. (2010). Is there only one "fenêtre" in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language*, 62(4), 421–437. doi: 10.1016/j.jml.2010.01.002

Bürki, A., Ernestus, M., Gendrot, C., Fougeron, C., & Frauenfelder, U. H. (2011). What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *The Journal of the Acoustical Society of America*, 130(6), 3980. doi: 10.1121/1.3658386

Bybee, J., & Hopper, P. (2001). Introduction to frequency and the emergence of linguistic structure. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 9–32). John Benjamins Publishing.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of don't in English. *Linguistics*, 37(4), 575–596. doi: 10.1515/ling.37.4.575

Caballero, G., & Kapatsinski, V. (2014). Perceptual functionality of morphological redundancy in Choguita Rarámuri (Tarahumara). *Language, Cognition and Neuroscience*, (July 2015), 1–10. doi: 10.1080/23273798.2014.940983

Chen, Q., & Mirman, D. (2015). Interaction between phonological and semantic representations: Time matters. *Cognitive science*, 39(3), 538–558. doi: 10.1111/cogs.12156

Chomsky, N. (2000). *New horizons in the study of language and the mind* (p. 230). Cambridge: Cambridge University Press.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73–111. doi: 10.1016/S0010-0277(02)00017-3

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In J.-F. L. Ny & W. Kintsch (Eds.), *Language and comprehension* (1st ed., pp. 287–300). Amsterdam: North-Holland Publishing Company.

Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology*, 37(3), 201–242. doi: 10.1006/cogp.1998.0693

Clopper, C. G., & Pierrehumbert, J. B. (2008). Effects of semantic predictability and regional dialect on vowel space reduction. *The Journal of the Acoustical Society of America*, 124(3), 1682–1688. doi: 10.1121/1.2953322

Cohen Priva, U. (2008). Using information content to predict phone deletion. In N. Abner & J. Bishop (Eds.), *Proceedings of the 27th West Coast Conference on Formal Linguistics* (pp. 90–98). Somerville, MA: Cascadilla.

Cohen Priva, U. (2012). *Sign and signal: deriving linguistic generalizations from information utility*. Doctoral dissertation, Stanford University.

Cohen Priva, U. (2015). Informativity affects consonant duration and deletion rates. *Laboratory Phonology*, 6(2), 243–278. doi: 10.1515/lp-2015-0008

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: the role of talker variation in lexical access. *Cognition*, 106(2), 633–664. doi: 10.1016/j.cognition.2007.03.013

De Jong, K. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, 32(4), 493–516. doi: 10.1016/j.wocn.2004.05.002

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283–321. doi: 10.1037/0033-295X.93.3.283

Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of the “model of the listener”. In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A swarthmore festschrift for lila gleitman* (pp. 105–130). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, 23(4), 517–542. doi: 10.1016/S0364-0213(99)00014-2

Demberg, V., Sayeed, A. B., Gorinski, P. J., & Engonopoulos, N. (2012). Syntactic surprisal affects spoken word duration in conversational contexts. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning* (pp. 356–367). Association for Computational Linguistics.

Drager, K. (2011). Speaker age and vowel perception. *Language and Speech*, 54(Pt 1), 99–121. doi: 10.1177/0023830910388017

Dryer, M. S. (2013). Expression of pronominal subjects. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online* (pp. 410–413). Leipzig: Max Planck Institute for Evolutionary Anthropology.

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *The Journal of the Acoustical Society of America*, 119(4), 1950–1953. doi: 10.1121/1.2178721

Elsness, J. (1984). That or Zero? A look at the choice of object clause connective in a corpus of American English. *English Studies*, 65(6), 519–533.

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*, 142, 27–41. doi: 10.1016/j.lingua.2012.12.006

Ernestus, M., & Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, 39(3), 253–260. doi: 10.1016/S0095-4470(11)00055-6

Ernestus, M., Baayen, R. H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and language*, 81(1-3), 162–173. doi: 10.1006/brln.2001.2514

Fedzechkina, M., Newport, E., and Jaeger, T. F. 2016. Balancing effort and information transmission during language acquisition: *Evidence from word order and case-marking. Cognitive Science*, 41(2), 416–446. doi: 10.1111/cogs.12346

Fedzechkina, M., Jaeger, T. F., and Newport, E. 2012. Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences* 109(44), 17897–17902. doi: 10.1073/pnas.1215776109

Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language*, 35(5), 724–755. doi: 10.1006/jmla.1996.0038

Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation*, 49, 209–246. doi: 10.1016/S0079-7421(08)00006-6

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40(4), 296–340. doi: 10.1006/cogp.1999.0730

Ferreira, V. S., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, 32(6), 669–692. doi: 10.1023/A:1026146332132

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition*, 96(3), 263–284. doi: 10.1016/j.cognition.2004.09.002

Finegan, E., & Biber, D. (2001). Register variation and social dialect variation: The register axiom. In P. Eckert & J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 235–267). Cambridge: Cambridge University Press.

Flickinger, D., & Wasow, T. (2013). A corpus-driven analysis of the Do-Be construction. In P. Hofmeister & E. J. Norcliffe (Eds.), *The core and the periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag* (pp. 35–63). CSLI Publications.

Foulkes, P., & Hay, J. B. (2015). The emergence of sociophonetic structure. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 292–313). Oxford: Blackwell.

Fox, N. P., Reilly, M., & Blumstein, S. E. (2015). Phonological neighborhood competition affects spoken word production irrespective of sentential context. *Journal of Memory and Language*, 83, 97–117. doi: 10.1016/j.jml.2015.04.002

Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing "the" as "thee" to signal problems in speaking. *Cognition*, 62(2), 151–67. doi: 10.1016/S0010-0277(96)00781-0

Frank, A. F. (2011). *Integrating linguistic, motor, and perceptual information in language production*. Doctoral dissertation, University of Rochester.

Frank, A. F., & Jaeger, T. F. (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Meeting of the Cognitive*

*Science Society* (pp. 939–944). Cognitive Science Society.

Friedman, L. F. (2013). Chart: The Art of "'Sup". Retrieved from https://www.psychologytoday.com/articles/201304/chart-the-art-sup

Fry, J. S. (2001). *Ellipsis and wa-marking in Japanese conversation*. Doctoral dissertation, Stanford University.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84(3), 474–496. doi: 10.1353/lan.0.0035

Gahl, S. (2015). Lexical competition in vowel articulation revisited: Vowel dispersion in the easy/Hard database. *Journal of Phonetics*, 49, 96–116. doi: 10.1016/j.wocn.2014.12.002

Gahl, S., & Garnsey, S. M. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, 80(4), 748–775. doi: 10.1353/lan.2004.0185

Gahl, S., & Garnsey, S. M. (2006). Knowledge of Grammar Includes Knowledge of Syntactic Probabilities. *Language*, 82(2), 405–410. doi: 10.1353/lan.2006.0089

Gahl, S., Jurafsky, D. S., & Roland, D. (2004). Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, & Computers*, 36(3), 432–443. doi: 10.3758/BF03195591

Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, 66(4), 789–806. doi: 10.1016/j.jml.2011.11.006

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, 62(1), 35–51. doi: 10.1016/j.jml.2009.09.002

Gallo, C. G. (2011). *Incremental interpretation and production across clauses*. Doctoral dissertation, University of Rochester.

Gallo, C. G., Jaeger, T. F., & Furth, K. (2010). A Database for the Exploration of Spanish Planning. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, … D. Tapias (Eds.), *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)* (pp. 1–4).

Gallo, C. G., Jaeger, T. F., & Smyth, R. (2008). Incremental Syntactic Planning across Clauses. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 1294–1299).

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37(1), 58–93. doi: 10.1006/jmla.1997.2512

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(20), 8051–8016. doi: 10.1073/pnas.1216438110

Givón, T. (1983). *Topic Continuity in Discourse: A quantitative cross-language study* (p. 492). John Benjamins Publishing.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1166–1183. doi: 10.1037/0278-7393.22.5.1166

Goldrick, M., Vaughn, C., & Murphy, A. (2013). The effects of lexical neighbors on stop consonant articulation. *The Journal of the Acoustical Society of America*, 134(2), EL172–EL177. doi: 10.1121/1.4812821

Guy, G. R. (1996). Form and function in linguistic variation. In G. R. Guy, C. Feagin, D. Schiffrin, & J. Baugh (Eds.), *Towards a social science of language: Papers in honor of william labov. volume 1: Variation and change in language and society* (pp. 221–252). Amsterdam: Benjamins Publishing Company.

Hanique, I., Ernestus, M., & Schuppler, B. (2013). Informal speech processes can be categorical in nature, even if they affect many different words. *The Journal of the Acoustical Society of America*, 133(3), 1644–1655. doi: 10.1121/1.4790352

Hare, M., McRae, K., & Elman, J. L. (2004). Admitting that admitting verb sense into corpus analyses makes sense. *Language and Cognitive Processes*, 19(2), 181–224. doi: 10.1080/01690960344000152

Hay, J. B., Pierrehumbert, J. B., Walker, A. J., & LaShell, P. (2015). Tracking word frequency effects through 130 years of sound change. *Cognition*, 139, 83–91. doi: 10.1016/j.cognition.2015.02.012

Hay, J. B., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4), 458–484. doi: 10.1016/j.wocn.2005.10.001

Haywood, S. L., Pickering, M. J., & Branigan, H. P. (2005). Do speakers avoid ambiguities during dialogue? *Psychological Science*, 16(5), 362–366. doi: 10.1111/j.0956-7976.2005.01541.x

Heller, D., Gorman, K. S., & Tanenhaus, M. K. (2012). To name or to describe: shared knowledge affects referential form. *Topics in Cognitive Science*, 4(2), 290–305. doi: 10.1111/j.1756-8765.2012.01182.x

Heller, J. R., & Goldrick, M. (2014). Grammatical constraints on phonological encoding in speech production. *Psychonomic Bulletin & Review*, 21(6), 1576–1582. doi: 10.3758/s13423-014-0616-3

Heller, J. R., & Pierrehumbert, J. B. (2011). Word burstiness improves models of word reduction in spontaneous speech. In *Architectures and Mechanisms for Language Processing (AMLaP) Conference*. Paris, France.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59, 91–117. doi: 10.1016/0010-0277(96)81418-1

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, 279(5354), 1213–1216. doi: 10.1126/science.279.5354.1213

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of English*. Cambridge: Cambridge University Press.

Hall, K., Hume, E., Jaeger, T. F., & Wedel, A. (submitted). *The message shapes phonology*.

Jaeger, T. F. (2006). *Redundancy and Syntactic Reduction in Spontaneous Speech*. Doctoral dissertation, Stanford University.

Jaeger, T. F. (2007). Rational speakers: Speakers help processing when it is most necessary. In *Architectures and Mechanisms for Language Processing (AMLaP) Conference*, Ghent, Belgium.

Jaeger, T. F. (2010). Redundancy and reduction: speakers manage syntactic information density. *Cognitive Psychology*, 61(1), 23–62. doi: 10.1016/j.cogpsych.2010.02.002

Jaeger, T. F. (2011). Corpus-based research on language production: Information density and reducible subject relatives. In E. M. Bender & J. E. Arnold (Eds.), *Language from a cognitive perspective: Grammar, usage, and processing. studies in honor of Tom Wasow* (pp. 161–197). Stanford, CA: CSLI Publications.

Jaeger, T. F. (2013). Production preferences cannot be understood without reference to communication. *Frontiers in Psychology*, 4, 230. doi: 10.3389/fpsyg.2013.00230

Jaeger, T. F., & Ferreira, V. S. (2013). Seeking predictions from a predictive framework. *Behavioral and Brain Sciences*, 36(4), 359–360. doi: 10.1017/S0140525X12002762

Jaeger, T. F., & Grimshaw, J. (2013). Information density affects both production and grammatical constraints. In *Architectures and Mechanisms for Language Processing (AMLaP) Conference*. Marseille, France.

Jaeger, T. F., & Norcliffe, E. J. (2009). The Cross-linguistic Study of Sentence Production. *Language and Linguistics Compass*, 3(4), 866–887. doi: 10.1111/j.1749-818X.2009.00147.x

Jaeger, T. F., & Wasow, T. (2006). Processing as a Source of Accessibility Effects on

Variation. In R. T. Cover & Y. Kim (Eds.), *Proceedings of the 31st annual meeting of the Berkeley Linguistic Society* (pp. 169–180). Ann Arbor, MI: Sheridan Books.

Jaeger, T. F., Furth, K., & Hilliard, C. (2012a). Incremental phonological encoding during unscripted sentence production. *Frontiers in Psychology*, 3(November), 481. doi: 10.3389/fpsyg.2012.00481

Jaeger, T. F., Furth, K., & Hilliard, C. (2012b). Phonological overlap affects lexical selection during sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(5), 1439–1449. doi: 10.1037/a0027862

Jaeger, T. F., Snider, N. E., Staum, L., & Jurafsky, D. S. (2006). (In)dependence of lexical and syntactic production: that-reduction and omission in spontaneous speech. In *Proceedings of the 19th Annual CUNY Conference on Human Sentence Processing* (p. 42). New York, NY.

Janssen, N., & Caramazza, A. (2009). Grammatical and phonological influences on word order. *Psychological Science*, 20(10), 1262–1268. doi: 10.1111/j.1467-9280.2009.02429.x

Jenkins, L. (2000). *Biolinguistics: Exploring the biology of language*. Cambridge: Cambridge University Press.

Johnson, K. (1997). The auditory/perceptual basis for speech segmentation. *OSU Working Papers in Linguistics*, 50, 101–113.

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous Speech: Data and Analysis. Proceedings of the 1st Session of the 10th International Symposium*. Tokyo, Japan: The National International Institute for Japanese Language.

Jurafsky, D. S., Bell, A., Gregory, M., & Raymond, W. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: John Benjamins Publishing Company.

Kahn, J. M., & Arnold, J. E. (2012). A processing-centered look at the contribution of givenness to durational reduction. *Journal of Memory and Language*, 67(3), 311–325. doi: 10.1016/j.jml.2012.07.002

Kahn, J. M., & Arnold, J. E. (2015). Articulatory and lexical repetition effects on durational reduction: Speaker experience vs. common ground. *Language, Cognition and Neuroscience*, 30(1–2), 103–119. doi: 10.1080/01690965.2013.848989

Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. (2017). Zipf's Law of Abbreviation and the Principle of Least Effort: Language users optimise a miniature lexicon for efficient communication. *Cognition*, 165, 45–52. http://doi.org/10.1016/j.cognition.2017.05.001.

Kamide, Y. (2012). Learning individual talkers' structural preferences. *Cognition*, 124(1), 66–71. doi: 10.1016/j.cognition.2012.03.001

Kempen, G., & Harbusch, K. (2004). A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In T. Pechmann & C. Habel (Eds.), *Multidisciplinary approaches to language production* (pp. 173–182). Berlin: Walter de Gruyter.

Kidd, C., & Jaeger, T. F. (2008). Prosodic phrasing and function word pronunciation. In *Experimental and Theoretical Advances in Prosody*. Cornell University, Ithaca, NY.

Kirov, C., & Wilson, C. (2012). The specificity of online variation in speech production. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 587–592). Austin, TX: Cognitive Science Society.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, 122(2), 148–203. doi: 10.1037/a0038695

Kohler, K. J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 69–92). Boston/London: Kluwer.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, 51(2), 141–178. doi: 10.1016/j.cogpsych.2005.05.001

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, 13(2), 262–268. doi: 10.3758/BF03193841

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, 56(1), 1–15. doi: 10.1016/j.jml.2006.07.010

Kravtchenko, E. (2014). Predictability and syntactic production: Evidence from subject omission in Russian. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 785–790). Austin, TX: Cognitive Science Society.

Kravtchenko, E., & Demberg, V. (2015). Semantically underinformative utterances trigger pragmatic inferences. In R. Dale, C. Jennings, P. Maglio, T. Matlock, D. Noelle, A. Warlaumont, & J. Yoshimi (Eds.), *Proceedings of the 37th annual meeting of the cognitive science society* (pp. 1207–1212). Austin, TX: Cognitive Science Society.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., … Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science (New York, N.Y.)*, 277(5326), 684–686. doi: 10.1126/science.277.5326.684

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, 20(1), 54–72. doi: 10.3758/s13423-012-0341-8

Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2007). Morphological predictability and acoustic duration of interfixes in Dutch compounds. *The Journal of the Acoustical Society of America*, 121(4), 2261. doi: 10.1121/1.2537393

Kurumada, C. (2011). *Syntactic context affects probability estimation: Evidence from Japanese relative clauses*. Manuscript. Stanford University.

Kurumada, C., & Jaeger, T. F. (2015). Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language*, 83, 152–178. doi: 10.1016/j.jml.2015.03.003

Labov, W. (1994). *Principles of linguistic change, volume 1: Internal factors* (p. 664). Malden, MA: Wiley-Blackwell.

Lee, H. (2006). Parallel optimization in case systems: Evidence from case ellipsis in Korean. *Journal of East Asian Linguistics*, 15(1), 69–96. doi: 10.1007/s10831-005-3004-1

Lee, M.-W., & Gibbons, J. (2007). Rhythmic alternation and the optional complementiser in English: New evidence of phonological influence on grammatical encoding. *Cognition*, 105(2), 446–456.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. doi: 10.1016/j.cognition.2007.05.006

Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In B. Schlökopf, J. Platt, & T. Hoffman (Eds.), *Advances in*

*Neural Information Processing Systems* (Vol. 19, pp. 849–856). Cambridge, MA: MIT Press.

Lindblom, B. (1990a). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403–439). Kluwer Academic Publishers.

Lindblom, B. (1990b). On the communication process: Speaker-listener interaction and the development of speech. *Augmentative and Alternative Communication*, 6(4), 220–230. doi: 10.1080/07434619012331275504

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, 9(3), 550–557. doi: 10.3758/BF03196312

Lombard, E. (1911). Le signe de l'elevation de la voix. *Ann Maladies Oreille, Larynx, Nez, Pharynx*, 37(25), 101–119.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4, 226. doi: 10.3389/fpsyg.2013.00226

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703. doi: 10.1037/0033-295X.101.4.676

Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, 126(2), 313–318. doi: 10.1016/j.cognition.2012.09.010

Manin, D. Y. (2006). Experiments on predictability of word in context and information rate in natural language. *Journal of Information Processes*, 229–236.

McDonald, J. L., Bock, J. K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, 25(2), 188–230. doi: 10.1006/cogp.1993.1005

Melnick, R. (2011). *Judgment studies of information-driven syntactic reduction.* Master's thesis, Stanford University.

Munson, B. (2007). Lexical Access, Lexical Representation, and Vowel Production. In J. Cole & J. I. Hualde (Eds.), *Laboratory phonology 9* (pp. 201–228). Berlin: Mouton de Gruyter.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18(1), 62–85. doi: 10.1177/0261927X99018001005

Norcliffe, E. J. (2009). *Head-marking in usage and grammar: A study of variation and change in Yucatec Maya*. Doctoral dissertation, Stanford University.

Norcliffe, E. J., & Jaeger, T. F. (2014). Predicting head-marking variability in Yucatec Maya relative clause production. *Language and Cognition*, 1–39. doi: 10.1017/langcog.2014.39

Norcliffe, E. J., Harris, A. C., & Jaeger, T. F. (2015). Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition and Neuroscience*, 30(9), 1009–1032. doi: 10.1080/23273798.2015.1080373

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, 115(2), 357–395. doi: 10.1037/0033-295X.115.2.357

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, 47(2), 204–238. doi: 10.1016/S0010-0285(03)00006-9

Ohala, J. J. (1988). Discussion of Bjoern Lindblom's "Phonetic Invariance and the adaptive nature of speech". In *Working models of human perception*. London, UK: Academic Press.

O'Seaghdha, P. G., & Marin, J. W. (2000). Phonological competition and cooperation in form-related priming: sequential and nonsequential processes in word production. *Journal of*

*Experimental Psychology: Human Perception and Performance*, 26(1), 57–73. doi: 10.1037/0096-1523.26.1.57

Pardo, J. S. (2017). Parity and disparity in conversational interaction. In E. M. Fernández & H. S. Cairns (Eds.), *Handbook of psycholinguistics*. Wiley-Blackwell.

Pardo, J. S., & Remez, R. E. (2006). The Perception of Speech. In *The handbook of psycholinguistics, 2nd edition* (pp. 201–248).

Pate, J. K., & Goldwater, S. (2015). Talkers account for listener and channel characteristics to communicate efficiently. *Journal of Memory and Language*, 78, 1–17. doi: 10.1016/j.jml.2014.10.003

Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language*, 87(3), 539–558. doi: 10.1353/lan.2011.0057

Peramunage, D., Blumstein, S. E., Myers, E. B., Goldrick, M., & Baese-Berk, M. M. (2011). Phonological neighborhood effects in spoken word production: An fMRI study. *Journal of Cognitive Neuroscience*, 23(3), 593–603. doi: 10.1162/jocn.2010.21489

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences of the United States of America*, 108(9), 3526–3529. doi: 10.1073/pnas.1012551108

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 29, 434–446. doi: 10.1044/jshr.2904.434

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–157). Amsterdam: John Benjamins Publishing Company.

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 101–139). Berlin: Mouton de Gruyter.

Pierrehumbert, J. B. (2003). Probabilistic Phonology : Discrimination and Robustness. In R. Bod, J. B. Hay, & S. Jannedy (Eds.), *Probability theory in linguistics*. Cambridge, MA: MIT Press.

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005a). Lexical frequency and acoustic reduction in spoken Dutch. *The Journal of the Acoustical Society of America*, 118(4), 2561. doi: 10.1121/1.2011150

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005b). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, 62(2-4), 146–159. doi: 10.1159/000090095

Post, M., & Jaeger, T. F. (2010). Word Production in Spontaneous Speech: Availability and Communicative Efficiency. In *The 23rd CUNY Sentence Processing Conference*. NYC, NY.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, 90(6), 2956. doi: 10.1121/1.401770

Race, D. S., & MacDonald, M. C. (2003). The use of "that" in the production and comprehension of object relative clauses. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 946–951). Mahwah, NJ: Cognitive Science Society.

Rapp, D. N., & Samuel, A. G. (2002). A reason to rhyme: Phonological and semantic influences on lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 564–571. doi: 10.1037/0278-7393.28.3.564

Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realization. *Cognition*, 61(1-2), 127–159. doi: 10.1016/S0010-0277(96)00722-6

Rischel, J. (1992). Formal linguistics and real speech. *Speech Communication*, 11,

379–392. doi: 10.1016/0167-6393(92) 90043-7
Roche, J. M., Dale, R., & Kreuz, R. J. (2010). The resolution of ambiguity during conversation: More than mere mimicry? In R. Camtrabone & S. Ohlsson (Eds.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 206–211). Austin, TX: Cognitive Science Society.
Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, 29(8), 912–927. doi: 10.1080/01690965.2013.854918
Rohdenburg, G. (2004). Grammatische Parallelen zwischen niederdeutschen Mundarten und Nichtstandardvarietäten im Englischen aus typologischer Sicht. *Niederdeutsches Jahrbuch*, 127, 85–122.
Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, 57(3), 348–379. doi: 10.1016/j.jml.2007.03.002
Roland, D., Elman, J. L., & Ferreira, V. S. (2006). Why is that? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition*, 98(3), 245–272. doi: 10.1016/j.cognition.2004.11.008
Rose, D. E. (2017). Predicting plurality: An examination of the effects of morphological predictability on the learning and realization of bound morphemes (Unpublished PhD dissertation). University of Canterbury, Christchurch, New Zealand. Chapter 4.
Rosenbach, A. (2008). Animacy and grammatical variation: Findings from English genitive variation. *Lingua*, 118(2), 151–171. doi: 10.1016/j.lingua.2007.02.002
Rosenfeld, R. (1996). A maximum entropy approach to adaptive statistical language modeling. *Computer, Speech and Language*, 10, 187–228. doi: 10.1006/csla.1996.0011
Sayeed, A. B., Fischer, S., & Demberg, V. (2015). Vector-space calculation of semantic surprisal for predicting word pronunciation duration. In *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers)* (pp. 763–773). Beijing, China: Association for Computational Linguistics. Retrieved from http://www.aclweb.org/anthology/P15-1074
Scarborough, R. (2010). Lexical and contextual predictability: Confluent effects on the production of vowels. In C. Fougeron, B. Kuehnert, M. Imperio, & N. Vallee (Eds.), *Laboratory phonology 10* (pp. 557–586). Berlin: Mouton de Gruyter.
Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of psycholinguistic research*, 29(2), 169–182. doi: 10.1023/A:1005192911512
Schertz, J. (2013). Exaggeration of featural contrasts in clarifications of misheard speech in English. *Journal of Phonetics*, 41(3-4), 249–263. doi: 10.1016/j.wocn.2013.03.007
Sevald, C. A., & Dell, G. S. (1994). The sequential cuing effect in speech production. *Cognition*, 53, 91–127. doi: 10.1016/0749-596X(90)90011-N
Seyfarth, S. (2014). Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition*, 133(1), 140–155. doi: 10.1016/j.cognition.2014.06.013
Seyfarth, S., Buz, E., & Jaeger, T. F. (2016). Dynamic hyperarticulation of coda voicing contrasts. *Journal of the Acoustic Society of America* 139(2), EL31–37. doi: 10.1121/1.4942544.
Shannon, C. E. (1948). A Mathematical Theory of Communication. *Mobile Computating and Communications Review*, 5(1), 3–55.

Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell System Technical Journal*, 30(1), 50–64. doi: 10.1002/j.1538-7305.1951.tb01366.x

Shintel, H., & Keysar, B. (2009). Less Is More: A Minimalist Account of Joint Action in Communication. *Topics in Cognitive Science*, 1(2), 260–273. doi: 10.1111/j.1756-8765.2009.01018.x

Shriberg, E. (1996). Disfluencies in switchboard. In H. T. Bunnell & W. Idsardi (Eds.), *Proceedings of International Conference on Spoken Language Processing* (pp. 11–14). Philadelphia, PA: IEEE.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41(1), 1–20. doi: 10.1146/annurev.ps.41.020190.000245

Smith, M., & Wheeldon, L. R. (2004). Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(3), 675–686. doi: 10.1037/0278-7393.30.3.675

Snedeker, J., & Trueswell, J. C. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and language*, 48, 103–130. doi: 10.1016/S0749-596X(02)00519-3

Sóskuthy, M., & Hay, J. (2017). Changing word usage predicts changing word durations in New Zealand English. *Cognition*, 166, 298–313. http://doi.org/10.1016/j.cognition.2017.05.032

Stent, A. J., Huffman, M. K., & Brennan, S. E. (2008). Adapting speaking after evidence of misrecognition: Local and global hyperarticulation. *Speech Communication*, 50(3), 163–178. doi: 10.1016/j.specom.2007.07.005

Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, 10(01), 1–15. doi: 10.1017/S0305000900005092

Strand, E. A. (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology*, 18(1), 86–100. doi: 10.1177/0261927X99018001006

Tagliamonte, S., & Smith, J. (2005). No momentary fancy! The zero in English dialects. *English Language and Linguistics*, 9(02), 289. doi: 10.1017/S1360674305001644

Tanaka, M. N., Branigan, H. P., McLean, J. F., & Pickering, M. J. (2011). Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language*, 65(3), 318–330. doi: 10.1016/j.jml.2011.04.009

Tanenhaus, M. K. (2013). All P's or mixed vegetables? *Frontiers in Psychology*, 4(May), 234. doi: 10.3389/fpsyg.2013.00234

Tanenhaus, M. K., & Brown-Schmidt, S. (2008). Language processing in the natural world. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 363(1493), 1105–1122. doi: 10.1098/rstb.2007.2162

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 217–262). New York: Academic Press.

Temperley, D. (2003). Ambiguity avoidance in English relative clauses. *Language*, 79(3), 464–484. doi: 10.1353/lan.2003.0189

Tily, H., & Kuperman, V. (2012). Rational phonological lengthening in spoken Dutch. *The Journal of the Acoustical Society of America*, 132(6), 3935–40. doi: 10.1121/1.4765071

Tily, H., & Piantadosi, S. T. (2009). Refer efficiently: Use less informative expressions for more predictable meanings. In K. V. Deemter, A. Gatt, M. V. Gompel, & E. J. Krahmer (Eds.), *Proceedings of the Workshop on the Production of Referring Expressions: Bridging the Gap Between Computational and Empirical Approaches to Reference* (pp. 1–8).

Tily, H., Gahl, S., Arnon, I., Snider, N. E., Kothari, A., & Bresnan, J. (2009). Syntactic probabilities affect pronunciation variation in spontaneous speech. *Language and Cognition*, 1(02), 147–165. doi: 10.1515/LANGCOG.2009.008

Torreira, F., & Ernestus, M. (2009). Probabilistic effects on French [t] duration. *10th Annual Conference of the International Speech Communication Association 2009 (INTERSPEECH 2009)*, 448–451.

Torreira, F., & Ernestus, M. (2011). Vowel elision in casual French: The case of vowel/e/in the word c'etait. *Journal of Phonetics*, 39(1), 50–58. doi: 10.1016/j.wocn.2010.11.003

Torreira, F., & Ernestus, M. (2012). Weakening of intervocalic/s/in the Nijmegen Corpus of Casual Spanish. *Phonetica*, 69(3), 124–148. doi: 10.1159/000343635

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, 39(3), 1429–43. doi: 10.1016/j.neuroimage.2007.09.054

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, 49(1), 2–7. doi: 10.1016/j.specom.2006.10.003

van de Ven, M., Tucker, B. V., & Ernestus, M. (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Memory & Cognition*, 39(7), 1301–1316. doi: 10.3758/s13421-011-0103-2

van Son, R. J. J. H., & Pols, L. C. W. (2002). Evidence for efficiency in vowel production. In *7th International Conference on Spoken Language Processing*. Institute of Phonetic Sciences, University of Amsterdam.

van Son, R. J. J. H., & Pols, L. C. W. (2003). How efficient is speech? *Proceedings Institute of Phonetic Sciences, University of Amsterdam*, 25, 171–184.

van Son, R. J. J. H., & van Santen, J. P. H. (2005). Duration and spectral balance of intervocalic consonants: A case for efficient communication. *Speech Communication*, 47(1-2), 100–123. doi: 10.1016/j.specom.2005.06.005

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, 122(4), 2306–2319. doi: 10.1121/1.2773966

Walker, A., & Hay, J. B. (2011). Congruence between "word age" and "voice age" facilitates lexical access. *Laboratory Phonology*, 2(1), 219–237. doi: 10.1515/labphon.2011.007

Wardlow Lane, L., & Ferreira, V. S. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1466–1481. doi: 10.1037/a0013353

Wasow, T. (1997). End-weight from the speaker's perspective. *Journal of Psycholinguistic Research*, 26(3). doi: 10.1023/A:1025080709112

Wasow, T., & Arnold, J. E. (2003). Post-verbal constituent ordering in English. In G. Rohdenburg & B. Mondorf (Eds.), *Determinants of grammatical variation in english* (pp. 119–154). Berlin: Mouton de Gruyter.

Wasow, T., Jaeger, T. F., & Orr, D. M. (2011). Lexical Variation in Relativizer Frequency. In H. Simon & H. Wiese (Eds.), *Expecting the unexpected: Exceptions in grammar* (pp. 175–195). Berlin: De Gruyter.

Wasow, T., Levy, R., Melnick, R., Juzek, T., & Zhu, H. (2015). Processing, prosody and optional *to*. In L. Frazier & E. Gibson (Eds.), *Explicit and implicit prosody in sentence processing. studies in honor of Janet Dean Fodor*. (pp. 133–158). Switzerland: Springer International Publishing. doi: 10.1007/978-3-319-12961-7_8

Watson, D. G., Arnold, J. E., & Tanenhaus, M. K. (2010). Corrigendum to Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 114(3), 462–463. doi: 10.1016/j.cognition.2010.01.007

Watson, D. G., Buxó-Lugo, A., & Simmons, D. C. (2015). The effect of phonological encoding on word duration: Selection takes time. In E. Gibson & L. Frazier (Eds.), *Explicit and implicit prosody in sentence processing* (pp. 85–98). Switzerland: Springer International Publishing. doi: 10.1007/978-3-319-12961-7_5

Weatherholtz, K. (2015). *Perceptual learning of systemic cross-category vowel variation*. Doctoral dissertation, The Ohio State University.

Weatherholtz, K., & Jaeger, T. F. (2016). Speech perception and generalization across speakers and accents. In *Oxford Research Encyclopedia of Linguistics, Dec. 2016*. doi: 10.1093/acrefore/9780199384655.013.95.

Wedel, A. (2006). Exemplar models, evolution and language change. *The Linguistic Review*, 23, 247–274. doi: 10.1515/TLR.2006.010

Wedel, A., & Sharp, R. (2014). Phonetic cues distinguishing lexical competitors are hyperarticulated in natural speech. In *Linguistics Society of America (LSA) Annual Meeting* (Vol. 1). Minneapolis, MN.

Wedel, A., Jackson, S., & Kaplan, A. (2013). Functional Load and the Lexicon: Evidence that Syntactic Category and Frequency Relationships in Minimal Lemma Pairs Predict the Loss of Phoneme contrasts in Language Change. *Language and Speech*, 56(3), 395–417. doi: 10.1177/0023830913489096

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: a corpus study. *Cognition*, 128(2), 179–186. doi: 10.1016/j.cognition.2013.03.002

Wheeldon, L. R. (2003). Inhibitory form priming of spoken word production. *Language and Cognitive Processes*, 18(1), 81–109. doi: 10.1080/01690960143000470

Wheeldon, L. R., & Lahiri, A. (1997). Prosodic Units in Speech Production. *Journal of Memory and Language*, 37(3), 356–381. doi: 10.1006/jmla.1997.2517

Wiechmann, D. (2015). *Understanding relative clauses: A usage-based view on the processing of complex constructions* (Vol. 268, p. 270). Berlin: Walter de Gruyter.

Zhao, Y., & Jurafsky, D. S. (2009). The effect of lexical frequency and Lombard reflex on tone hyperarticulation. *Journal of Phonetics*, 37(2), 231–247. doi: 10.1016/j.wocn.2009.03.002

Zipf, G. K. (1929). Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology*, 40, 1–95.

Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort* (p. 573). Oxford, UK: Addison-Wesley Press.

# 4 Production in Bilingual and Multilingual Speakers

DANIELA PAOLIERI[1], LUIS MORALES[2], AND TERESA BAJO[1]

[1] University of Granada
[2] Universidad Loyola Andalucía

## Introduction

Bilingualism refers to the coexistence of more than one language system within an individual (Hakuta, 2009). The ability to communicate in more than one language covers a large spectrum of proficiencies, from having a native-like dominance of different languages to the ability to communicate at almost any proficiency level (Cenoz, Hufeisen, & Jessner, 2003). A balanced level of proficiency in several languages is rather infrequent if we take into account the different dimensions of communicative competence, including linguistic, pragmatic, sociolinguistic, discourse, and strategic competence (Celce-Murcia, Dörnyei, & Thurrell, 1995). Grosjean's (1982) definition of bilingualism invokes, rather than proficiency, the notion of regular use of two (or more) languages (or dialects) in their everyday lives.

Researchers have estimated that at least half of the world's population is multilingual to some degree. To illustrate, the European Commission published a report in 2006 that asked Europeans about their native language and their knowledge of other languages. This report showed that approximately 56% of the inhabitants of 25 European countries were able to speak a second language fluently enough to have a conversation in it, even if they did not speak the two languages on a daily basis. Furthermore, two-thirds of the world's children grow up in a plurilingual environment (Crystal, 2003). Although counting the number of users of a single or multiple language is very difficult, this percentage gives us an idea of how extensive multilingualism can be. The question of how different languages interact at the cognitive and behavioral levels has been of long-standing interest to psycholinguists as well as to neurologists, clinicians, and educators (Hakuta, 2009).

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

Moreover, in our view, a deep knowledge of the cognitive and brain mechanisms involved in language processing can only be achieved if they are also explored from the perspective of bilingual speakers. Additionally, the way in which the mind of a bilingual copes with different languages may shed some light on the processes that otherwise might remain hidden in monolinguals (Kroll *et al*., 2012).

Preparing words in speech production is normally a fast and accurate process, although the underlying structure is exceedingly complex (Levelt, Roelofs, & Meyer, 1999). A range of linguistic stages are involved in speech production. Speech is the final expression of concepts and sensations, translated into a linguistic form that involves lexical, syntactic, morphological, phonological and phonetic encoding before the beginning of articulation (Caramazza, 1997; Dell, 1986; Levelt, 1989; Levelt *et al*., 1999; Starreveld & La Heij, 1995, 1996). According to most accounts, the activated semantic representations spread activation to the corresponding lexical representations. Because of this assumption, most models of word production have the shared assumption that lexical selection is a competitive process (Levelt *et al*., 1999; Roelofs, 1992; Starreveld & La Heij, 1996) that is necessary to decide which lexical representations should be selected for further processing (see also Caramazza & Hillis, 1991; Mahon *et al*., 2007, for a different account). (See also the following section for a summary of the most important theories in the context of monolingual speech production and in the context of production of signed utterances; see Wilbur, this volume.)

This chapter aims to review the main aspects of language production in bilingual speakers that also applies to speakers of more than two languages. After reviewing the most important theories regarding language activation in bilinguals we will present evidence of different language production capacities that bilinguals demonstrate in everyday life and we will focus on the mechanism that allows bilinguals to accomplish these complex tasks in an easy way.

## Theories of language production in bilinguals

Several theories of bilingual lexical processing assume a parallel activation of the lexicons of the two languages during reading, speaking and listening (Dijkstra, 2005; Hoshino & Thierry, 2011; Kroll, Sumutka, & Schwartz, 2005; Marian & Spivey, 2003). Despite the growing interest in bilingual language processing, some questions are still unsettled. Which linguistic levels of the two languages are activated during bilingual lexical selection and how do these levels interact? How does a bilingual control their two languages during speech processing, selecting the lexical items he/she intends to produce, in the language in which he/she wants to communicate? Thus, one of the central issues regarding bilingual speech production can be formulated in the following way: How do bilinguals retrieve words from one of the two languages selectively when both words express the same conceptual content? For example, how does an English-Spanish bilingual employ the word *mariposa* in one situation and *butterfly* in another situation, even though both words have the same meaning?

Several perspectives have been proposed to explain the organization of the lexical system in bilinguals. Finkbeiner, Gollan, and Caramazza (2006) have labeled the difficulty to decide between the two translations equivalent lexical nodes that are activated with a common semantic representation, as the "hard problem" of bilinguals. Traditionally, two different models have been proposed as solutions to this problem: those assuming selective activation of the intended language in bilinguals, and those assuming non-selective language activation.

## *Language-selective activation models*

Language-selective activation models (e.g., Costa & Caramazza, 1999; Costa, Miozzo, & Caramazza, 1999; Finkbeiner *et al.*, 2006a, 2006b; La Heij, 2005) suggest that the selection of a lexical entry uniquely affects the lexical entry corresponding to the intended language. According to this view, lexical access in bilinguals should entail similar processes to those involved in monolingual speakers during lexical selection. The intention to speak in one language determines which candidates become active and the two languages are considered as being functionally separate. From this view, there may be activation of words within the language not in use, but the activation of those words does not make them candidates for selection (Costa *et al.*, 1999; Costa & Caramazza, 1999). Kroll, Bobb, Misra, and Guo (2008) consider the proposal of the language-selective model as a "mental firewall" where the language cue effectively signals the correct activated alternatives. In this context, Finkbeiner *et al*. (2006b) assume that the language cue acts to set the activation level higher for candidates in the target language, avoiding potential competition between them at the point where selection occurs. In a similar language-specific proposal, La Heij (2005) shifts the locus of the bilingual's hard problem up to the level of concept selection, and not at the lexical level (Costa *et al.*, 1999; Costa & Caramazza, 1999), suggesting that only selected concepts, or, more appropriately, preverbal messages, activate their corresponding lexical nodes. This process should be similar to what happens when monolinguals need to choose between seemingly equivalent words to express a concept with subtle differences in meaning. In this way, the selection of one of two translations equivalent lexical nodes will be similar to the selection of words used in different registers carrying similar meanings (Levelt, 1989; for example: slang, formal language, or euphemisms, see Figure 4.1a). According to La Heij (2005), no additional activation or inhibition processes at the lexical level are needed, because the preverbal message contains a language cue (Poulisse & Bongaerts, 1994) that ensures that the word in the intended language reaches the highest activation level.

Most of the evidence supporting these models (and others) generally comes from psycholinguistic paradigms, which allow for the study of language processing and the mental processes involved while producing words in real time. The picture–word naming interference task is a variant of the classic Stroop task (McLeod, 1991), in which participants are usually instructed to name a picture as quickly and accurately as possible while ignoring a superimposed distractor word. Since naming latencies are affected by the relationship between the picture's name and the

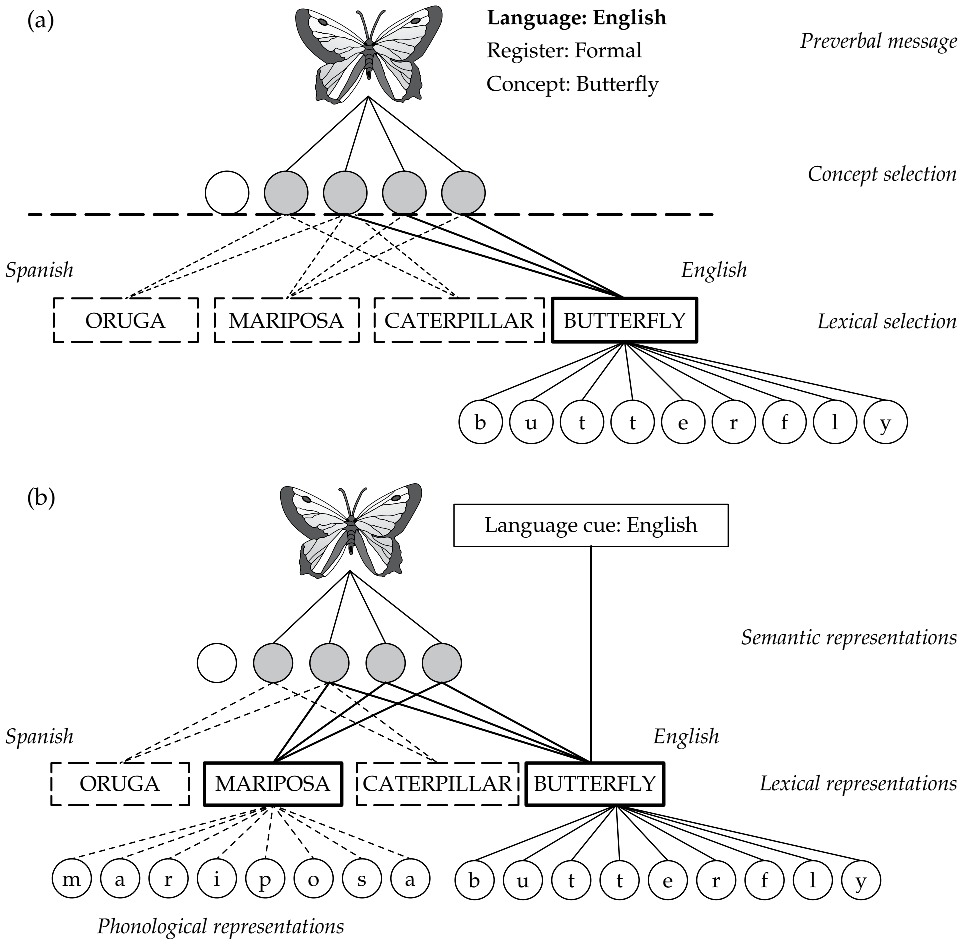



**Figure 4.1** **(a)** Schematic representation of *Language-Selective Activation Models* (adapted from La Heij, 2005) for an English-Spanish bilingual naming the picture of a butterfly in English. Due to the language cue (part of the preverbal message) the intended name will reach the highest activation level and will thus be selected. **(b)** Schematic representation of a general Non-Selective Language Activation Model (adapted from Costa *et al*., 2006 and Kroll *et al*., 2008). For an English-Spanish bilingual, when naming the picture of a butterfly in Spanish *mariposa*, the language not in use (English) will be also active. A language cue represents the intention to name the object in one of the two languages. Note that in the course of naming this picture, there may be several lexical representations activated (for example, semantically related items such as *caterpillar*, *spider*, *ant*, etc.) in both languages and there can be activation of abstract candidates at the lexical level or among phonological representations. (Figure continues on next page.) **(c)** Schematic representation of naming a picture with a Cognate name (left- lemon in Dutch *citroen* and in French *citron*) and a Non-Cognate name (right- apple in Dutch *appel* and in French *pomme*) for a Dutch-French bilingual. **(d)** Schematic representation of naming a picture with a Gender-Incongruent name (left: *cartepillar*, in Spanish *oruga*$_{FEM}$ and in Italian *bruco*$_{MAS}$) and Gender-Congruent name (right: butterfly, in Spanish *mariposa*$_{FEM}$ and in Italian *farfalla*$_{FEM}$) for an Italian-Spanish bilingual (adapted from Costa *et al*., 2000).

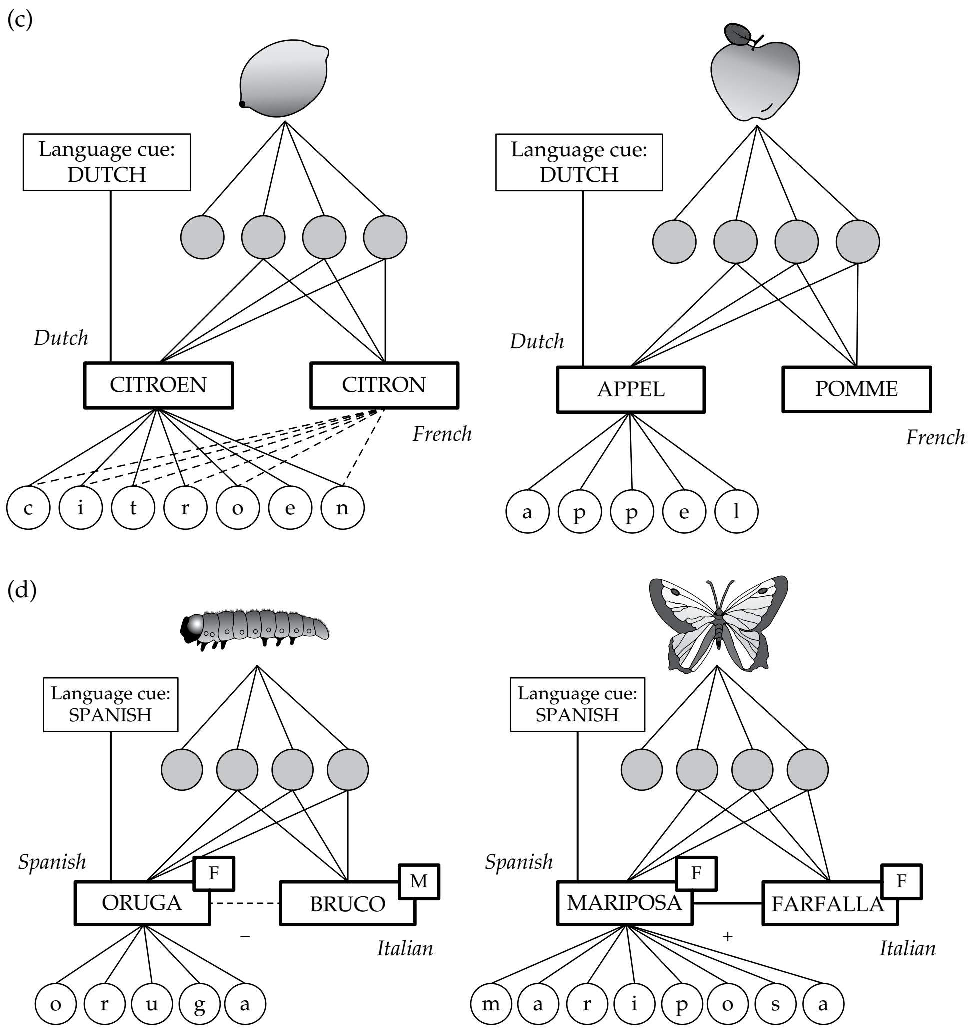



**Figure 4.1** (*Continued*)

distractor word, this paradigm became a useful tool to study the processes involved during lexical access (e.g., Glaser & Duengelhoff, 1984; Rayner & Springer, 1986; Rosinski, Golinkoff, & Kukish, 1975). In the bilingual version of this task, distractor words can be presented in one of the languages, in order to see how naming in one language is altered by the presence of the other. Various pieces of evidence for the activation of the non-target language when planning single-word utterances have been reported for Spanish and Catalan (Costa *et al*., 1999), Spanish and English (Costa & Caramazza, 1999), and Dutch and English (Hermans *et al*., 1998; Hermans, 2004) languages. In a series of picture–word tasks, Costa *et al*. (1999) reported lexical connections between the two systems of bilingual Catalan–Spanish speakers. They found interference effects when participants had to name pictures presented with semantically related words for both same and different language conditions, relative to when they were presented with semantically unrelated words. For example, when naming the picture of a table in Catalan, semantically related distractors such as *silla* and *cadira* (meaning "chair" in Spanish and Catalan, respectively) produced an equal amount of interference. Costa *et al*. (1999) proposed that the lexical selection mechanism only considers for selection those lexical representations belonging to the target language, and that distractor words are thought to activate lexical-semantic representations directly (via an input orthographic lexicon). According to these authors, this implies that is not necessary to make the assumption that selection is not language-specific, since the distractor words *silla* and *cadira* are equally able to activate the lexical node *cadira* (Costa *et al*., 1999; Costa & Caramazza, 1999). Costa and Caramazza (1999) also found a cross-language identity effect for less proficient English–Spanish bilinguals, whose two languages were more different than those of Catalan–Spanish bilinguals. Moreover, Costa *et al*. (1999) consider an additional argument in favor of language-selective models, that is, the finding that picture naming in L2 is facilitated by the presence of translation equivalent distractors in L1 compared with an unrelated word in L1 (Costa *et al*., 1999; see also Hermans, 2004). These findings led Costa *et al*. (1999) to conclude that the lexical selection mechanism must not consider the activation levels of lexical representations belonging to the non-target language, because, for example, the distractor word in L1 *mesa*, that should have been a strong competitor for naming the correct word in L2 *taula* (both meaning table in Spanish and Catalan, respectively), is not taken into consideration by the lexical selection process.

More recently, Guo and Peng (2006) have replicated and extended these behavioral findings using event-related potentials (ERP). In particular, they showed that in Chinese–English bilinguals (languages with different scripts) the same translation facilitation effect is also produced, and that this condition induces N400 amplitude reductions compared to unrelated distractor words. However, as argued by Kroll and Gollan (2014) and Hermans (2004), the finding of this translation facilitation effect within the picture-word interference task does not in itself provide unique evidence in helping to decide between selective and non-selective access models. Indeed, it is true that in the picture-word interference task, the two languages of a bilingual are involved and that the presence of distractor words in L1 could send back some activation to the L2 lexical representations by bottom-up processing of the lexical distractors (Costa, Caramazza, & Sebastián-Gallés, 2000; Costa, La Heij, & Navarrete, 2006). Therefore, in order to resolve this issue, other recent studies have restricted the experimental situation to only one language to corroborate that the two languages of a bilingual become activated in parallel (Morales, Paolieri, & Bajo, 2011; Paolieri *et al*., 2010a) or to examine whether bilinguals activate translations in the native language while they incidentally process words presented in their second language (Wu & Thierry, 2012).

## *Non-selective language activation models*

Most of the evidence, in fact, suggests that the intention to plan speech in one language alone is not sufficient to restrict activation to that language, thus providing support for language non-selective activation models in bilinguals. A large number of studies have tried to determine how lexical access occurs in bilinguals, and how the specific properties of the native language-L1 influence and affect linguistic processing in the alternative language-L2 (Kroll & de Groot, 2005). In language production there is evidence from a number of sources demonstrating that activation spreads to the two lexical systems of the bilinguals (Costa *et al*., 1999; Costa *et al*., 2000; Costa, 2005; Hanulovà, Davidson, & Indefrey, 2011; Hoshino & Kroll, 2008; Ju & Luce, 2004; Kroll *et al*., 2012; Kroll & Stewart, 1994) and that L2 processing is not completely autonomous, but is instead affected by the earlier acquired L1 at different levels of representation (Bordag & Pechmann, 2007; Costa *et al*., 2000, Hermans *et al*., 1998; Macizo & Bajo, 2006; Paolieri *et al*., 2010; Salamoura & Williams, 2007). Accordingly, most of the models postulate that a shared semantic system spreads activation to the lexical representations of both languages of a bilingual (Costa, 2005; Costa *et al*., 1999; de Bot, 2000; Green, 1986, 1998; Hermans *et al*., 1998; Kroll & Stewart, 1994; Poulisse & Bongaerts, 1994). As a consequence, lexical access in bilinguals is a complicated process since the selection of a given concept may activate, at least two different lexical entries (see Figure 4.1b). And although these models converge on the assumption of non-selective access, they reach different conclusions with respect to the locus of selection. Some argue that selection occurs at the level of the lemma or abstract lexical representation, whereas others suggest that cross-language interaction extends down to the phonology (Hermans *et al*., 1998, Kroll *et al*., 2008; Poulisse & Bongaerts, 1994).

Non-selective language activation models allow competition for selection such that candidates within and across languages actively compete with alternatives in the unintended language, which are eventually inhibited to allow accurate production to proceed (e.g., Green, 1998). Distinct cues for language membership may bias the access to different candidates in the intended language or allow those in the unintended language to be inhibited. Kroll, Bobb, and Wodniecka (2006) considers the possibility that the degree of the non-intended language activity will depend on several factors: the language of production, proficiency, the context in which spoken production occurs, features of the two languages, and the task being performed. For example, as argued by Kroll *et al*. (2008), during production in L1 there may be little evidence of L2 influence because L1 is more skilled, and has a faster time course (e.g., Bloem & La Heij, 2003). In contrast, during production in L2, multiple influences of L1 on L2 are possible (e.g., Costa *et al*., 2000, 2006; Hermans *et al*., 1998; Hoshino & Kroll, 2008).

Accordingly, several paradigms have been used to determine whether and at which point the language not-in-use is activated during language production.

*Cognate effects in picture naming:* The capacity to produce names correctly is one of the most important components of successful communication. Thus, it is not surprising that picture naming is one of the most broadly studied skills in psycholinguistics (Gollan *et al*., 2007). In simple L2 picture naming tasks, bilingual participants are asked to name pictures of single objects using their second language. Unlike the bilingual version of the picture-word interference task, this task allows for an assessment of the assumptions about the architecture of the bilingual lexical system without a forced activation of the language not in use.

One of the first phenomena that have been interpreted as revealing the presence of activation of the bilinguals' two languages is the "cognate advantage effect." Cognates are words that overlap across languages in their meaning and their phonological/orthographic form (e.g., Dutch *appel*, meaning apple in English). Cognate status has become a usual manipulation in bilingual research (for a review see Friel & Kennison, 2001). Usually, bilinguals are faster at naming a picture in one language when the picture's name is a cognate noun (Colomé & Miozzo, 2010; Costa *et al*., 2000; see Figure 4.1c as an example of a cognate noun in Dutch and French). To illustrate, Costa *et al*. (2000) observed shorter naming latencies for pictures with cognate names (e.g., Spanish *gato* and Catalan *gat*, meaning "cat" in English) compared with non-cognate names (e.g., Spanish *mesa* and Catalan *taula*, meaning "table" in English) in Catalan-Spanish bilinguals naming in both their dominant and non-dominant language. In contrast, no differences were found in Spanish monolinguals. Moreover, although the cognate effect was present in both languages, it was larger for naming in the non-dominant language.

The cognate facilitation effect has also been observed when features of the two languages are markedly different, such as when the two languages use different written scripts. For example, Hoshino and Kroll (2008) found identical cognate effects for both Japanese-English bilinguals and Spanish-English bilinguals, independently of the fact that there can be no orthographic overlap in Japanese and English (only the phonology can be shared), whereas for Spanish and English the overlap is possible at both the orthographic and phonological levels. The facilitation for naming pictures with cognate nouns suggests that the lexical candidate in the irrelevant language is activated at the level of phonology during the planning of single word utterances. However, recent ERP studies have found a cognate effect at early stages of lexical access with a remarkably similar pattern during both first and second language naming (Christoffels, Firk, & Schiller, 2007; Strijkers, Costa, & Thierry, 2010). Strijkers *et al*. (2010) suggest that this early cognate effect may reflect an induced word frequency effect rather than genuine parallel activation of the two languages. They conducted an ERP study with Spanish–Catalan bilinguals and Catalan–Spanish bilinguals performing a picture naming task in the L1 or the L2. ERP results showed early effects between 150 and 200 ms post-target presentation, both for frequency and cognate effects. Consequently, Strijkers *et al*. (2010) proposed that these two effects might have the same origin at the lexical level. In particular, due to the phonological overlap between the cognate words, lexical representations in both languages are strongly activated, and cognate lexical representations should have a higher frequency than non-cognate lexical representations, because the former are activated more often (irrespective of the language of utterance). However, Kroll & Gollan (2014) also consider the possibility that the increased frequency for cognates may have its origin in a process that requires the activation of both languages down to the level of phonology.

On the basis of cognate facilitation effects alone, some criticisms have been raised against the idea of parallel activation of the two languages in bilinguals (see also Costa *et al*., 2006, for a critical discussion). However, the study of cross-language interactions at different levels of representation could offer new and robust evidence to support the idea that L2 word production is not totally autonomous.

*Grammatical gender:* An issue that is probably one of the most neglected topics in the field of language research is the role of the syntactic information of words, such as grammatical gender, and the extent of the interactions between the two languages in the bilingual mental lexicon.

Grammatical gender is an inherent lexical feature of nouns that exists in many languages (Corbett, 1991) and is stored at a level that is different from conceptual and phonological representations (Levelt *et al*., 1999). Recent evidence suggests that this property is automatically activated in the process of lexical access (Alario *et al*., 2008; Cubelli *et al*., 2005; Paolieri *et al*., 2010b; but see Schiller & Caramazza, 2003). Thus, if grammatical gender is represented as an abstract nominal feature and is always available in lexical processing (Cubelli *et al*., 2005; Cubelli & Paolieri, 2008), an effect of grammatical gender should be observed when the names of the stimulus picture in the two languages share the same grammatical gender. Recently, a large number of studies demonstrating that grammatical gender interacts between languages in bilingual speech production (Bordag & Pechmann, 2007, 2008; Lemhöfer, Spalek & Schriefers, 2008; Morales *et al*., 2011; Paolieri *et al*., 2010a; see also Salamoura & Williams, 2007, using a translation task) support the idea that both languages are simultaneously active in the bilingual mind. For instance, Paolieri and colleagues (2010a) asked Italian-Spanish speakers to translate words or name a series of pictures in L2 by producing either the bare noun or the gender marked definite article ($el_{MASC}$ or $la_{FEM}$). Regardless of the task and the type of response, results revealed slower naming latencies with incongruent gender words between languages (e.g., $seta_{FEM}$, $fungo_{MASC}$, "mushroom" in Spanish and Italian, respectively) relative to congruent gender words (e.g., $bufanda_{FEM}$, $sciarpa_{FEM}$, "scarf"). To explain the grammatical gender effect it is important to note that both languages are assumed to be simultaneously active in the bilingual mind and that congruent gender nouns share more between-language information than incongruent gender nouns (see Figure 4.1d). Because of this, words with a similar gender value across languages are rapidly accessed with regard to words that do not match in gender, and faster response latencies are observed. Moreover, the advantage in processing words with the same gender across languages has also been found in German-Dutch (Lemhöfer *et al*., 2008), Czech-German (Bordag & Pechmann, 2007) and Greek-German (Salamoura & Williams, 2007) bilinguals, supporting the existence of a an interaction at the grammatical gender level between languages (for a different account see Costa *et al*., 2003). As with the cognate effect, the gender congruency effect in bilinguals has also been observed when features of the two languages are different, as is the case when the two languages utilize different written scripts, different gender values, and different systems of gender agreement such as Russian and Spanish (Paolieri *et al*., 2013).

*Phonology and articulation:* Additional cross-language effects have been found in the study of language interactions in bilinguals up to the phonological level. In a seminal study, Hermans and colleagues investigated interactions at the phonological level in Dutch-English speakers (Hermans *et al*., 1998; Hermans, 2004) using the bilingual version of the picture-word interference task. Participants were asked to name a series of pictures in their L2 (e.g., the picture of a mountain, *berg* in Dutch) while ignoring distractor words that could be phonologically related to the English name of the picture (e.g., *mouth*), phonologically related to the translation of the picture noun (e.g., *berm*, meaning "bench" in English) or semantically related (e.g., *valley*). Results revealed slower naming latencies in the two latter conditions, suggesting that the two lexical systems of bilinguals are simultaneously activated. The authors propose that lexical selection becomes more difficult when the more activated distractor word receives extra activation from either semantically related words or through translation words. In contrast, results revealed faster naming latencies when picture nouns and distractor words shared a phonological relationship. The authors argue that lexical selection is facilitated when the noun of the picture receives extra activation from phonological segments of the distractor word (but see de Groot & Nas, 1991).

Further evidence for co-activation at the phonological level comes from studies where language interactions have been observed when the two languages are completely different, for example when one language is written or spoken and the other language is signed (e.g., Morford *et al*., 2011) or in languages that utilize diverse scripts (e.g., Hoshino & Kroll, 2008; Guo & Peng, 2006; Paolieri *et al*., 2013) suggesting that the locus of the interaction is at the phonological level.

Taken together, all this evidence shows that cross-language interactions can be observed from the earlier to the final production stages, thus exerting their effects from the lexical to phonological word levels. Further, a study by Jacobs, Gerfen, and Kroll (2005) revealed that language non-selectivity continues to influence the execution of L2 speech. In their study, English-Spanish bilinguals who varied in their L2 proficiency were asked to name cognate and non-cognate words in Spanish. In the measures of speech execution (i.e., articulatory duration and voice onset time, VOT) only the least proficient learners showed an overall effect of cognate status, as revealed by their shorter articulatory durations and the more English-like VOTs for the cognate words.

*Transfer effects:* Although cross-language effects provide strong evidence for non-selective language activation in bilinguals, researchers have also explored whether some properties of the native language not only affect L2 production, but can also be transferred to the second language. For instance, Hoshino and Thierry (2011) found that Spanish-English speakers are slower at naming pictures in L2 while ignoring distractor words phonologically related to the Spanish or English name of the pictures, demonstrating that the transfer of phonology native information is present during word production in L2. Recently, Morales, Paolieri, Cubelli, and Bajo (2015), using the picture-word interference task, investigated how grammatical gender of the native language influences language production in bilinguals when speaking in a language lacking this property. Spanish-English bilinguals named pictures in English-L2 while ignoring L1 distractor words that had the same or different gender of the target Spanish translation. Language immersion was also manipulated by including a group of no-immersed bilingual speakers (in an L1 context) at the moment of the experiment (Spain), and another group immersed in a country where L2 was the dominant language of use (USA). The results showed that the group of non-immersed speakers was influenced by their native gender information even when they were asked to name the pictures in English, a language in which grammatical gender is not present for nouns. This pattern of results was not observed in the group of immersed bilinguals, suggesting that access to L1 is attenuated during language immersion. Similarly, congruency effects were not present in the monolinguals, indicating that the materials themselves were not a possible factor accounting for the results in the bilingual groups. The conflict in this task resulted from a co-activation of the Spanish and English linguistic systems, and from the transfer of Spanish gender features to English grammar. The transfer of Spanish-gender effects to English can be explained by assuming that the lexical entry corresponding to the target noun in L1 is active along with its corresponding translation during the production of the noun in L2, thus interfering with the distractor word. Since grammatical gender is not present in English, only the gender information of L1 could be responsible for the difference found in naming latencies between the two conditions. Therefore, when the target and distractor share gender, the interference delays the selection of the target noun and latencies are prolonged.

*Co-activation in sentential contexts:* Typically, co-activation has been observed with simple tasks, such as naming words (e.g., Schwartz, Kroll, & Diaz, 2007) or naming pictures (e.g., Costa *et al.*, 2000), but it has also been shown to persist in sentence context (e.g., Schwartz and Kroll, 2006). However, as argued by Hartsuiker, Pickering, and Veltkamp (2004), most research into this question has only considered the representations of words or concepts, and there has been little consideration of how bilinguals represent syntax, in particular whether they have two entirely separate syntactic stores, one for each language, or if some syntactic information is shared between the languages.

The structural (or syntactic) priming between languages refers to the tendency of speakers to mimic the same structural pattern as one that was previously encountered (Bock, 1986; Garrod & Pickering, 2004; Pickering & Garrod, 2004; for a review see Pickering & Ferreira, 2008). The idea is to consider bilingual syntactic representations within an extension of Pickering and Branigan's (1998) model. Pickering and Branigan (1998) argued that combinatorial nodes are shared between lemmas, so that all verbs that can be used in the passive, for instance, are linked to the same passive node within the lemma stratum, which is a level of lexical representation that encodes syntactic information (Levelt *et al.*, 1999). This proposal can be extended to bilingual lexical-syntactic representations in words that are also "tagged" for their language by the language node (Dijkstra & van Heuven, 2002; van Heuven, Dijkstra, & Grainger, 1998). Moreover, shared representations are "tagged" for both languages, and non-shared representations are "tagged" for one or the other language (Bernolet *et al.*, 2007; Hartsuiker *et al.*, 2004; see also Hartsuiker & Pickering, 2008). The activation of the lemma plus one of the combinatorial nodes leads to the activation of the grammatical structure, unspecified for language. A working hypothesis proposed by Pickering and Ferreira is that "bilinguals share as much grammatical information as they can" (2008, p. 41).

## Translation

Another task that bilinguals accomplish in daily life is rephrasing a message from one language into another—the translation task. This task is one of the most employed paradigms to study the dynamics of the activation of lexical information in bilinguals (Kroll & Stewart, 1994), including both the L2 to L1 "backward" translation and the translation from L1 to L2 that has been labeled "forward" translation (La Heij, Hooglander, Kerling, & van der Velden, 1996).

Although it is possible to distinguish different types of translation tasks (depending on the modality of the input, output or temporal parameters; for example, simultaneous, consecutive, or self-paced translation; see Christoffels & de Groot, 2005; Macizo & Bajo, 2006), most theories of translation agree that there are three common processes: analysis and understanding of the message source, language switching between the two linguistic codes, and production of the message in the target language. Therefore, the translator has to analyze the source message at the lexical, syntactic and discourse level, and then perform planning and lexical selection to correctly produce the message in the target language (Ruiz, Paredes, Macizo, & Bajo, 2008).

The Revised Hierarchical Model (RHM; Kroll & Stewart, 1994) assumes that the conceptual representations are shared among the languages, while the lexical representations are language-specific (see also Kroll & de Groot, 1997). This model was initially proposed to account for asymmetries in translation by late bilinguals for whom the L1 is still the dominant language. The L1 was hypothesized to have privileged access to meaning, whereas the L2 was thought to require mediation of the L1 translation equivalent until the bilingual acquired sufficient skill in the L2 to access meaning directly (Kroll, van Hell, Tokowicz & Green, 2010). Recent evidence suggests that semantic mediation can also be involved in backward translation from L2 to L1 (Duyck & Brysbaert, 2004; Sunderman & Kroll, 2006) despite the fact that forward translation, from L1 to L2, was more likely to engage semantics than backwards translation (e.g., Kroll & Stewart, 1994; Sholl, Sankaranarayanan, & Kroll, 1995; see Kroll *et al.*, 2011, and Kroll & Ma, this volume, for a discussion on this issue).

It seems to be generally accepted that translation is conceptually mediated in proficient bilinguals (e.g., Brysbaert & Duyck, 2010; de Groot, Dannenburg, & van Hell, 1994; de Groot & Poot, 1997; Francis & Gallard, 2005; La Heij *et al.*, 1996; Kroll *et al.*, 2011). Christoffels, Ganushchak, and Koester (2013) used ERPs to investigate the temporal course of translation production in Dutch-English bilinguals, a task in which a word is presented in one language and participants have to produce the translation equivalent in the other language. This task has high ecological validity, since it combines word comprehension and word production in two different languages, requiring a considerable amount of language control. In order to increase the need for cognitive control during translation production, Christoffels *et al.* (2013) presented participants with interlingual homographs (IH) or "false friends," stimuli in which the same orthographical form has different meanings in two languages, (e.g., *room* in Dutch means cream in English). The IHs were translated at a much slower rate, and elicited more errors and more negative N400 amplitudes when compared with control words. These results suggest that participants were not able to prevent the activation of the irrelevant meaning, and they are interpreted in terms of an increased lexical-semantic competition with these stimuli (see also, Macizo, Bajo, & Martín, 2010). Interestingly, their results also showed differences in the amplitudes of the P2 and N400 components depending on the translation direction, with a combination of a larger P2 amplitude for L1→L2 translation and a larger N400 amplitude for the opposite L2→L1 translation. No differences in ERPs between languages and word types were found in a naming experiment where participants read aloud the same words in L1 or L2. Christoffels *et al.* (2013) conclude that the task goal appears to influence how words are processed. Similar conclusions were reached by Macizo and Bajo (2006) in a study using bilinguals and professional translators. In their study, they showed that when participants read for translation, on-line and global comprehension was affected by lexical ambiguity, cognate status of the words, and memory load. However, when participants were asked to understand and repeat the sentences, these manipulations did not have any effect. Similarly, Ruiz *et al.* (2008) extended these findings to syntactic processing by manipulating the syntactic congruency between the presented sentences and their translations. Participants were faster reading the syntactically congruent sentences when they were reading for later translation, indicating that they were looking for syntactic matches while processing the source text.

This pattern of results provides support for the "horizontal" approach of translation (Gerver, 1976; see Danks & Griffin, 1997, for a similar approach). According to this view, translation involves establishing semantic matches between the lexical and syntactic entries in the two languages in a continuous parallel manner, so that the two languages of the bilinguals start to interact very early in the translation process from the moment in which comprehension starts. This view contrasts with the proposal of a "vertical view," in which comprehension and reformulation are independent processes that proceed in a sequential manner, and therefore they should impose similar demands on resources (Seleskovitch, 1976, 1999; Seleskovitch & Lederer, 1995). Note that the horizontal view is consistent with the non-selective activation hypothesis in which the two languages of the bilinguals interact during bilingual production.

## Bilingual cognitive control

The advantage held by bilinguals in communicative competence relative to monolinguals is evident (see Bialystok, Craik, Green, & Gollan, 2009 for nonlinguistic advantages, and Kroll & Bialystok, 2013 for the consequences of bilingualism for both language and cognition). However, bilingualism also entails a number of disadvantages in language production, such as reduced verbal fluency scores (Gollan, Montoya, & Werner, 2002; Rosselli *et al*., 2000) and more retrieval failures than monolinguals (Gollan & Acenas, 2004; Gollan & Silverberg, 2001). Importantly, these disadvantages are found even when bilinguals are tested in their L1 (e.g., Gollan & Acenas, 2004). Experimental evidence has shown that in bilinguals, conceptual activation spreads not only to the lexical entries corresponding to the language in use, but also to the lexical system of the alternative language, thus causing interference during selection of the intended lexical entries (Hoshino & Thierry, 2011; Kroll & Stewart, 1994; Kroll, Sumutka, & Schwartz, 2005). Hence, it is important to identify which cognitive processes are involved in the control of languages in bilinguals and how people speaking several languages select the appropriate lexical entries, given the interference caused by the simultaneous activation of the two languages.

Spontaneous slips of the tongue are interesting sources of information when testing theories of speech production (Dell, 1995). It is generally agreed that semantically related lexical errors reflect co-activation of semantically related lexical candidates during a conceptually driven retrieval process (e.g., Garrett, 1984; Vigliocco, Vinson, Paganelli, & Dworzynski, 2005). In the context of a bilingual's production, the presence of L1 intrusions in L2 production has been considered to support the parallel activation of the two languages activated from the same semantic system. This activation may lead to a malfunction of the lexical selection mechanism, thus selecting the translation in the non-intended language instead of the target word from the proposed language (Poulisse & Bongaerts, 1994; Poulisse, 1999, for a review). As shown by Dijkstra (2003), language errors might be particularly expected in multilinguals, not least because considering that already monolinguals must be able to quickly select a word from a lexicon composed of more than 50,000 words (see Aitchison, 1987), whilst proficient bilinguals must have at least 10,000 additional L2 words to select from.

This means that, during language production, several different words may be the possible targets. It is surprising, however, given the large amount of lexical competition during bilingual production, that the costs associated with the capacity of processing more than one language seems to be relatively mild. High-proficient bilinguals seem to be able to master the control of the two languages in a very natural and efficient way (e.g., Costa *et el*., 2000; Gollan, Sandoval, & Salmon, 2011; Poulisse, 1999; Poulisse & Bongaerts, 1994). This would require the existence of a language control mechanism that monitors lexical selection (Costa *et al*., 1999). Moreover, the importance of a control mechanism has also been recognized in cases of aphasia: Bilingual speakers who suffered neurological damage cannot properly control language selection, leading to pathological language mixing (Green & Abutalebi, 2008, see also Calabria, Marne, Romero-Pinel, Juncadella & Costa, 2014).

Indeed, recent evidence suggests that bilinguals use general mechanisms of control to achieve errorless language-selective production (Green, 1998) that are thought to share some features with the more general executive control system (e.g., Abutalebi & Green, 2007, 2008). In relation to this issue, Gollan *et al*. (2011) observed that failures in language control increase with aging-related declines in executive control, providing robust evidence for the role of executive control in maintaining language selection.

The most important model proposed to explain mechanisms of language control in bilinguals is the Inhibitory Control model (IC) proposed by Green (1998). In the IC model, language processing involves different levels of control or regulation by modifying levels of activation of the items in the language networks.

A key concept in the IC model is the language task schema. The language task schema allows bilinguals to select the appropriate task by suppressing the non-intended task. A language task schema regulates the output from the word identification system by altering the activation levels of representations in that system and by inhibiting outputs from the system. In addition, an internal lexical-semantic mechanism exerts control by inhibiting the competing lexical representations from the non-intended language. The locus of word selection is the lemma level in Levelt *et al*.'s (1999) terms and selection involves the use of language tags, and resolution of the competition from the non-intended language requires attentional resources. An assumption of the IC model is that inhibition is proportional to the activation level of the words to be inhibited, so that the higher the activation of the competing entries, the higher the inhibition needed to solve it. Recently, Abutalebi and Green (2007) proposed that cognitive control allows the correct selection of the lexical item in the target language and to keep it free from non-target language interferences with the integration of separable neural systems. These systems include: (i) the prefrontal cortex, that is involved in the mechanism of language switching and language selection and is involved in executive functions, response selection, response inhibition, and working memory; (ii) the anterior cingulate cortex, that detects response conflict and triggers a top-down signal from the prefrontal cortex to modulate the non-target representations; (iii) the basal ganglia, involved in language and lexical selection; and (iv) the inferior parietal lobule, implicated in the maintenance of representations and the working memory process.

A variety of evidence supporting the importance of inhibitory control process comes from studies using the language switching paradigm, where the response language varies in an unpredictable manner during the task (Meuter & Allport, 1999). These studies usually reveal asymmetric language switching costs, since bilinguals take longer to switch into L1—the dominant language—than to L2—their non-dominant language (Costa & Santesteban, 2004; Kroll & Stewart, 1994;

Meuter & Allport, 1999; Thomas & Allport, 2000; see Bobb & Wodniecka, 2013 for a recent review on switching data). According to the IC model (Green, 1998), the use of L2 leads to stronger inhibition of L1, so that later switching to L1 (the participants' native language) is more costly because of the need to overcome stronger inhibition. Moreover, the inhibitory mechanism seems to act at specific levels of representation within the lexical system, as has been demonstrated by a series of recent studies (Macizo *et al*., 2010; Martín, Macizo, & Bajo, 2010). In these studies, Spanish-English bilinguals had to judge the relationship between a pair of words presented in L2. The results revealed that participants were slower in responding to the pairs when they included an homograph along with a word related to the L1 translation to the homograph (e.g., pie-toe, with *foot* being the Spanish translation of *pie*), suggesting that bilinguals could not avoid the influence of their native language while performing the task. More importantly, however, response times were even slower when subsequent trials included the English translation of the Spanish meaning of the homograph presented on the previous trial (i.e., foot-toe), indicating that lexical entries were specifically inhibited. Because of this inhibition, participants took longer to reactivate these words and responses times increased. Similarly, Levy and colleagues (Levy *et al*., 2007; but see Runnqvist & Costa, 2012) have adapted the so-called Retrieval-Induced Forgetting paradigm (Anderson, Bjork, & Bjork, 1994) to study how inhibitory control can act at specific levels within the lexical system to solve between-language competition in bilinguals. In their study, English-Spanish bilinguals were asked to name a series of pictures in L2 once, 5 or 10 times with the assumption that the greater the repetition of trials in L2, the higher the inhibition of L1. Following this phase, the accessibility to the corresponding words in the native language was measured using a rhyme test (e.g., shake-sn___, to recall the word snake), showing that naming pictures in Spanish 5 or 10 times led to decreased recall of the corresponding English names relative to those that were named only once. Moreover, presenting semantic cues (e.g., venom-sn___) did not produce the forgetting effect of repeatedly named pictures in L2, leading to the conclusion that phonological first language attrition arises from specific inhibition of the phonological native language representations during second language use. Additionally, Morales *et al*. (2011) showed how inhibitory processes are in charge of managing language control in bilinguals at the specific level of grammatical gender. In their study, the grammatical gender knowledge of Italian-L1 is inhibited while performing a task in Spanish-L2 where this representation produced interference and competition between languages. Again, this study was specifically designed to explore the involvement of inhibitory mechanisms in the resolution of gender effects in bilinguals. During task 1, Italian-Spanish bilingual participants produced the name of a series of pictures that could either be congruent or incongruent in gender between the two languages. Critically, the pictures were practiced 1 or 5 times. The main prediction was that the higher the number of naming trials in L2, the greater inhibition of L1, which would specifically act at the grammatical gender level in order to resolve the competition arising at the grammatical level (Levy *et al*., 2007). After this naming phase, participants were asked to produce, in their native L1, the definite articles of the same pictures that they previously practiced in L2. The results revealed a grammatical gender congruency effect that increased for those words practiced five times in L2. Therefore, these results support the existence of an inhibitory mechanism that might be involved in the suppression of the native gender features during language production in bilinguals. This procedure allows for the attainment of two independent indexes. First, the interference effect found in task 1 was taken as an index of non-selective activation and between-language connections at the level of grammatical gender. Second, the additional time observed in task 2 to reactivate the incongruent nouns more practiced in the previous task was taken as an index of the inhibition of the gender representations of the nouns. This methodological feature makes this procedure advantageous over other procedures that provide indirect measures of inhibition such asymmetrical costs in the code switching (Meuter & Allport, 1999). In addition, most of these procedures employed mixed language conditions (Christoffels *et al.*, 2007; Costa & Santesteban, 2004; Meuter & Allport, 1999; Verhoef, Roelofs, & Chwilla, 2009), which possibly increases the level of activation of the two languages and biases the cross-language interaction (Wu & Thierry, 2010). In Morales *et al.* (2011) study, however, task 1, which was critical for creating between-language gender competition, was carried out in only one language context (i.e., the participants' L2 Spanish). This procedure is in line with the body of empirical research that attempts to study language inhibition in bilinguals without affecting the language mode of the interlocutor during the critical conditions (Levy *et al.*, 2007; Macizo *et al.*, 2010; Martín *et al.*, 2010). In general, these studies suggest that inhibition acts at a local level to overcome lexical interference from the stronger language.

Recently, Guo, Liu, Misra, and Kroll (2011) and Misra, Guo, Bobb, and Kroll (2012) have provided evidence that inhibitory mechanisms could potentially operate either at a local level, inhibiting specific lexical candidates, or at a global level, entirely inhibiting one of the bilingual's languages (e.g., de Groot, 2011; de Groot & Christoffels, 2006; Neumann, McCloskey, & Felio, 1999). Guo *et al.* (2011) used functional magnetic resonance imaging (fMRI) to examine the neural correlates associated with a mixed-language naming task, which is assumed to require local inhibition, and a blocked-language naming task, requiring the inhibition of the global irrelevant language. They found that each of these conditions produced activation of neural areas associated with cognitive control and inhibitory processes with different patterns of brain activation: the dorsal anterior cingulate cortex and the supplementary motor area seemed to play central roles during language switching (local inhibition), whereas the dorsal left frontal gyrus and the parietal cortex appeared to be essential during language blocking (global inhibition) (see also Abutalebi & Green, 2007, 2008). Similarly, Misra *et al.* (2012) used both ERPs and behavioral measures to investigate the same issue in a more natural environment. In their task, participants were asked to name the same set of pictures in either L1 followed by L2, or in the reverse order. Under these conditions, we would expect to find facilitation in the form of repetition priming, because the pictures to be named in each language were identical. However, if naming in a given language produces the inhibition of the other language, then priming should be reduced or eliminated. They found that the hypothesized priming was observed when pictures were named in the L2 following the L1. In contrast, an inhibitory pattern was observed when pictures were named in the L1 following the L2 and that this pattern was maintained over the course of the two blocks of L1 naming. These results are taken as evidence for the existence of an inhibitory pattern that operates at a global level, affecting the entire language of the bilingual, and that this persists over time. This may contrast with other recent evidence that indicates that inhibition is applied specifically over words (i.e., at a local level) and that it does not last longer than 750 ms (e.g., Martín *et al.*, 2010). However, taken together, these different sources of complementary evidence suggest that bilinguals may differently overcome the cross-language competition by suppressing the non-intended language either at a global or local level, with the two types of suppression involving different time courses.

## *The role of language immersion*

The notion that learning a second language can lead to a loss of access to the native language (Seliger & Vago, 1991) has also been explored in the context of language immersion. Linck, Kroll, and Sunderman (2009) suggest that L2 immersion facilitates the learning of a second language as a result of the suppression of the native language. Consequently, the activation of the more dominant L1 is reduced, and its negative influence on L2 becomes attenuated. Linck and collaborators (2009) showed that L2 language immersion produced temporal inhibition of L1, improving L2 learning by attenuating the negative influence of L1. In their experiment, English-Spanish speakers, immersed in an L2 context, were exposed to a production and comprehension task. The results revealed that this immersed group outperformed classroom learners of Spanish non-immersed in a Spanish context. But, more importantly, their results also showed that immersed learners inhibited their L1 while living in the L2 context, supporting the notion that bilinguals must launch inhibitory processes to suppress one of the languages when using the other (Abutalebi & Green, 2008; Green, 1998). Morales *et al.* (2015), presented evidence of Spanish-English speakers being influenced by the grammatical gender of their native language during a production task in L2. In a picture-word task, the participants were slower at naming pictures in English when these were paired to distractor words that shared gender with the target noun. In contrast, this influence was not observed in a group of participants immersed in an L2 context at the time of the experiment, for a period of at least two years. In this case, their naming latencies were not affected by the gender relationship between the target and distractor noun, demonstrating that immersion can restrict the influence of the native language on L2 processing. This result is important because it adds to existing evidence showing that immersion experience modulates the activation of the more dominant language during spoken production, in congruency with the inhibitory account provided by Linck *et al.* (2009) and the IC model (Green, 1998), which further suggests that bilinguals need to inhibit the language not in use to enable selective language access. Therefore, although initially lexical entries of both languages are active, inhibitory control would be exerted on the more dominant and competing language (i.e., L1, usually), which in turn leads to a greater cost in reactivating the native language when it is again needed.

## *Expertise in translation*

Experience in simultaneous (verbal) interpretation also seems to modulate language co-activation. The role of expertise in translation has previously been explored in relation to the linguistic and cognitive processes involved in translation and interpreting tasks (Christoffels & de Groot, 2005; Christoffels, de Groot, & Kroll, 2006; Ibáñez, Macizo, & Bajo, 2010; Macizo & Bajo, 2006). Translators are a special type of multilingual individual not only because they usually master three or more languages at a very proficient level, but also because language use of each of these languages differs from that of other types of bilingual. Despite differences among the existing varieties of translation tasks, the main characteristic of the translation performance is that the translator has not only to understand and reformulate a message from one language to another, but also she/he has to maintain the two relevant languages active and to switch continually between them. Therefore, translators have to manage the activation of two languages and be continuously coping with the interference coming from the parallel activation of the two languages in the translation task. Although the evidence suggests that translators and bilinguals activate the two relevant languages during on line comprehension for later translation (Macizo & Bajo, 2006, Ruiz *et al*., 2008), results obtained by Ibáñez *et al*. (2010) suggest that bilinguals may differ in the way they negotiate their two (or more) languages. In their study, Ibáñez *et al*. asked professional translators and untrained bilinguals to read and understand sentences presented word by word, and repeat them in the language of presentation (Spanish: L1 or English: L2) once they had finished reading them. To explore the non-selective activation of both languages they introduced cognate words (e.g., zebra/cebra, in English/Spanish, respectively) in the sentences and compared their reading times to control non-cognate words. The presence of cognate effects (faster reading of cognates relative to control words) was taken as an index of between-language activation (Dijkstra, Grainger, & van Heuven, 1999; Kroll & Stewart, 1994; Macizo & Bajo, 2006). In addition, in order to explore the nature of the lexical selection mechanism, they adapted the language switching paradigm (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999) to a sentence reading task. Thus, the sentences were presented in Spanish (L1) or English (L2) in an unpredictable manner. The presence of an asymmetrical switching cost (larger switching cost in the dominant L1 than in the less dominant L2) was taken as an index of inhibitory control. This procedure also had the advantage of mimicking comprehension in a bilateral translation where translators listen and comprehend sentences uttered from two or more speakers of different languages. The results showed that translators, unlike the control bilinguals, were faster at processing cognate words as compared to control words. Furthermore, they did not show an asymmetrical language switching cost (index of inhibition), whereas the bilingual group showed slower responses when switching from their L2 to their L1 (asymmetrical switching cost). These results suggest that translators differ from bilinguals in the way they control their languages during comprehension, in that they kept them both active (cognate effects) and showed no evidence of inhibition. The results of the present study suggest that translators do not use inhibitory processes to control for the concurrent activation of their two languages. The fact that the translator showed cognate effects during reading suggests that neither local nor global inhibition was used to control their languages. To further provide support for this observation, Martín (2010) used the negative priming with the interlingual homographs procedure using professional translators. Consistent with the data reported by Ibañez *et al*. (2010), the translators showed evidence of language co-activation, since they were slower when interlingual homographs were presented during the first trial. However (and unlike bilingual controls), they did not show slower responses when the translation of the irrelevant meaning was presented on the second trial after presentation of the homograph on the first trial. Although not conclusive, the results concerning the cognitive abilities of professional translators seem to suggest that language control in translators is of a proactive nature and more related to monitoring and updating (Köpke & Nespoulous, 2006, Yudes, Macizo, & Bajo, 2012; see also Costa *et al*., 2006; Costa, Santesteban, & Ivanova, 2006, for discussion of this view for very balanced bilinguals).

The modulating role that both immersion and experience have in language selection is consistent with the idea recently proposed by Green and Abutalebi (2013) within the Adaptive Control Hypothesis, stating that the context in which bilinguals acquire and speak their L2 could determine how language selection proceeds. Such regulation requires sensitivity to external input and the capacity for internal direction. Green and Abutalebi (2013) argue that the neural networks that support bilingual language processes are necessarily tuned differently in response to the requirement to engage these processes differentially. For example, habitual code-switching bilinguals may engage inhibitory mechanisms differently to a bilingual who uses each language in separate contexts. Similarly, immersed individuals or translators may adapt their language selection mechanism according to the context of L2 use provided by their immersion experience or their professional work.

## Conclusions

The joint activation of the two languages in bilinguals requires an active mechanism that negotiates cross-language activation and facilitates language selection. Evidence suggests that the act of planning speech in a second language requires the inhibition of the native language, which then has negative consequences for speech planning in L1 (see Kroll *et al*., 2008). This idea has also received support from studies using ERPs and functional magnetic resonance imaging (fMRI) (e.g., Rodríguez-Fornells *et al*., 2005). Moreover, non-selective language activation does not appear to be a rule, since there are conditions that restrict speech planning to one language (Kroll *et al.*, 2006). One example is that being immersed in a second language country can constrain the activation of the native language by inhibiting its level of activation (Linck *et al.*, 2009). Overall, from this perspective it is not surprising that bilinguals develop abilities for negotiating cross-language competition that confers them enhanced cognitive control (Bialystok *et al.*, 2004; Costa, Hernández, & Sebastián-Gallés, 2008), since brain areas associated with inhibitory processing function seem to be recruited by bilinguals to select the appropriate language (Abutalebi & Green, 2007, 2008).

## Acknowledgments

The writing of this chapter was supported in part by doctoral research grant from the Andalusian Government (P08-HUM-03600) to Luis Morales; grants PSI2013-46033-P, PCIN-2015-165-C02-01, and PCIN-2015-132 from the Spanish Ministry of Economy and Competitiveness to María Teresa Bajo and Daniela Paolieri; and grant P08-HUM-03600 to María Teresa Bajo.

## REFERENCES

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20, 242–275. doi: 10.1016/j.jneuroling.2006.10.003.

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes*, 23, 557–582. doi: 10.1080/01690965.2011.613209.

Aitchison, J. (1987). *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Blackwell.

Alario, F. X., Ayora, P., Costa, A., & Melinger, A. (2008). Grammatical and nongrammatical contributions to closed-class word selection. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 34, 960–981. doi: 10.1037/0278-7393.34.4.960.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1063–1087. doi: 10.1037/0882-7974.19.2.290.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 931–949. doi: 10.1037/0278-7393.33.5.931

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging*, 19, 290–303.

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, 10, 89–129. doi: 10.1177/1529100610387084.

Bloem, I. & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, 48, 468–488. doi: 10.1016/S0749-596X(02)00503-X.

Bobb, S. C., & Wodniecka, Z. (2013). Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology*, 25, 568–585. doi: 10.1080/20445911.2013.792822.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387. doi: 10.1016/0010-0285(86)90004-6.

Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism:Language and Cognition*, 10, 299–314. doi: 10.1017/S1366728907003082.

Bordag, D., & Pechmann, T. (2008). Grammatical gender in translation. *Second Language Research*, 24, 139–166. doi: 10.1177/0267658307086299.

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after 25 years of service? *Bilingualism: Language and Cognition*, 13, 359–371. doi: 10.1017/S1366728909990344.

Calabria, M., Marne, P., Romero-Pinel, L., Juncadella, M., & Costa, A. (2014). Losing control of your languages: a case study. *Cognitive neuropsychology*, 31, 266–286. doi: 10.1080/02643294.2013.879443

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177–208. doi: 10.1080/026432997381664

Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, 349, 788–790.

Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 6, 5–35.

Cenoz, J., Hufeisen, B., & Jessner, U. (eds.). (2003). *The Multilingual Lexicon*. Dordrecht: Kluwer Academic Publishers.

Christoffels, I. K., & de Groot, A. M. B. (2005). Simultaneous interpreting: A cognitive perspective. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 326–348). Oxford University Press.

Christoffels, I. K., de Groot, A. M. B., & Kroll, J. F. (2006). Memory and language skill in simultaneous interpreting: The role of expertise and language proficiency. *Journal of Memory and Language*, 54, 324–345. doi: 10.1016/j.jml.2005.12.004

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain research*, 1147, 192–208. doi: 10.1016/j.brainres.2007.01.137

Christoffels, I. K., Ganushchak, L., & Koester, D. (2013). Language conflict in translation: An ERP study of translation production. *Journal of Cognitive Psychology*, 25, 646–664. doi: 10.1080/20445911.2013.821127.

Colomé, A., & Miozzo, M. (2010). Which words are activated during bilingual word production? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 96–109. doi: 10.1037/a0017677.

Corbett, G.G. (1991). *Gender*. Cambridge: Cambridge University Press.

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 308–325). Oxford University Press.

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Bilingualism: Language and Cognition*, 2, 231–244. doi: 10.1006/jmla.1999.2651

Costa, A., Caramazza, A., & Sebástian-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1283–1296. doi: 10.1037TO278-7393.26.5.1283

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, 106, 59–86. doi: 10.1016/j.cognition.2006.12.013

Costa, A., Kovacic, D., Franck, J., & Caramazza, A. (2003). On the autonomy of the grammatical gender systems of the two languages of a bilingual. *Bilingualism: Language and Cognition*, 6, 181–200. doi: 10.1017/S1366728903001123

Costa, A., Heij, W. L., & Navarrete E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, 9, 137–151. doi: 10.1017/S1366728906002495

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365–397. doi: 10.1006/jmla.1999.2651

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50, 491–511. doi: 10.1016/j.jml.2004.02.002.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly-proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32, 1057–1074. doi: 10.1037/0278-7393.32.5.1057.

Crystal, D. (2003). *English as a Global Language*. 2nd edition. Cambridge: University Press

Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R. (2005). Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language*, 53, 42–59. doi: 10.1016/j.jml.2005.02.007

Cubelli, R., & Paolieri, D. (2008). The processing of grammatical gender as abstract lexical feature. In *Language and cognition: A long story. Festschrift in honour of Ino Flores D'Arcais* (pp. 73–86). Padova, Italy: Cleup.

Danks, J. H., & Griffin, J. (1997). Reading and translation. In H. J. Danks, G. M. Shreve, S. B. Fountain, & M. K. McBeath (Eds.), *Cognitive processes in translation and interpreting* (pp. 161–175). Thousand Oak, CA: Sage.

De Bot, K. (2000). A bilingual production model: Levelt's 'speaking'model adapted. The bilingualism reader, 420–442. doi: 10.1093/applin/13.1.1.

De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. Psychology Press.

De Groot, A. M. B., & Christoffels, I. K. (2006). Language control in bilinguals: monolingual tasks and simultaneous interpreting. *Bilingualism: Language and Cognition*, 9, 189–201. doi: 10.1017/S1366728906002537.

De Groot, A. M. B., Dannenburg, L., & Van Hell, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language*, 33, 600–629. doi: 10.1006/jmla.1994.1029.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and non-cognates in compound bilinguals. *Journal of Memory and Language*, 30, 90–123. doi: 10.1016/0749-596X(91)90012-9.

De Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning*, 47, 215–264. doi: 10.1111/0023-8333.71997007.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence

production. *Psychological Review*, 93, 283–321.
Dell, G. S. (1995). Speaking and misspeaking. *An invitation to cognitive science*, 1, 183–208.
Dijkstra, T. (2003). Lexical processing in bilinguals and multilinguals: The word selection problem. In *The multilingual lexicon* (pp. 11–26). Springer Netherlands.
Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford University Press.
Dijkstra, A., Grainger, J., & van Heuven, W. J. B. (1999). Recognizing cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41, 496–518. doi: 10.1006/jmla.1999.2654.
Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5, 175–197. doi: http://dx.doi.org/10.1017/S1366728902003012.
Duyck, W., & Brysbaert, M. (2004). Forward and backward number translation requires conceptual mediation in both balanced and unbalanced bilinguals. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 889–906. doi: 10.1037/0096-1523.30.5.889.
Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006a). Lexical selection in bilingual speech production does not involve language suppression. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1075–1089. doi: 10.1037/0278-7393.32.5.1075.
Finkbeiner, M., Gollan, T., & Caramazza, A. (2006b). Bilingual lexical access: What's the (hard) problem? *Bilingualism: Language and Cognition*, 9, 153–166. doi: 10.1017/S1366728906002501
Francis, W. S., & Gallard, S. L. (2005). Concept mediation in trilingual translation: Evidence from response time and repetition priming patterns. *Psychonomic Bulletin & Review*, 12, 1082–1088. doi: 10.3758/BF03206447.
Friel, B. M., & Kennison, S. M. (2001). Identifying German–English cognates, false cognates, and non-cognates: Methodological issues and descriptive norms. *Bilingualism: Language and Cognition*, 4, 249–274. doi: 10.1017/S1366728901000438.
Garrett, M. F. (1984). The organization of processing structure for language production: Applications to aphasic speech. *Biological Perspectives on Language*, 172–193.
Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8, 8–11. doi: 10.1016/j.tics.2003.10.016.
Gerver, D. (1976). Empirical studies of simultaneous interpretation: A review and a model. In R. W. Brislin (Ed.), *Translation: Applications and research* (pp. 165–207). New York: Gardiner.
Glaser, W. R., & Duengelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 640–654. doi: 10.1037//0096-1523.10.5.640
Gollan, T. H., & Acenas, L. A. (2004). What is a TOT? Cognate and translation effects on tip-ofthe-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30, 246–269. doi: 10.1037/0278-7393.30.1.246.
Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society*, 13, 197–208. doi: 10.1017/S1355617707070038.
Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*, 16, 562–576. doi: 10.1037/0894-4105.16.4.562.

Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection *Psychological Science*, 22, 1155–1164. doi: 10.1177/0956797611417002.

Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition*, 4, 63–83. doi: 10.1017/S136672890100013X.

Green, D. W. (1986). Control, activation and resource. *Brain and Language*, 27, 210–223. doi: 10.1016/0093-934X(86)90016-7

Green, D. W. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition*, 1, 67–81. doi: 10.1017/S1366728998000133

Green, D. W., & Abutalebi, J. (2008). Understanding the link between bilingual aphasia and language control. *Journal of Neurolinguistics*, 21, 558–576. doi:10.1016/j.jneuroling.2008.01.002.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25, 515–530. doi: 10.1080/20445911.2013.796377

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge: Harvard University Press.

Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *NeuroImage*, 56, 2300–2309. doi: 10.1016/j.neuroimage.2011.03.049

Guo, T., & Peng, D. (2007). Speaking words in the second language: From semantics to phonology in 170 milliseconds. *Neuroscience Research*, 57, 387–392. doi: 10.1016/j.neures.2006.11.010

Hakuta, K. (2009). Bilingualism. In L. Squire (ed.) *New Encyclopedia of Neuroscience*. Elsevier.

Hanulovà, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes*, 26, 902–934. doi: 10.1080/01690965.2010.509946

Hartsuiker, R. J., & Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychologica*, 128, 479–489. doi: 10.1016/j.actpsy.2007.08.005

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish/English bilinguals. *Psychological Science*, 15, 409–414. doi: 10.1111/j.0956-7976.2004.00693.x

Heij, W. L., Hooglander, A., Kerling, R., & Van Der Velden, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory and Language*, 35, 648–665. doi: 10.1006/jmla.1996.0034.

Hermans, D. (2004). Between-language identity effects in picture-word interference tasks: A challenge for language-nonspecific or language-specific models of lexical access? *International Journal of Bilingualism*, 8, 115–125. doi: 10.1177/13670069040080020101

Hermans, D., Bongaerts, T., de Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1, 213–229. doi: 10.1017/S1366728998000364

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, 106, 501–511. doi: 10.1016/j.cognition.2007.02.001

Hoshino, N., & Thierry, G. (2011). Language selection in bilingual word production: electrophysiological evidence for cross-language competition. *Brain Research*, 1371, 100–109. doi: 10.1016/j.brainres.2010.11.053

Ibáñez, A., Macizo, P., & Bajo, M. (2010). Language access and language selection

in professional translators. *Acta Psychologica*, 135, 257–266. doi: 10.1016/j.actpsy.2010.07.009
Jacobs, A., Gerfen, C., & Kroll, J. F. (2005). Inhibiting first language phonology in planning and producing speech in a second language. Paper presented at the European Society for Cognitive Psychology, Leiden, The Netherlands.
Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, 15, 314–318. doi: 10.1111/j.0956-7976.2004.00675.x
Köpke, B., & Nespoulous, J. L. (2006). Working memory performance in expert and novice interpreters. *Interpreting*, 8, 1–23. doi: 10.1075/intp.8.1.02kop
Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, 25, 497–514. doi: 10.1080/20445911.2013.799170
Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, 128, 416–430. doi: 10.1016/j.actpsy.2008.02.001
Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9, 119–135. doi: 10.1017/S1366728906002483
Kroll, J. F., & de Groot, A. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives*. Mahwah, N.J.: Lawrence Erlbaum Associates.
Kroll, J. F., & de Groot, A. (2005). *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford University Press.
Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Valdés Kroff, J. R. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 229–262). San Diego, CA: Academic Press.
Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.), *The Oxford Handbook of Language Production*. Oxford University Press.
Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Language and Memory*, 33, 149–174. doi: 10.1006/jmla.1994.1008.
Kroll, J. F., Sumutka, B. M., & Schwartz, A. I. (2005). A cognitive view of the bilingual lexicon: Reading and speaking words in two languages. *International Journal of Bilingualism*, 9, 27–48. doi: 10.1177/13670069050090010301.
Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition*, 13, 373–381. doi: 10.1017/S136672891000009X.
La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll and A. M. B. de Groot (Eds.) *Handbook of Bilingualism*, (pp. 289–307). Oxford University Press.
Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, 59, 312–330. doi: 10.1016/j.jml.2008.06.005.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA, USA: MIT Press.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75.

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language. The role of retrieval induced forgetting during second-language acquisition. *Psychological Science*, 18, 29-34. doi: 10.1111/j.1467-9280.2007.01844.x

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, 20, 1507–1515. doi: 10.1111/j.1467-9280.2009.02480.x

Macizo, P., & Bajo, M. T. (2006). Reading for repetition and reading for translation: do they involve the same processes? *Cognition*, 99, 1–34. doi: 10.1016/j.cognition.2004.09.012

Macizo, P., Bajo, M. T., & Martín, M. C. (2010). Inhibitory processes in bilingual language comprehension: Evidence from Spanish–English interlexical homographs. *Journal of Memory and Language*, 63, 232–244. doi: 10.1016/j.jml.2010.04.002

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, 109, 163–203.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 503–535. doi: 10.1037/0278-7393.33.3.503

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, 6, 97–115. doi: 10.1017/S1366728903001068

Martín, M. C., Macizo, P., & Bajo, M. T. (2010). Time course of inhibitory processes in bilingual language processing. *British Journal of Psychology*, 101, 679–693. doi: 10.1348/000712609X480571

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25–40. doi: 10.1006/jmla.1998.2602

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, 67, 224–237. doi: 10.1016/j.jml.2012.05.001.

Morales, L., Paolieri, D., & Bajo, M.T. (2011). Grammatical gender inhibition in bilinguals. *Frontiers in Psychology*, 2:284. doi: 10.3389/fpsyg.2011.00284

Morales, L., Paolieri, D., Cubelli, R., & Bajo, M. T. (2015). Transfer of Spanish grammatical gender to English: Evidence from immersed and non-immersed bilinguals. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728914000017

Morford, J. P., Wilkinson, E., Villwock, A., Piñar, P. & Kroll, J. F. (2011). When deaf signers read English: Do written words activate their sign translations? *Cognition*, 118, 286–292. doi: 10.1016/j.cognition.2010.11.006

Neumann, E., McCloskey, M. S., & Felio, A. C. (1999). Cross-language positive priming disappears, negative priming does not: Evidence for two sources of selective inhibition. *Memory & Cognition*, 27, 1051–1063. doi: 10.3758/BF03201234.

Paolieri, D., Cubelli, R., Macizo, P., Bajo, M. T., Lotto, L., & Job, R. (2010a). Grammatical gender processing in Italian and Spanish bilinguals. *Quarterly Journal of Experimental Psychology*, 63, 1631–1645. doi: 10.1080/17470210903511210

Paolieri, D., Lotto, L., Morales, L., Bajo, T., Cubelli, R., & Job, R. (2010b). Grammatical gender processing in romance languages: evidence from bare noun production in Italian and Spanish. *European Journal of Cognitive Psychology*, 22, 335–347. doi: 10.1080/09541440902916803

Paolieri, D., Padilla, F., Koreneva, O., & Macizo, P. (2013). *Gender congruency effect in second language recognition and production: Evidence from Russian-Spanish bilinguals*. Poster presented at the 18th Meeting of the European Society for Cognitive Psychology, Budapest, Hungary.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39, 633–651. doi: 10.1006/jmla.1998.2592

Pickering, M. J., & Ferreira, V. S. (2008). Structural Priming: A Critical Review. *Psychological Bulletin*, 134, 427–459. doi: 10.1037/0033-2909.134.3.427

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–190. doi: 10.1017/S0140525X04000056

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production*. Amsterdam, The Netherlands: John Benjamins.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15, 36–57. doi: 10.1093/applin/15.1.36

Rayner, K., & Springer, C. J. (1986). Graphemic and semantic similarity effects in the picture-word interference task. *British Journal of Psychology*, 77, 207–222. doi: 10.1111/j.2044-8295.1986.tb01995.x

Rodríguez-Fornells, A., Van Der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience*, 17, 422–433. doi: 10.1162/0898929053279559

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142. doi: 10.1016/0010-0277(92)90041-F

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development*, 46, 247–253. doi: 10.2307/1128859

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., et al. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology*, 7, 17–24. doi: 10.1207/S15324826AN0701_3

Ruiz, C., Paredes, N., Macizo, P., & Bajo, M. T. (2008). Activation of lexical and syntactic target language properties in translation. *Acta Psychologica*, 128, 490–500. doi: 10.1016/j.actpsy.2007.08.004

Runnqvist, E., & Costa, A. (2012). Is retrieval-induced forgetting behind the bilingual disadvantage in word production? *Bilingualism: Language and Cognition*, 15, 365–377. doi: 10.1017/S1366728911000034

Salamoura, A., & Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: Evidence from Greek and German. *Bilingualism: Language and Cognition*, 10, 257–275. doi: 10.1017/S1366728907003069

Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language*, 48, 169–194. doi: 10.1016/S0749-596X(02)00508-9

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56, 153–171. doi: 10.1016/j.jml.2006.10.002

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, 55, 197–212. doi: 10.1016/j.jml.2006.03.004

Schwartz, A., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, 22, 106–129. doi: 10.1080/01690960500463920

Seleskovitch, D. (1976). Interpretation: A psychological approach to translating. In R. W. Brislin (Ed.), *Translation: Applications and Research* (pp. 92–116). New York: Gardner.

Seleskovitch, D. (1999). The teaching of conference interpretation in the course of the last 50 years. *Interpreting*, 4, 55–66. doi: 10.1075/intp.4.1.07sel

Seleskovitch, D., & Lederer, M. (Eds.). (1995). *A systematic approach to teaching interpretation*. Silver Spring, MD: RID.

Seliger, H. W., & Vago, R. M. (Eds.). (1991). *First language attrition*. Cambridge University Press.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological Science*, 6, 45–49. doi: 10.1111/j.1467-9280.1995.tb00303.x

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 686–698. doi: 10.1037/0278-7393.21.3.686

Starreveld, P. A. & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 96–918. doi: 10.1037/0278-7393.22.4.896

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, 20, 912–928. doi: 10.1093/cercor/bhp153

Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, 28, 387–422. doi: 10.1017/S0272263106060177

Thomas, M. S. C., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language*, 43, 44–66. doi: 10.1006/jmla.1999.2700

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39, 458–483. doi: 10.1006/jmla.1998.2584

Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition*, 110, 84–99. doi: 10.1016/j.cognition.2008.10.013

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: implications for language learning and language use. *Journal of Experimental Psychology: General*, 134, 501–520. doi: 10.1037/0096-3445.134.4.501

Wu, Y. J., & Thierry, G. (2010). Investigating bilingual processing: The neglected role of language processing contexts. *Frontiers in Psychology*, 1:178. doi: 10.3389/fpsyg.2010.00178

Wu, Y. J., & Thierry, G. (2012). How reading in a second language protects your heart. *The Journal of Neuroscience*, 32, 6485–6489. doi: 10.1523/JNEUROSCI.6119-11.2012

Yudes, C., Macizo, P., & Bajo, T. (2012). Coordinating comprehension and production in simultaneous interpreters: Evidence from the articulatory suppression effect. *Bilingualism: Language and Cognition*, 15, 329–339. doi: http://dx.doi.org/10.1017/S1366728911000150

# 5 Production of Signed Utterances

RONNIE B. WILBUR

Purdue University

## Introduction

To fully understand research on the production of sign language, we must review some issues related to similarities and differences in signing and speech. Unlike most spoken-language research, excepting minority/endangered languages, we also need to review the historical context, which has affected the framing of questions and the types of research conducted.

Briefly, in 1880 at the Conference on Education of the Deaf in Milan, a resolution was passed to ban sign languages from deaf education, declaring that "oral education was superior," despite acknowledged success of sign-based education for many decades. Sign language research did not begin again until 1960 when William Stokoe published his linguistic analysis of American Sign Language (ASL), identifying sign components as handshape, location (place of articulation) and movement (Stokoe, 1960). Research picked up steam in the 1970s and has continued ever since. Klima and Bellugi (1979) demonstrated that ASL had minimal pairs based on these components (and added orientation of the palm). They also identified morphological compounding, derivation, and inflection of various types, as well as creative use of language (e.g., performance and poetry).

However, the shadow of the ban on ASL did not disappear. Much research was conducted in light of questions that challenged the legitimacy of sign languages as languages, treated them as iconic pictures in the air or just gestures, or suggested that ASL was "incomplete" English and could be "improved" by using English word order, adding artificial signs to mirror English morphology, and even adding speech to "complete" the message. To some extent, these myths still exist, along with another myth that "sign language is universal." The results reported here may seem obvious to (psycho-) linguistically sophisticated readers, and many do parallel spoken language quite nicely. However there are modality differences

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

between speech and signing that affect how the full picture is drawn, and these must be dealt with. Only those aspects critically relevant to understanding the research will be dealt with here; readers are referred to Pfau, Steinbach, and Woll (2012) for a more complete overview of sign language research.

Another problem is the absence of video analysis technology or methodology comparable to that for speech (e.g., software like Praat; Boersma & Weenink, 2015). Digital analysis software is still in development (cf. ELAN, http://tla.mpi.nl/tools/tla-tools/elan/). To conduct phonetic level research requires special equipment, such as motion capture equipment, and there is no standard analytical procedure. This will necessarily limit what can be covered here. Sign language acquisition is covered in Chapter 30 of this volume; readers are referred to Corina and Spotswood (2012) for review of neurolinguistic studies of signing, speech, and gesture.

Most psycholinguistic research addresses a simple question: What is involved in native-like fluent use of sign language? The answer is not at all simple. First, production in the visual modality makes a difference on sign production itself (each will be discussed in more detail below). There is a combination of simultaneous and sequential information. Sign construction begins with simultaneous layering of handshape, palm/finger orientation, place of articulation, and movement. However, movement brings with it two *sequential* timing slots per syllable, one each for starting/ending specifications of movement (Brentari, 1998, 2012; Crasborn, 2012; Wilbur, 2011a). Furthermore, most sign languages tend to be predominantly monosyllabic, much like Mandarin (Klima & Bellugi, 1979; Wilbur, 2011a). In addition, there are markers (head/face/body, collectively "nonmanuals") that can be specified lexically, morphologically, syntactically, and/or semantically in addition to, and distinct from, non-linguistic affective marking (Anderson & Reilly, 1998; Wilbur, 2011b). This information is simultaneous, co-occurring while the hands make signs. However, simultaneity needs to be put in proper context: utterances are composed of signs in sequence. Thus, the modality of perception and production (signed vs. speech) can make a difference, referred to as the "modality effect," but it does not have to, nor does it always.

Another difference results from general differences among languages. Typologically, English and ASL are not in the same groups on a variety of parameters. Despite having basic SVO order, they nonetheless differ: ASL is pro-drop, English is dummy subject; ASL has preference for focus in final position, English permits stress movement for focus within a sentence; ASL uses verbal classifiers, English does not; English has overt definite/indefinite determiners, ASL uses vertical space (lower, higher, respectively); English requires verb tense, ASL does not; English marks case on pronouns, ASL does not. This list is far from exhaustive.

This chapter will consider production from a prosodic perspective, beginning at the lowest levels of prosody (syllables and their internal structure), progressing upwards to prosodic issues such as stress at both word (sign) and larger phrasal/sentential units. Because sign languages users are predominantly second language learners (whether deaf or hearing, very few are native, learning ASL in the home), a major prosodic issue is that of determining what contributes to signer fluency. Finally, we consider implications of syntactic and prosodic differences between ASL and English, especially for the pedagogical idea that deaf children should be educated with a simultaneous combination of speaking (English) and signing (English or ASL), which cannot be produced satisfactorily beyond simple-sentence level.

## Basic prosody in sign languages

There is widespread agreement that sign languages, like spoken languages, have hierarchically structured prosody. This means that there are different levels of prosodic structure and that higher levels of prosody are built on the structure of lower levels. Summaries can be found in Brentari (1998); Tang, Brentari, González, and Sze (2010); Sandler and Lillo-Martin (2006); Ormel and Crasborn (2012a). The standard Prosodic Hierarchy (1) holds for ASL (Sandler, 2012).

(1) mora > syllable > prosodic word > phonological phrase > intonational phrase > phonological utterance

Having said that, we immediately need to take a detour about non-sequential information being transmitted on the face, head, and body along with the hands. Everything the signer does while signing is visible, not all of which is grammatically significant. A clear example of communicatively important but not grammatically driven visible signals is display of signer emotion or evaluation of signed content. Signers use their faces to show emotional states ("affective use") such as happiness, surprise, sadness, anger, and so on, while at the same time including grammatical information. In particular, eyebrow position in ASL can be meaningfully lowered for *wh*-questions ("who, what, when, where, why, how many, which"), meaningfully raised in a several constructions (yes/no questions, topics, conditionals, relative clauses, others), and otherwise neutral. Brow position is also used to convey emotions—generally "up" for happy and surprised, and "down" for anger and disgust. Weast (2008) analyzed how affective uses interact with grammatical uses. She observed that emotional state *constrains* grammatical use: for example, raised eyebrows for yes/no questions are raised higher in happy and surprised states than in angry or sad states. Thus both pieces of information are present if you know what to look for.

When multiple nonmanuals are produced simultaneously, phonological formation must be distinct enough to permit easy determination of which cues are present and for how long. Features enabling this include: abrupt onset/offset; single/repeated; articulator: head (tilt, turn, nod), eyes (blink, gaze), mouth (upper/lower lip, corners, tongue), shoulder/body. Grammatical marker domains are beyond the scope of this chapter, but are related to which articulators are involved, whether they are domain or edge markers (e.g., blinks at phrasal end), and other factors. The appropriate marker choice and their proper domains are part of what fluent signers are expected to know. Their presence needs to be kept in mind even though much of the discussion here will focus on the hands.

## Production studies of syllable prosody

There is a general consensus that syllables exist in sign languages (Wilbur, 2011a; Brentari, 2012; Sandler, 2008; Van der Kooij & Crasborn, 2008). In both speech and sign, syllables can be counted, tapped, and measured for duration. Here we focus on a controversy concerning the structure of signed syllables.

Recall that some parts of a signed syllable are produced simultaneously (handshape, place of articulation, orientation) whereas movement unfolds over time. The impact of this combination is that, unlike spoken syllables with sequences of sounds hierarchically arranged into onset and rhyme (nucleus and coda), signed syllables lack this type of internal temporal organization. Instead, there is a division within syllables of those components that do not change during the syllable ("Inherent Features" in Brentari's model) and those that do ("Prosodic Features"). Prosodic features can include change of handshape, for example, the syllable is specified for one handshape at the beginning (initial timing slot) and another one at the end (second/final timing slot). In such cases, the movement consists of change from one shape to another. Other possibilities include change of contact with the body, change of palm orientation, and change of location. When stated this way, it might appear that the movement is in some way epenthetic, the emergent result of starting in one configuration and ending in another. From such a perspective, movement might not be phonologically specified because it is redundant with the specifications of initial and final handshape/location, and so on. Indeed such arguments have been made, but have been convincingly countered using phonological and experimental studies. There is still a need for movement specifications in the phonological representations of signs that have movements that are not simple "change from configuration one to configuration two" (see Brentari (1998) for detailed discussion).

Because what is actually produced is not exactly what the linguistic model represents (as seen, e.g., in actual speech as opposed to linguists' phonemic representations of speech), there has been debate about the linguistic representation of the internal structure of signed syllables, namely whether there are sequential segments, parallel to consonants and vowels in speech. Segmental models have been offered by Sandler (2012); Sandler and Lillo-Martin (2006), and earlier by Liddell (1984, 1990, 1993). The Sandler model treats syllables as composed of Location (L) and Movement (M) segments ("LML"). Liddell suggested the segments were Holds and Movements ("HMH"). In both models, movement is just the result of changing from the first segment to the last. From the prosodic model perspective, there are only two timing slots, so that sequences of three segments such as those predicted by the HMH/LML models do not occur. We turn here to the experimental evidence.

## *Syllable internal structure and representation*

When asked to tap to spoken syllables, listeners tap regularly, and usually to a particular location inside stressed syllables, namely to the location of the consonant release and onset of the vowel, referred to as the Perceptual Center (Allen, 1972; Marcus, 1975). If segmental models of signed syllables were correct, similar tapping behavior would be expected at some identifiable temporal location inside the syllable, corresponding to a significant visible event. If the prosodic model is correct, there should be no such perceptual center. A parallel tapping study conducted on ASL indicates that signers do not target a particular point in the syllable, supporting the idea that signed syllables lack internal Perceptual Centers (Allen, Wilbur & Schick, 1991; Wilbur & Allen, 1991). This result can only be predicted if the sign syllable is composed of constantly changing movement (smoothly changing muscular activity), meaning there is no single timepoint which attracts perceptual attention. Thus, at the perceptual/phonetic level there is no support for syllable-internal segmental structure comparable to speech.

## *Evidence for syllable structure from backward signing*

A phenomenon known as backward speaking provides further insights into syllable structure and phoneme awareness (Cowan & Leavitt, 1981, 1990). Studies of people who have the ability to "talk backward" allow researchers to determine levels at which such reversals can be made. Backward talkers segment spoken words into "phonemic or syllabic" units, and then reverse their order while maintaining the syntax of the intended sentence, and can use either orthography or phonology for reversal. For orthographic reversers, the word "terrace" is said as /ɛkarɛt/, including pronunciation of final "silent e" and adjusting the pronunciation of the letter "c" followed by back vowels to /k/. In contrast, phonological reversers say/saerEt/for "terrace," showing that they are reversing the segment order.

Lacking writing systems, orthographic reversal is unavailable for backward signing, thus we expect parallels to phonological reversals. If so, and if there were segments inside the syllable, we would expect to see segmental reversals. But this is not what is observed (Wilbur & Petersen, 1997).

The HMH model represents the sign THINK (viewable at http://www.lifeprint.com/dictionary.htm) as two segments M and H, with M having the feature "[approach] (to forehead)" and H having the feature "[contact]" (with the forehead). Reversing MH to HM should yield H[contact] M[approach](to forehead). Signers actually produce [contact] at the forehead followed by "move *away*" (from forehead), not the predicted result. This result is correctly predicted only if the starting ("not at forehead") and ending ("closer to forehead") locations of the "approach" movement itself were reversed. That is, if "approach" were instead treated as two timing slots with features [-contact] and [+contact] with respect to the forehead, the reverse [+contact] [-contact] results in movement starting in contact with the forehead and then moving *away* from the forehead. The HMH model separates movement toward the forehead from the ending contact with the forehead. To describe the actual production, we first take "approach" and combine it *with* contact, and then reanalyze the movement as a sequence of non-contact to contact; as a result, the two segments "approach" and "contact" no longer exist, having been replaced by a sequence of features: [-contact] on the first timing slot, [+contact] on the second timing slot.

Likewise, incorrect predictions concerning backward forms are obvious with the verb FLY (viewable at www.lifeprint.com/dictionary.htm), represented as the single segment M. With only one segment, the reverse should be *the same* as the original, because from a segmental perspective, there is nothing available to reverse (think about trying to reverse the English determiner "a"). But backward signing reverses the movement *direction* of FLY, comparable to the direction reversal in THINK. Proponents of segmental models could argue that the representation should be HMH and that backward production reverses the two Hs, but evidence for those two Hs would need to be provided. Even with such an argument, the lack of analogy with spoken sequences can be seen: the backward form of *cat* /kæt/is/tæk/, with the vowel unchanged. In the backward form of FLY, the movement *is* changed. Signers simply do not produce what is predicted by segmental models.

With this evidence, Wilbur and Petersen (1997) argue that movement is not *inside* the syllable (M in HMH/LML models), but that movement *is* the syllable, a conception of "syllable" that takes movement as a dynamic gesture with no linguistically meaningful internal specifications (see gestural phonology approaches, e.g., Mauk & Tyrone, 2008; Tyrone & Mauk, 2010). This view translates into support for Brentari's postulation of two timing slots with associated features (a prosodic analysis). Thus, speech and signing both have syllables, but their internal structure is radically different, reflecting differences between auditory and visual transmission.

## Production studies of stress and sentence level prosody

A logical next question is what happens to syllables under stress. While disyllabic signs, compounds, and multisyllabic forms resulting from reduplication do exist, most lexical items are monosyllabic. Common generalizations like "stress antepenultimate syllables" are not found in sign languages. Also, ASL does not have distinctive lexical stress parallel to the English pair/'permit/and/per'mit/. However there are a few rules that may be observed for locating stress in multisyllabic signs.

### *Locating stress in signs*

We again need to discuss differences between the two modalities. The speech signal is turned on/off by the speaker, meaning the listener receives speech surrounded by strategically placed silences. In contrast, the visual signal is constantly present even when the signer is not signing. This makes every movement of the signer's hands and face of potential linguistic importance. One consequence is that every sign is preceded and followed by transition movement: changes in handshape/location necessary to get from the end of one sign to the beginning of the next. Thus, there is much movement that tends to be ignored, perceptually by viewers and analytically by sign language researchers. This is also true with stress assignment.

Multisyllabic signs have three possible forms, all restrictive with respect to stress placement (stress is predictable) (Wilbur, 2011a). First, multisyllabic signs may result from lexicalization of reduplicated forms (two-lexical movements with transitional movement between; ASL signs NAME and CHAIR).[1] In these forms, only the first syllable is prominent. Second, a sign may be a lexical disyllable, that is, if the *morpheme* itself requires two syllables. There are two types. In one, the second movement must be rotated 180 degrees from the first (e.g., back-and-forth, side-to-side, up-and-down), that is, it is a *return* to initial position; in these, prominence is equal on both syllables. In the other type, second syllable movement is rotated 90 degrees from the first (a cross movement, e.g., vertical, then horizontal), such that there is a short transition between the first lexical movement and start of the rotated second movement. Prominence is also equal on both syllables in this second type, meaning the transition movement is ignored for stress assignment. Thus, all lexical disyllables have equal stress on both syllables. With the exception of disyllables (which are lexically exceptional in being marked as disyllabic), stress assignment at the lexical level follows the Basic Accentuation Principle (Kiparsky & Halle, 1977): stress the leftmost syllable/vowel. There is thus no modality difference at the syllable–metric structure interface.

Finally, a sign may have two syllables if it is a compound, but then the first syllable is reduced compared to the second. Like spoken languages, ASL compounds pattern with phrases, following the Nuclear Stress Rule (Halle & Vergnaud, 1987): stress the most prominent syllable of the rightmost lexical item. Thus, phrasal stress assignment in ASL also does not show a modality effect.

## *Locating stress in sentences*

One major difference between ASL and English is their sentence level flexibility with respect to main stress. English maintains canonical word order and can move stress to different sentence-internal locations to mark the main focus (contrast the pair "No, *John* left early" versus "No, John left *early*"). ASL prefers sentence-final stress, resulting in different word orders to ensure that focused constituents receive stress in final position (see Discourse Effects).

Languages like English allowing stress movement are considered [+plastic], whereas languages like ASL and Spanish are [-plastic], where [plastic] is a typological feature reflecting ability of a language to bring stress and information focus together by shifting stress. This difference, among others, makes speaking English and signing ASL simultaneously impossible for actual conversations (about which more below).

## *What does stress look like?*

In speech, three signal characteristics may be affected under linguistic stress: duration, pitch, and amplitude. In contrast, sign languages use duration and movement amplitude, but pitch is unavailable. However, other visible characteristics can be used, such as speed, acceleration, muscle tension, height in signing space, and facial expression.

Wilbur and Schick (1987) compared signs in stressed and unstressed positions within narratives. For example, the sign DIE is targeted as stressed in (2) and unstressed in (3):

(2) SHOCK IX-1. DISCOVER GOOD FRIEND DIE. THINK HEART-ATTACK, NOT-KNOW... SEEM SICK IX-3, NOT-KNOW IX-1.
"I was shocked! I found out a good friend died! I think he had a heart attack, but I'm not sure … he did seem sick, but I don't really know."

(3) POSS-1 FRIEND MARRY AGAIN. WIFE FIRST DIE LONG-AGO. NOW HAVE WIFE.
"My friend married again. His first wife died a long time ago. Now he has another wife."

Stressed signs are set off from surrounding signs by "sharper transition boundaries," produced vertically higher in the signing space, and show increased muscle tension compared to unstressed counterparts.

Thus for syllable structure and stress, there are similarities and differences across modalities. In both, syllables can be counted and measured, and stress assignment rules are very similar. However, modality makes a tremendous difference in syllable-internal structure and production of stressed syllables. Put in perspective, we can say that linguistically, from the syllable level up (words, phrases), languages function pretty much the same regardless of modality. However, from the internal structure of the syllable down (phonetically), modality makes a huge difference. In the remainder of this chapter, we focus on the differences, assuming that most similarities need not be mentioned.

# Production studies of signing rate effects

Changes in signing rate provide linguistic information about ASL sentence production. If a head nod is present at normal and slow rates but missing in fast, we can infer that it is not a grammatical necessity. Because we do not yet know exactly which things are grammatical and which are signers' expressiveness, everything is subject to detailed investigation, leading to rather slow progress in this area.

## *Blinking and breathing*

Speakers planning sentence production consider the phrasing of what they plan to say and whether to breathe at phrasal pauses. Instead, signing is produced independent of breathing but with visual constraints related to blinking (Grosjean, 1977, 1979; Grosjean & Lane, 1977). Eyeblinks for signing and breathing for speech share a number of common elements. Physiological constraints interact with linguistic functions and structure, regulating blinks (for signing) and breaths (for speech) within the linguistic utterance. Speakers do not breathe in the middle of words but signers may breathe anywhere (Grosjean & Lane, 1977). Speakers may blink anywhere whereas signers do not blink in the middle of signs (Baker & Padden, 1978).

Two kinds of blinks are relevant here. Voluntary blinks (slower, longer) are added to the signed signal, analogous, perhaps, to adding loudness to stressed items in speech; these blinks overlap with signs. Inhibited periodic blinks are physiological responses for eye wetting. Signers use this blink type at phrase boundaries. Interestingly, addressees tend to blink at grammatical boundaries before the signer gets there, indicating they anticipate these boundaries. Baker and Padden (1978) suggest this aids linguistic processing of incoming information. Comparing three signing rates, Wilbur (2009) found that at the faster rate, there were fewer syntactic junctures filled with blinks whereas slower signing showed blinks at more locations. Thus we find that blinks change with rate, supporting the view that they are interacting with the linguistic signal.

Parallel to studies of spoken languages, Grosjean and Lane (1977) found that the longest pauses in signed stories appeared at boundaries between two sentences, that shorter pauses appeared between constituents of conjoined sentences, and that the shortest pauses appeared between sentence-internal constituents. These results show that signed sentences are organized hierarchically with respect to syntax as reflected by prosody.

Bellugi and Fischer (1972) compared the time to relate a story in both speech and signing. The story took about the same amount of time to produce in both modalities. However, modality made a difference: 50% more spoken words than signs were needed, still produced in the same amount of time, since spoken words take less time to produce than signs. One implication is that, at some processing level, there is an optimum time or rate for transmission of information regardless of modality. The finding that signed English sentences (signs in English word order) increased story time by almost 50% can be interpreted as a potential problem for signed English usage, in terms of perception, production, and memory processing.

## Nonmanuals

Rate changes also affect other nonmanuals (Wilbur, 2009). Increased signing rate decreased not only sign and pause duration, but also brow raises, brow lowering, and eye blink durations (and number of blinks). However, not everything is affected by signing rate. In particular, when facial articulations (e.g., brow position) are present because of a lexical requirement (e.g., lowered brows on the sign STRUGGLE) or signer affective state, signing rate tends to have less effect. In contrast, nonmanuals affected by signing rate are syntactically and semantically determined. For example, raised brows can be required for several phrases in sequence, for example, a topic or conditional followed by a yes/no question. At fast speed, these raised brows may simply blend together, creating one longer brow raise (hence reducing number of brow raises). However, the signed material covered by them remains the same across different rates.

## Studies of prosody and signer fluency

Fluent signers understand and use coordinated face and hand articulations with appropriate prosodic phrasing. Assessing signer fluency is a critical sociolinguistic issue, both for determining who is fluent enough to serve as teachers or interpreters and for tracking progress of sign language learners. Unfortunately, algorithms for describing fluent prosodic structure are still in development.

### *Understanding prosodic structure*

We begin with experimental techniques for understanding prosodic cues, which include tapping and cue judgments. A tapping paradigm compared native ASL signers and sign-naive hearing English speakers (Allen, Wilbur & Schick, 1991; for speech, Allen, 1972). Subjects watched repeated signed narratives and were instructed to "tap the rhythm" with a metal wand on a copper plate, generating acoustic signals for subsequent analysis. [One narrative was repeated, but before watching it again, subjects were instructed to "tap the syllables" (discussed earlier)]. Results demonstrate that both groups tap rhythmically to signed stimuli. Stimuli features—repetition, primary stress, phrase final position—influence whether observers treat it as a rhythmic beat. However, sign-naive subjects routinely tapped to secondary stresses, whereas native signers ignored them, reflecting knowledge of the language.

Because signing and gesture occur in the same modality, it is sometimes not clear which is which; it is often difficult to know when sentences end. Comparing signer judgments with non-signers is one experimental technique for resolving these issues. González (2011) asked four groups (ASL signers, Hong Kong Sign Language (HKSL) signers, hearing non-signers, hearing second language (L2) ASL students) to identify ASL prosodic boundaries. Non-signers and L2 signers were more accurate with a broader range of boundary cues; surprisingly, HKSL signers were more accurate than ASL signers, indicating that knowing the language can result in distraction due to language processing. Likewise, Brentari, Nadolske, and Wolford (2012) identified relative strengths of boundary cues for native and hearing L2 ASL signers and non-signers: sign duration, holds, transition between signs, pause duration (hold plus transition), blinks, drop hands, and nonmanual position changes (brow, head, torso).

Brentari, González, Seidl, and Wilbur (2011) reported three studies of prosodic cue perception. In one, Deaf ASL users and hearing non-signers revealed strong sensitivity to visual cues. Presented with sign strings excerpted from larger contexts (example 4; target underlined), they were equally accurate at identifying presence of intonational breaks between two signs (e.g., BIG STILL), relying on presence of pauses. Non-signers also relied on drop hands or holds. The question arises of why non-signers perform so similarly to signers, and what role gesture familiarity might play in this performance.

(4) (a) ANIMAL TEND THEIR STRANGE. SNAKE BIG STILL MOVE FAST CAN. ALWAYS HAVE PLENTY EAT.
"Animals have strange characteristics. Big snakes still can move fast. [They] always have plenty to eat."
(b) YESTERDAY MORNING MY GARAGE I SAW SNAKE BIG. STILL MOVE FAST CAN ALWAYS. CHASED IT.
"Yesterday morning I saw a big snake in my garage. [It] can still always move fast! [I] chased it all over."

Their second study tested nine-month-old hearing babies on the same stimuli with a looking-preference paradigm (habituation to a stimulus, then exposure to another stimulus; longer looking at the second stimulus indicates discrimination between the two). Despite lack of exposure to ASL and no extensive exposure to gestures, infants were sensitive to the visual cues used by the adults for intonational phrases.

The third study explored smaller prosodic units, with a cross-linguistic paradigm. Groups included users of ASL, Croatian Sign Language, and Austrian Sign Language, as well as hearing speakers of English, Croatian, and German. Stimuli consisted of 168 nonsense signs, 48 structured like real ASL signs (e.g., possible but non-existent "blick" in English), and 120 impossible lexical items (e.g., "bnick" for English). Combinations of handshape, place of articulation, and movement were tested for cue conflict that would lead subjects to respond whether a stimulus could be just one sign or had to be two. Groups used the same strategy: one value equals one word. Differences among signing groups reflected that stimuli used ASL handshapes whereas handshape inventories (hence sign structure constraints) differ across sign languages. No differences were found for sequential movement but signers were more sensitive to simultaneous information than non-signers. Brentari, González, Seidl, and Wilbur (2011) conclude that there are elements in the sign phonology inventory that are not distinct from gesture ("continuous"), and also elements that are clearly discontinuous. Prosodic cues (pauses, holds) are continuous such that non-signers perform like signers; these cues should then be universal to all sign languages (see Crasborn, van der Kooij, & Ros, (2012) for boundary cues in Sign Language of the Netherlands [NGT]). In contrast, cues like handshape are used differently by signers than non-signers, reflecting a discontinuity between sign language and gesture.

## *Signer fluency*

Parallel to research on spoken fluency, signing judges have rated signer fluency on a variety of criteria. Kantor (1978) asked judges to view fluent and non-fluent signers and determine: (1) whether each was native or L2 signer, and (2) what cues they used to make decisions. Judges easily identified native Deaf and L2 hearing signers, mentioning facial expression, exaggerated mouthing (early oral training), rhythm, speed, fluidity and use of space. Lupton (1998) identified important production cues as smooth and steady instead of choppy, hesitant, and jerky. Less fluent signers in her study also used excessive mouth movements and showed less eye contact, facial expression, and body movements. These latter aspects were correlated with syntactic abilities; thus the concept of fluency is not simply a rhythmic/motoric notion.

It is often assumed that adult sign learners already have motor coordination to fluently produce sentences. Lupton and Zelaznik (1990) demonstrated that this is not the case. Adult ASL L2 students did not achieve bilateral coordination of their hands in two-handed signs until about 12 weeks into their first semester course.

## Production studies of slips of the hand

Like slips of the tongue, signers can make "slips of the hand." Hohenberger and Leuninger (2012) report that in both ASL and German Sign Language, such slips involve phonetic features (handshape, movement, etc.) with handshape representing the largest portion. Wrong word selection, accounting for nearly half the slips, reflect higher level lemma activation. They note that sign language production seems to focus on stacked/vertical representations, related to simultaneity of features, whereas speech focuses more on serialized/horizontal representations, related to sequential production. Finally, they observe that all slip categories reported for speech are also found in signing, providing support for an amodal language processor. One difference is during error monitoring, signers focus on feedback from their internal representations whereas speakers focus on external representations (auditory feedback).

## Instrumental attempts to capture production prosody

Sophisticated technology parallel to speech research is newly developed for signing. Research has investigated place of articulation (sign lowering/raising) as an effect of height of preceding/following signs (Mauk, 2003; Mauk & Tyrone, 2008, 2012; Tyrone & Mauk, 2010; Russell, Wilkinson & Janzen, 2011). There have been two perceptual studies of signing and speech coarticulation effects (Grosvald, 2009; Grosvald & Corina, 2012). Other research has contributed to understanding motor disorders in sign language (Tyrone, 2007; Tyrone & Woll, 2008; Tyrone, Atkinson, Marshall, & Woll, 2009), initially studied by hand-made measurements (Brentari, Poizner, & Kegl, 1995).

Ormel and Crasborn (2012b) report that transitions between signs have lower velocity than sign movements, helping viewers distinguish lexical from transitional movements. Wilbur (1999) reports results for two effects, stress and phrase position. Instrumental results document significant phrase final lengthening of sign duration. Interestingly, stress did not affect duration, only peak velocity. These results were recently confirmed by Wilbur and Malaia (in press), who report one ASL signer producing the same narratives from Wilbur and Schick (1987). Comparing stressed and unstressed targets, peak velocity was confirmed for marking stress. This study also showed that productions of sufficient stimuli by a single signer can achieve comparable power and results to older carrier phrase methods with multiple signers, enabling the field to move toward analysis of more natural signing in longer narratives. Further such studies may lead to measures of sentence prosody to capture differences between fluent signers and learners, or movement-disrupted (e.g., Parkinson's) signing.

Instrumental research can also investigate other functions marked by kinematic variables. Malaia and Wilbur (2012) compared two groups of ASL verbs denoting different event structures (having an end-state, "telic," or not, "atelic"; Wilbur 2008, 2010). Signs denoting telic events (e.g., HIT, ARRIVE) had sharper "end-marking", produced by rapid deceleration to a stop, compared to atelics (TRAVEL). This end-state marking could be thought of as a verb suffix, comparable to English past tense, as for instance when "walked" is pronounced/walkt/, that is, as a single syllable. Thus, movement kinematics of monosyllabic signs are modified to simultaneously show multiple morphemes. Malaia, Wilbur, and Milković (2013) found similar end-marking on telic signs in Croatian Sign Language.

## Discourse effects on utterance production

Like speech, there are formal/informal signing "registers," larger signs and signing space for "shouting" across rooms and smaller for whispering (Quinto-Pozos, Mehta, & Reynolds, 2006). But unlike speech, all sign interaction is face-to-face, even when using web video. Thus, an orderly system of visual turn taking cues (hand position, eye contact) is necessary to ensure that participants do not miss anything (Baker-Schenk, 1983). When preparing to sign, waiting for a turn, or to interrupt, hands assume "half-rest position," generally waist level. Higher hands are a strong indication to the signer to yield the floor. A signer can ignore an interruption by not establishing eye contact with interruptor. Conversations cannot begin without eye contact, after which the addressee must continue to watch, but the signer is free to look away, for organizing thought or maintaining the floor. Sign learners must adjust to being *constantly* watched by their addressees, an uncomfortable feeling for most hearing non-signers.

### *Specific discourse differences between ASL and English*

ASL has flexible word order related to its preference for sentence stress/focus in final position in contrast to English. Both languages also have syntactic ways of putting information in focus and old information into the background. These differences are seen more fully when utterances are produced in discourse and narratives (Wilbur, 2012).

## *Foregrounding effects*

Presentation of information in a sentence is structured according to the producer's belief regarding the addressee's knowledge and attentional state. It must be clear what is new (focus), what is old/shared (topic), and what is intended as correction (contrast) (Lambrecht, 1994). Focus may be marked syntactically, lexically, or prosodically. Like Hungarian, with fixed focus position preverbally, ASL prefers sentence-final focus. Consider (5):

(5) What bothered me the most about his behavior is that he seemed to think no-one noticed.

Focus is conveyed by the syntactic form "what X is Y" (wh-cleft), with non-focused X "bothered me the most about his behavior," and focused Y "that he seemed to think no-one noticed." Primary stress could occur on four words inside the focus, depending on speaker intent: "he," "seemed," "no-one," and "noticed," with "noticed" being non-contrastive/non-emphatic and the others being contrastive/emphatic. ASL has a parallel wh-cleft construction (6) with the non-contrastive interpretation, with stress on NOTICE. The non-focused material is marked with a brow raise ("br") (Wilbur & Patschke, 1999). To arrive at interpretations with focus on "he," "seemed," or "no-one" requires complete rephrasing of the entire construction (see Wilbur & Patschke, 1999).[2]

<u>                                                            br</u>
(6) POSS-3 BEHAVIOR BOTHER IX-1 WHAT, IX-3 SEEM THINK NO-ONE NOTICE.
3sg-possessive behavior bother 1sg what, 3sg seem think no-one notice
"What bothered me about his behavior was he seemed to think that no-one noticed."

English and ASL also have a cleft construction to focus NPs: for English "it-cleft" (7a) and for ASL, "THAT-cleft" (7b). Unlike brow raise in (6), which is on *non-focused* material, here brow raise is on *focused* material (Wilbur & Patschke, 1999). THAT serves as a focuser.

(7) (a) Look! It's the thief that robbed McDonalds!
<u>      br</u>
(b) THIEF THAT, ROB MCDONALDS!
"It's the thief that robbed McDonald's!"

ASL has other lexical focusers, ONLY(-ONE) "only/only-one" and SAME "even" (Wilbur & Patschke, 1998). ASL again differs from English: (1) the focused item is put before the focuser sign and marked with brow raise, (2) the focuser sign may be marked with lean forward ("even") or back ("only"), and (3) head nod ("hn") may be added to emphasize the focused sign. Van der Kooij, Crasborn, & Emmerik (2006) report similar results for use of leans in NGT.

## *Backgrounding effects*

Backgrounding information in ASL is not a mere matter of de-stressing items. Example (7) showed wh-cleft backgrounding, in which backgrounded information is contained in a wh-clause with WHAT at the end, marked by brow raise, and set off from focus by a pause (the comma). Unlike English wh-clefts which use only "what," in ASL the wh-word can be almost any wh-word (8). Also, the wh-word comes after the non-focused material, not before. Note that English does not allow direct translation for WHERE and WHO, requiring "the place where" and "the person who."

(8) ______________________br

(a) ELLEN SEE KIM PUT BOOK WHERE, TABLE
"The place where Ellen saw Kim put the book was the *table*."
______________________br

(b) ELLEN SEE KIM PUT-ON-TABLE WHAT, BOOK
"What Ellen saw Kim put on the table was the *book*."
______________________br

(c) ELLEN SEE BOOK PUT-ON-TABLE WHO, KIM
"The person who Ellen saw put the book on the table was *Kim*."

ASL is a discourse configurational language (tracking new and old information that leads to structural changes), whereas English is not (Kiss, 1995) [alluded to with the [plastic] feature]. New information introduced into discourse as focus in one sentence becomes old in the next sentence. Old information may be omitted (pro-drop), used in pronoun form, or otherwise put in background constructions. Two studies illustrate this observation.

Analysis of "The fox and the stork" signed in ASL shows differing treatment of the fox, which invites the stork to dinner and is thus more salient as host, and the stork, which is newer as the guest. The stork is referred to overtly in 44% of subject slots as compared to 18% for the fox. The fox never occurs as direct object; the stork is referred to overtly in 25% of object slots. This analysis gives us an estimate of relative topicality, with the fox more topical (fewer overt references) and the stork less topical (referenced almost three times more) (Wilbur, 1994).

A study of information structure in Croatian Sign Language (HZJ) shows similar changes in word order for focus and reduction or omission of old information (Milković, Bradarić-Jončić, & Wilbur, 2007). Analyzing multiple signers' narratives elicited from pictures, several mechanisms show reduced contextual significance: establishment of spatial location and eye gaze to it to refer to old referents; signing using both hands with the non-dominant hand (H2) for backgrounded information; use of classifiers as pronominal indicators; and complex noun phrases (relative clauses) allowing a single occurrence of a noun to simultaneously serve multiple functions.

Example (9) illustrates a boy crossing the street and being hit by a car. The boy is first introduced with a noun sign, and then is never overtly mentioned again.

Next, the first car is introduced with the dominant hand but kept as backgrounded information on the non-dominant hand while the dominant hand describes the boy continuing to walk without seeing the second car. Then the non-dominant hand, now referring to the second car, shows the car hitting the boy. In the last sentence, the boy falls, but only the verb FALL is signed, as everything else has already been established and can therefore be omitted.

(9) [HZJ]
BOY WALK. SEE CAR1-CL (H1) WALK NOT-SEE CAR2-CL (H1) WALK (H1) FALL
(H2)CAR1-CL----------------- (H2) HIT
"The boy was crossing the street and saw one car. But he didn't see a second car coming from the other direction. That car hit him, and he fell."

These devices—permissible omission and simultaneous representation of multiple referents—permit information to be conveyed without separate signs for every referent, which would create longer constructions.

## Implications of comparing ASL and English

As indicated, English can shift stress within a sentence ([+plastic]) but ASL is discourse configurational ([-plastic], see example 8). ASL and English word order overlap in the simplest of sentences, basically subject-verb-object (SVO) assertions. ASL uses many alternate structures found in languages other than English. Most sign learners are not aware of this distinction though and, despite willingness to acknowledge that English and ASL have different syntactic structures, they nonetheless fall back on trying to put ASL signs into English word order. This (hearing person) signing variety is known as "signed English" or "pidgin sign English," and when paired with speech, as simultaneous communication (SC, or Simcom). Given the recent emergence of sign language linguistic research, there remains a belief among some (even deaf) people that English word order is the "right" way to sign, or that it makes sign language "better" or "more educated." These myths persist despite considerable effort to overcome them. Hence there are many hearing people interacting with deaf children who know only signed English and little to no ASL beyond the signs themselves. Professions in which this tends to be true are teachers of the deaf, audiologists, speech-language pathologists, and educational interpreters (with mainstreamed deaf children). As better trained signers enter these fields, more professionals will be able to appropriately switch between ASL and more English style signing as judged by the needs of their target addressees. Meanwhile, this situation has led to several investigations of how signed English and SC compare to spoken English and ASL, and to questions of signer fluency. In the next section, we deal with the first of these, and in the following section with signer fluency.

## *Trying to speak and sign at the same time*

Trying to speak and sign at the same time involves ASL signs in English word order, that is, signed English (SE) plus speech (SC). While it is impossible to simultaneously speak two languages, or to speak English and sign ASL at the same time, the use of SE is possible because the words are provide by ASL and the syntax is provided by English. What is produced is a coded form of English, allowing communication only if the deaf addressee knows enough English to understand what is signed. With SC, two articulation channels (spoken and signed) are used, interacting with each other in significant ways.

When producing SE, a signer must choose what to do about English words/morphemes that do not exist in ASL (because they are coded differently). Multiple varieties of SE exist (some actually have brand names); we will deal only generally with these differences. Signers can choose to fingerspell these morphemes, simply omit them, or use artificial signs selected from brand name programs. Different rules exist in these programs for deciding what to do; for example, whether one should fingerspell a word, or use one sign, e.g. "boat, ship", as part of another word, for example, "friendship"—meaning the sign for "friend" followed by the sign for "boat, ship." These decisions can make comprehension particularly difficult for the deaf viewer.

ASL conveys the same information through derivational and inflectional morphological modifications (aspect, agreement, classifiers) to sign movement (Wilbur, Klima, & Bellugi, 1983; Wilbur, 2015). ASL uses fewer prepositions than English because ASL can show certain meanings through the use of location in space. ASL also does not have separate signs for determiners "a/an" and "the" but has a spatial mechanism for indicating definite (lower in signing space) and indefinite (higher in signing space). When SE translates these forms into separate signs, each requires independent articulation. Thus, SE sentences have more signs per sentence and takes at least 50% longer to produce than either spoken English or ASL, which are roughly comparable to each other, reflecting their status as natural languages evolved for perception and production by humans (Bellugi & Fischer, 1972; Wilbur & Nolen, 1986).

## *Speaking and signing rates conflict*

When speech is added to SE, both speech and signing are disrupted. Most attention has been given to changes in the sign channel, assuming it causes the disruption (Johnson, Liddell, & Erting, 1989; Marmor & Petitto, 1979; Whitehead, Sciavetti, Whitehead, & Metz, 1995), with few studies concentrating on speech itself. One study targeted Key Word Signing (KWS), a form of SC in which only "key words" (subjects, objects, verbs) are signed while the entire sentence is spoken. Listeners of just the audio detected prosodic distortions in speech, such as increased pause length, greater number of pauses, slowed speech rate, and consonant prolongation (Windsor & Fristoe, 1989). If signing rate were the primary factor influencing speech, then in theory speech distortions should have been nonexistent in KWS.

Wilbur and Petersen (1998) compared sign and speech produced by two groups of fluent hearing signers (professionals from speech-language pathology, audiology, teachers of deaf children, and parents of deaf children), one group that knew ASL since birth but who used SE on a daily basis for professional functions (teaching, interpreting), and another that did not know ASL but had used SE on a daily basis for at least ten years. The groups, stimuli, and tasks were chosen to permit investigation of SC under best possible circumstances.

Conditions included speech, SE, and SC. For both groups, speech in SC took longer than speech alone, whereas SE alone took longer than in SC. Increased speech duration in SC was due to greater syllable duration, number of gaps, and gap duration. Decreased duration of signing in SC resulted from shorter sign duration, decreased gap duration, and increased sign omissions. In short, both groups speeded up their signing and slowed their speech. These data argue that, under the best of circumstances, speech input to deaf children in SC is distorted from normal speech. Furthermore, increased omissions in the signing channel means the English that is signed in SC is also not normal. There are significant differences between the two groups with respect to how they speed up their signing to keep up with their speech in SC.

As background, Mallery-Ruganis and Fischer (1991) argued that for effective communication using SC, the number of sign omissions is not as important as whether the meaning is distorted ("impermissible deletions"). ASL is a "pro-drop" language; like Spanish and Italian (unlike English and French), ASL can drop redundant pronouns. As mentioned, once an ASL topic is introduced, it may be omitted without effect on the meaning. But *content* words cannot be deleted without affecting the meaning. When producing SC, the ASL-fluent group omitted fewer categories that are generally not deletable, and used ASL knowledge to compensate for deletions. For example, they used sign modifications: instead of separate signs for "Show me the book," the sign SHOW was modified in location to agree with object "me" (toward the signer), reducing two signs to one but still showing verb and object. Establishing BOOK first in a location in space makes it definite, but changes word order compared to English. Other ASL verbs permit modifications for both subject and object, potentially reducing three signs to one. Similarly, the artificial sign –ING for English suffix "-ing" can be avoided by movement reduplication. Without such strategic knowledge, the group lacking ASL skills is at a disadvantage because the only strategy they have is sequential production of signs; to speed up their signing, they simply omit signs.

The differences between the two groups extend to use of the face. To illustrate what the ASL-fluent group of signers could do that the non-ASL group could not do, consider example (10) from Liddell (1978; notation modified):

```
      ____________________br
             ___________mm
(10) MAN  FISH[continuous]
```

(10) MAN FISH[continuous] "Is the man fishing with relaxation and enjoyment?"

The verb "to fish" displays movement modification (reduplication for continuous), lower mouth position meaning "relaxed, with enjoyment" ("mm"), and upper face marking (brow raise) for the yes/no question ("q"). The two signs convey what takes eight words/signs in English. Signers who knew ASL used these markers while producing SC. They produced raised brows on 47 of 48 yes/no questions, omitting the artificial question marker sign. Signers who did not know ASL used minimal and occasionally incorrect nonmanuals. For example, only 9 of 49 yes/no questions had brow raise; 26 were incorrectly marked with brow lowering. Other nonmanuals (blinks, negative headshakes) clearly differed between the two groups even though they were attempting to produce the same content.

In sum, when attempts are made to provide signing and speaking at the same time in the belief that this provides "full input" to deaf children, there is a significant difference between theory and practice. In theory English is provided in both channels. In practice, the signed channel is far from complete with respect to what is said, and even the speech is disturbed due to the presence of signing. These data argue that a full bilingual bimodal situation, in which there is both ASL and English at appropriate times and levels for the children, provides more comprehensible input in both languages.

This observation is supported by recent findings that performance on standardized tests of English and math is primarily predictable from deaf children's assessed ASL fluency (Hrastinski & Wilbur, 2016).

## Conclusion

Research on sign language production shows that as natural languages processed by human cognition and brain capacities, sign languages have developed in accordance with the possibilities and constraints of the visual modality in which they are produced. Like natural languages produced in the auditory modality, sentence structure and information flow vary in language-specific ways. At the same time, given production and perception constraints, it seems that all sign languages have developed as pro-drop languages to reduce production time. Another modality-induced difference would be the reliance on simultaneity of information as reflected by the use of nonmanuals (e.g., brow raise, leans) to convey multiple pieces of grammatical information without requiring longer sequences of separate signs. Comparison of the information transfer capacity of ASL shows that it is, as expected, in the same range as spoken languages (Malaia, Borneman, & Wilbur, 2016).

Much of the research described herein was motivated by two questions: (1) how are sign languages like other well-studied (albeit spoken) languages, and (2) what effect does the modality have on sign language perception and production? As we have seen, the answers are not always easy to explain. It is clear that sign languages have syllables, but their internal structure is not like those in spoken languages. It is clear that ASL has predictable lexical and phrasal stress, but that its sentential stress assignment ([-plastic]) is typologically different from English ([+plastic]). Also different is ASL's status as a pro-drop language, whereas English is a "dummy-subject" language. ASL uses more simultaneous mechanisms than English, such as representing more than one referent at a time (using two hands), representing a referent by using an already established spatial location, allowing a verb to move between two locations to indicate subject and object, using facial markers instead of separate signs (e.g., lowered brows to mark *wh*-questions), and many more not discussed here. In addition to the interesting contributions these differences make to understanding the nature of human language, there is a practical application to understanding what goes wrong when people attempt to speak and sign at the same time. The typological and production conflicts documented by these studies should make it clear that neither the human brain nor the motor system should be expected to successfully tackle such a task. Each language evolved to fit the modality of perception and production in a way that is comfortable for the humans who use it. Trying to put them together is like trying to speak two languages at the same time—it cannot be done.

## NOTES

1 ASL sign glosses are written in small capitals. Pointing signs to people or locations in space are considered indexes, abbreviated IX: IX-1 is first person, IX-2 second, and IX-3 third. POSS is possessive. There is no case marking or tense on verbs, but verbs may be marked for agreement with arguments, and may show temporal aspect marking (habitual, iterative, etc.). Further details are available in Pfau, Steinbach, and Woll (2012).

2 The domain of nonmanuals is shown by a line above the signs that co-occur with it. *br* indicates brow raise.

## REFERENCES

Allen, G. D. (1972). The location of rhythmic stress beats in English speech. Parts I and II. *Language and Speech* 15, 72–100, 179–195. doi: 10.1177/002383097201500110

Allen, G. D., Wilbur, R. B., & Schick, B. S. (1991). Aspects of rhythm in American Sign Language. *Sign Language Studies* 72, 297–320. doi: 10.1353/sls.1991.0020

Anderson, D. E., & Reilly J. S. (1998). PAH! The acquisition of adverbials in ASL. *Sign Language & Linguistics* 1, 117–142. doi: 10.1075/sll.1.2.03

Baker-Shenk, C. (1983). *A microanalysis of the nonmanual components of questions in American Sign Language*. Doctoral dissertation, University of California, Berkeley.

Baker, C. & Padden, C. (1978). Focusing on the nonmanual components of ASL. In P. Siple (Ed.), *Understanding language through sign language research*, 27–57. New York: Academic Press.

Bellugi, U., & Fischer, S. D. (1972). A comparison of sign language and spoken language: Rate and grammatical mechanisms. *Cognition* 1, 173–200. doi: 10.1016/0010-0277(72)90018-2

Boersma, P., & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.09, retrieved 1 June 2015 from www.praat.org/

Brentari, D. (1998). *A prosodic model of sign language phonology*. Cambridge, MA: The MIT Press.

Brentari, D. (2012). Phonology. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language. An international handbook (HSK - Handbooks of linguistics and communication science)*, 21–54. Berlin: Mouton de Gruyter.

Brentari, D., González, C., Seidl, A., & Wilbur, R.B. (2011). Sensitivity to visual prosodic cues in signers and nonsigners. *Language and Speech* 54, 49–72. doi: 10.1177/0023830910388011

Brentari, D., Nadolske, M., & Wolford, G. (2012). Can experience with co-speech gesture influence the prosody of a sign language? Sign language prosodic cues in bimodal bilinguals. *Bilingualism: Language and Cognition* 15, 402–412. doi: 10.1017/S1366728911000587

Brentari, D., Poizner, H., & Kegl, J. (1995). Aphasic and Parkinsonian signing: Differences in phonological disruption. *Brain & Language* 48, 69–105. doi: 10.1006/brln.1995.1003

Corina, D., & Spotswood, N. (2012) Neurolinguistics. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language. An international handbook (HSK - Handbooks of linguistics and communication science)*, 739–762. Berlin: Mouton de Gruyter.

Cowan, N., & Leavitt, L. (1981). Talking backward: Exceptional speech play in late childhood. *Journal of Child Language* 9, 481–491. doi: 10.1017/S0305000900004827

Cowan, N., & Leavitt, L. (1990). Speakers' access to the phonological structure of the syllable in word games. *Chicago Linguistics Society* 26, 45–59.

Crasborn, O. A, (2012). Phonetics. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language. An international handbook (HSK - Handbooks of linguistics and communication science)*, 4–20. Berlin: Mouton de Gruyter.

Crasborn, O. A., Van der Kooij, E., & Ros, J. (2012). On the weight of phrase-final prosodic words in a sign language. *Sign Language & Linguistics* 15, 11–38. doi: 10.1075/sll.15.1.02cra

González, C. (2011). *Perception of prosody in American Sign Language*. Doctoral dissertation, Purdue University.

Grosjean, F. (1977). The perception of rate in spoken and sign languages. *Perception and Psychophysics* 22, 408–413. doi: 10.3758/BF03199708

Grosjean, F. (1979). A study of timing in a manual and a spoken language: American Sign Language and English. *Journal of Psycholinguistic Research* 8, 379–405. doi: 10.1007/BF01067141

Grosjean, F., & Lane, H. (1977). Pauses and syntax in American Sign Language. *Cognition* 5, 101–117. doi: 10.1016/0010-0277(77)90006-3

Grosvald, M. A. (2009). *Long-distance coarticulation: A production and perception study of English and American Sign Language*. Doctoral dissertation, University of California at Davis.

Grosvald, M., & Corina, D. (2012). The perceptibility of long-distance coarticulation in speech and sign: A study of English and American Sign Language. *Sign Language & Linguistics* 15, 73–103. doi: 10.1075/sll.15.1.04

Halle, M., & Vergnaud, J. -R. (1987). *An essay on stress*. Cambridge, MA: The MIT Press.

Hohenberger, A., & Leuninger, H. (2012). Production. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language. An international handbook (HSK - Handbooks of linguistics and communication science)*, 711–738. Berlin: Mouton de Gruyter.

Hrastinski, I., & Wilbur, R. B. (2016). Academic achievement of deaf and hard of hearing children in an ASL/English bilingual program. *Journal of Deaf Studies and Deaf Education* 21, 156–170. doi: 10.1093/deafed/env072

Johnson, R. E., Liddell, S. K., & Erting, C. J. (1989). Unlocking the curriculum: Principles for achieving access in Deaf education. (Available from Gallaudet Research Institute, Gallaudet University, Washington, DC)

Kantor, R. (1978). Identifying native and second-language signers. *Communication and Cognition* 11, 39–55.
Kiparsky, P., & Halle, M. (1977). Towards a reconstruction of the Indo-European accent. In L. Hyman (Ed.), *Studies in stress and accent. Southern California Occasional Papers in Linguistics* 4, 209–238.
Kiss, K. E. (1995). *Discourse configurational languages*. Oxford: Oxford University Press.
Klima, E., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.
Kooij, E. van der, Crasborn, O. A., & Emmerik, W. (2006). Explaining prosodic body leans in Sign Language of the Netherlands: pragmatics required. *Journal of Pragmatics* 38, 1598–1614. doi: 10.1016/j.pragma.2005.07.006
Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge, U.K.: Cambridge University. doi: 10.1017/cbo9780511620607
Liddell, S. K. (1978). Nonmanual signals and relative clauses in American Sign Language. In P. Siple (Ed.), *Understanding language through sign language research*, 59–90. New York: Academic Press.
Liddell, S. K. (1984). THINK and BELIEVE: Sequentiality in American Sign Language signs. *Language* 60, 372–399. doi: 10.2307/413645
Liddell, S. (1990). Structures for representing handshape and local movement at the phonemic level. In S. Fischer & P. Siple (Eds.), *Theoretical issues in sign language research: Linguistics*, 37–66. Chicago: University of Chicago Press.
Liddell, S. (1993). Holds and positions: Comparing two models of segmentation in ASL. In G. Coulter (Ed.), *Current issues in ASL phonology: Vol. 3. Phonetics and phonology*, 189–210. San Diego, CA: Academic Press. doi: 10.1016/b978-0-12-193270-1.50014-5
Lupton, L. (1998). Fluency in American Sign Language. *Journal of Deaf Studies and Deaf Education* 3, 320–328. doi: 10.1093/oxfordjournals.deafed.a014359
Lupton, L., & Zelaznik, H. (1990). Motor learning in sign language students. *Sign Language Studies* 67, 153–174. doi: 10.1353/sls.1990.0020
Malaia, E., & Wilbur, R. B. (2012). Kinematic signatures of telic and atelic events in ASL predicates. *Language and Speech* 55, 407–421. doi: 10.1177/0023830911422201
Malaia E., Borneman J., & Wilbur R. B. (2016). Assessment of information content in visual signal: Analysis of optical flow fractal complexity. *Visual Cognition* 24(3), 246–251. doi: 10.1080/13506285.2016.1225142
Malaia, E., Wilbur, R. B., & Milković, M. (2013). Kinematic parameters of signed verbs at the morpho-phonology interface. *Journal of Speech, Language and Hearing Research* 56, 1677–1688. doi: 10.1044/1092-4388(2013/12-0257)
Mallery-Ruganis, D., & Fischer, S. (1991). Characteristics that contribute to effective simultaneous communication. *American Annals of the Deaf* 136, 401–408. doi: 10.1353/aad.2012.0485
Marcus, S. (1975). *Perceptual centres*. Doctoral dissertation, University of Cambridge.
Marmor, G. S., & Petitto, L. A. (1979). Simultaneous communication in the classroom: How well is English grammar represented? *Sign Language Studies* 23, 99–136. doi: 10.1353/sls.1979.0010
Mauk, C. (2003). *Undershoot in two modalities: Evidence from fast speech and fast signing*. Doctoral dissertation, University of Texas, Austin.
Mauk, C. E., & Tyrone, M. E. (2008). Sign lowering as phonetic reduction in American Sign Language. In R. Sock, S. Fuchs, & Y. Laprie (Eds.), *Proceedings of the 2008 International Seminar on Speech*

*Production*, 185–188. INRIA: http://issp2008.loria.fr/proceedings.html.

Mauk, C. E., & Tyrone, M. E. (2012). Location in ASL: Insight from phonetic variation. *Sign Language & Linguistics* 15(1), 128–146.

Milković, M., Bradarić-Jončić, S., & Wilbur, R. B. (2007). Information status and word order in Croatian sign language. *Clinical Phonetics and Linguistics* 21, 1007–1017. doi: 10.1080/02699200701659227

Ormel, E., & Crasborn, O. (2012a). Prosodic correlates of sentences in signed languages: A literature review and suggestions for new types of studies. *Sign Language Studies* 12, 279–315. doi: 10.1353/sls.2011.0019

Ormel, E., & Crasborn, O. (2012b). Kinematic properties of lexical and transitional movements in Sign Language of the Netherlands: A cyberglove study. In E. Efthimiou & G. Kouroupetroglou (Eds.), *Proceedings of the 9th International Gesture Workshop, May 25-27, 2011, Athens, Greece*, 80–83.

Pfau, R., Steinbach, M., & Woll, B. (Eds.). (2012). *Sign language. An international handbook (HSK - Handbooks of linguistics and communication science)*. Berlin: Mouton de Gruyter. doi: 10.1515/9783110261325

Quinto-Pozos, D., Mehta, S., & Reynolds, W. (2006). *The effects of register on the production of an ASL text*. 9th International Conference on Theoretical Issues in Sign Language Research (TISLR). Florianópolis, Brazil.

Russell, K., Wilkinson, E., & Janzen, T. (2011). ASL sign lowering as undershoot: A corpus study. *Laboratory Phonology*, 2, 403–422. doi: 10.1515/labphon.2011.015

Sandler, W. (2008). The syllable in sign language: Considering the other natural modality. In B. Davis & K. Zajdo (Eds.), *The syllable in speech production*, 379–408. New York: Taylor Francis.

Sandler, W. (2012). Visual prosody. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language. An international handbook (HSK - Handbooks of linguistics and communication science)*, 55–76. Berlin: Mouton de Gruyter. doi: 10.1515/9783110261325.55

Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139163910

Stokoe, W. (1960). *Sign language structure: An outline of the visual communication system of the American deaf*. Studies in Linguistics, Occasional Papers, No. 8. Buffalo, NY: University of Buffalo.

Tang, G., Brentari, D., González, C., & Sze, F. (2010). Cross-linguistic variation in prosodic cues. In D. Brentari (Ed.), *Sign Languages: Cambridge language surveys*, 519–542. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511712203.024

Tyrone, M. E. (2007). Simultaneity in atypical signers – Implications for the structure of signed language. In M. Vermeerbergen, L. Leeson, & O. Crasborn (Eds.), *Simultaneity in Signed Language: Form and Function*, 317–335. Amsterdam: John Benjamins. doi: 10.1075/cilt.281.14tyr

Tyrone, M. E., & Woll, B. (2008). Sign phonetics and the motor system: Implications from Parkinson's disease. *Signs of the time: Selected papers from TISLR 2004*, 43–68. Hamburg: Signum Press.

Tyrone, M. E., & Mauk, C. E. (2010). Sign lowering and phonetic reduction in American Sign Language. *Journal of Phonetics*, 38, 317–328. doi: 10.1016/j.wocn.2010.02.003

Tyrone, M. E., Atkinson, J. R., Marshall, J., & Woll, B. (2009). The effects of cerebellar ataxia on sign language production: A case study. *Neurocase*, 15, 419–426. doi: 10.1080/13554790902893097

Van der Kooij, E., & Crasborn, O. (2008). Syllables and the word-prosodic system in Sign Language of the Netherlands. *Lingua*, 118, 1307–1327. doi: 10.1016/j.lingua.2007.09.013

Weast, T. (2008). *Questions in American Sign Language: A quantitative analysis of raised and lowered eyebrows*. Doctoral dissertation, University of Texas, Arlington.

Whitehead, R. L., Schiavetti, N., Whitehead, B. H., & Metz, D. E. (1995). Temporal characteristics of speech in simultaneous communication. *Journal of Speech and Hearing Research*, 38, 1014–1024. doi: 10.1044/jshr.3805.1014

Wilbur, R. B. (1994). Foregrounding structures in ASL. *Journal of Pragmatics*, 22, 647–672. doi: 10.1016/0378-2166(94)90034-5

Wilbur, R. B. (1999). Stress in ASL: Empirical evidence and linguistic issues. *Language and Speech*, 42, 229–250. doi: 10.1177/00238309990420020501

Wilbur, R. B. (2008). Complex predicates involving events, time and aspect: Is this why sign languages look so similar? In J. Quer (Ed.), *Signs of the time: Selected papers from TISLR 2004*, 217–250. Hamburg: Signum Press.

Wilbur, R. B. (2009). Effects of varying rate of signing on ASL manual signs and nonmanual markers. *Language and Speech* 52, 245–285. doi: 10.1177/0023830909103174

Wilbur, R. B. (2010). The semantics-phonology interface. In D. Brentari (Ed.), *Cambridge Language Surveys: Sign Languages*, 357–382. Cambridge University Press. doi: 10.1017/cbo9780511712203.017

Wilbur, R. B. (2011a). Sign syllables. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds). *The Blackwell Companion to Phonology*, 1309–1334.

Wilbur, R. B. (2011b). Nonmanuals, semantic operators, domain marking, and the solution to two outstanding puzzles in ASL. *Sign Language & Linguistics*, 14, 148–178. doi: 10.1075/sll.14.1.08wil

Wilbur, R. B. (2012). Information structure. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language. An international handbook (HSK - Handbooks of linguistics and communication science)*, 462–489. Berlin: Mouton de Gruyter.

Wilbur, R. B. (2015). Word formation in sign languages. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word Formation. An international handbook (HSK - Handbooks of linguistics and communication science)*, 2225–2251. Berlin: Mouton de Gruyter. doi: 10.1515/9783110375732-037

Wilbur, R. B., &. Allen, G. D. (1991). Perceptual evidence against internal structure in ASL syllables. *Language and Speech*, 34, 27–46. doi: 10.1177/002383099103400102

Wilbur, R. B., Klima, E. S., & Bellugi, U. (1983). Roots: On the search for the origins of signs in ASL. *Chicago Linguistic Society*, 19, 314–336.

Wilbur, R. B., Malaia, E. (In press). A new technique for analyzing narrative prosodic effects in sign languages using motion capture technology In A. Huebel & M. Steinbach (Eds.), *Linguistic foundations of narration in spoken and sign languages*. Amsterdam: John Benjamins.

Wilbur, R. B., & Nolen, S. B. (1986). Duration of syllables in ASL. *Language and Speech*, 29, 263–280. doi: 10.1177/002383098602900306

Wilbur, R. B., & Petersen, L. (1997). Backwards signing and ASL syllable structure. *Language and Speech*, 40, 63–90. doi: 10.1177/002383099704000104

Wilbur, R. B., & Patschke, C. (1998). Body leans and marking contrast in ASL. *Journal of Pragmatics*, 30, 275–303. doi: 10.1016/S0378-2166(98)00003-4

Wilbur, R. B., Patschke, C. (1999). Syntactic correlates of brow raise in ASL. *Sign Language & Linguistics*, 2, 3–40. doi: 10.1075/sll.2.1.03wil

Wilbur, R. B., & Petersen, L. (1998). Modality interactions of speech and signing in simultaneous communication. *Journal of Speech, Language and Hearing Research*, 41, 200–212. doi: 10.1044/jslhr.4101.200

Wilbur, R. B., & Schick, B. S. (1987). The effects of linguistic stress on ASL signs. *Language and Speech*, 30, 301–323. doi: 10.1177/002383098703000402

Windsor, J., & Fristoe, M. (1989). Key word signing: Listeners' classification of signed and spoken narratives. *Journal of Speech and Hearing Disorders*, 54, 374–382. doi: 10.1044/jshd.5403.374.

# 6 Parity and Disparity in Conversational Interaction

JENNIFER S. PARDO

Montclair State University

## Introduction

Conversation in social interaction is a primary site for language use. However, Chomsky's emphasis on competence over performance set the mainstream agenda for early research in psycholinguistics, which focused on relatively encapsulated, mechanistic approaches to language. By separating language from culture, researchers made progress on specific questions about relationships between language and cognition. Of course, this revolutionary approach fostered a dialectical turn to research on social aspects of language use (Searle, 1976). In that vein, early research by luminaries such as Robert Krauss, Herb Clark, and Howard Giles demonstrated a profound influence of the target audience on language use in social interaction. While communicating linguistic messages, utterances also serve as a medium for nonlinguistic signals about a talker's identity and orientation to an addressee and their situation. This interplay evokes both parity and disparity across multiple aspects of conversational interaction. In some ways, talkers become more similar, but at the same time, they maintain and communicate their individual identities. Thus parity, while useful for mutual understanding of a linguistic message, is not necessarily carried through to every facet of spoken communication.

Conversation can be messy and unpredictable, making it an impractical target for rigorous scientific investigation. Controlled studies have introduced constraints on free-form conversation by using interviews, referential communication tasks, or completely constrained communication frames (such as semi-scripted sentence prompts). The upshot is that the form and function of a message is influenced in dramatic ways by attributes of a target addressee and social settings of language use. Accordingly, "the meaning of a message is more usefully thought of as something that is negotiated between a sender and receiver" than as something completely determined by a sender's aim (Krauss 1987, p. 86). Acknowledging the

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

importance of an addressee in spoken communication departs from traditional psycholinguistic conceptualizations of communication as message transmission from sender to receiver, but a complete appreciation for language use in its most natural setting demands that the field take this leap.

The following sections survey research on conversational interaction starting with early work from Robert Krauss' *Perspective-Taking* and Herb Clark's *Coordination* approaches to language use. Both approaches emphasize language use as a coordinated activity among interlocutors who attempt to build shared meaning. In a roughly parallel vein, Howard Giles and colleagues developed *Communication Accommodation Theory* (CAT), which explores social constraints on language variation during conversational interaction. This approach was driven by findings that talkers code-switch between dialectal variants in a way that often increases their similarity to an interlocutor's dialect. That is, a talker will adjust their pronunciation and vocabulary to approach or even match that of their target audience. While CAT enumerated many social constraints on speech accommodation, recent research on alignment/convergence in dialogue reveals the importance of these phenomena for constraining basic cognitive mechanisms of language use in social interaction. The final section offers a framework for tying these approaches together under a general entrainment perspective on coordinated behavior.

## Perspective-taking and coordination in language use

Some of the earliest studies demonstrating perspective-taking in referential communication were conducted by Robert Krauss and colleagues (and beautifully summarized in Krauss, 1987). These studies indicate that expectations about a target addressee influence the form of referential descriptions in ways that have implications for their communicative function. The basic paradigm involves a matching game with sets of cards depicting novel ambiguous figures that each evoke a complex descriptive utterance. This paradigm has been used both in non-interactive settings with hypothetical target addressees, as well as in task-oriented conversational studies.

In a series of non-interactive studies, talkers were asked to provide descriptions of ambiguous figures that could be used to identify them at a later time (Fussell & Krauss, 1989; Krauss, 1987; Krauss, Vivekananathan, & Weinheimer, 1968). In order to examine the impact of perspective-taking on message formulation, talkers provided separate descriptions of these figures for themselves, for a friend, and for a stranger. With regard to content, messages intended for strangers were longer, and comprised less diverse vocabulary and more literal descriptions than messages for the self. With regard to accuracy in matching the items with their descriptions, messages intended for the self yielded better performance than those for strangers, despite being shorter and less generic. Finally, when a target addressee was a specific friend, matching accuracy fell between that of messages for the self and for a stranger and was influenced by degree of friendship (Krauss, 1987). As a talker's knowledge of a target addressee increased (from stranger to friend to self in this case), descriptions became shorter and less literal, as well as more obscure to those who didn't share common ground with a talker.

When these figures are used in task-oriented conversational studies, talkers actively engage with target addressees in a matching task. In these tasks, each member of a pair of interlocutors receives a set of figures, but their figures are in different orders. One member is designated the Director in the task and the other member is the Matcher. The goal is for interlocutors to converse so that a Matcher can place their figures in the order dictated by a Director's set. Because the figures are ambiguous, the main activity of the task revolves around development of referential terms for the figures. As interlocutors played multiple rounds of the matching game with the same set of figures in different orders, descriptions of figures became shorter and more figurative, and matching accuracy increased, similar to the pattern found when comparing descriptions for strangers with those for friends and the self in non-interactive studies (Krauss, Bricker, Garlock, & MacMahon, 1977; Clark & Wilkes-Gibbs, 1986; Krauss & Bricker, 1967; Krauss & Weinheimer, 1964). Furthermore, degree of reduction in message length was influenced by whether or not a Matcher could provide verbal feedback to a Director at all (Krauss & Weinheimer, 1966), and whether feedback was delayed (Krauss & Bricker, 1967), indicating that Matchers actively participate in a process of referential description tailoring.

This tailoring by Matchers was demonstrated most dramatically in an interactive conversational study by Schober and Clark (1989; see also Clark & Wilkes-Gibbs, 1986; Wilkes-Gibbs & Clark, 1992). In their study, they contrasted an autonomous view of communication as a series of individual acts of language production and comprehension with a view that communication involves the accumulation of common ground between interlocutors through an active process of grounding (i.e., coordination), in which addressees tailor messages to be specifically designed for them. The study involved 10 pairs of talkers playing a matching game with ambiguous figures that were recycled over the course of six trials. On each trial, a Director and Matcher both had 16 figures, and the Director instructed the Matcher in placing 12 of them in a pre-determined order. This design permitted measures of number of words used to identify each figure, accuracy in figure identification, and timing of figure identification. In a crucial twist, the recorded conversations were presented to 40 additional Matchers (dubbed Overhearers), who completed the matching task in a non-interactive manner. Half of the Overhearers had access to entire conversations (Early Overhearers), while the other half could only access recordings from the third trial on (Late Overhearers).

Taking a viewpoint that language occurs in autonomous production and comprehension steps, there should be no difference in matching accuracy between (active) Matchers and Early Overhearers who had access to an entire conversation. Because both Matchers and Early Overhearers were strangers to Directors at the outset, the first referential descriptions were presumably designed for any generic stranger, and Early Overhearers had access to the same content as Matchers, including their feedback, questions, and so on. If common ground in communication merely accumulates via distinct acts of language production and comprehension, then shared content should yield equivalent performance between Early

Overhearers and Matchers. However, if, as Schober and Clark hypothesized, common ground is generated through an active process of grounding between interlocutors, with Matchers evoking messages that are specifically designed for them, then Early Overhearers should not perform as well as active Matchers.

During the task, the number of words per figure uttered by Directors started at an average of 73 in trial 1 and decreased steadily to asymptote at around 13 by trial 4. Likewise, the number of turns and amount of time spent per figure decreased from trial 1 to trial 6. This pattern contrasted with a slight increase in figure identification accuracy for both Matchers and Overhearers over the course of the trials. However, Matchers outperformed Early Overhearers in figure identification on all trials, with an average of 99% versus 88% correct identification, and Late Overhearers (who started with trial 3) fared even worse at 68% correct (chance performance could be as high as 20% for the last figure in a trial, presuming it was not already misidentified). An additional manipulation enabled some Overhearers to pause conversation recordings while completing the task, but their performance was no different from those who could not pause the recordings. With respect to timing of figure identification, Overhearers were much more variable relative to Matchers. Finally, these patterns were replicated in a second experiment in which Overhearers were present during the live exchanges, but unable to collaborate directly with the interacting talkers. According to Clark and Wilkes-Gibbs (Clark & Wilkes-Gibbs, 1986; Wilkes-Gibbs & Clark, 1992), both members of a conversing pair negotiate shared meaning through a process involving presentation, acceptance/rejection, and mutual recognition phases, which serve to minimize collaborative effort in grounding reference.

These findings challenge a traditional conceptualization of communication as message transmission from a sender to a receiver via an informational medium. According to the encoding-decoding paradigm, the sender is the cause of a message that a receiver simply decodes (see Krauss & Chiu, 1998; Krauss & Fussell, 1996). By adopting this paradigm, one can justify separating communication into distinct acts of language production and language perception that can be analyzed and studied separately. The organization of the field of psycholinguistics into language production and comprehension implicitly validates this perspective. However, research on language use in conversation belies this division by demonstrating that the form and efficacy of a message depend on expected and actual knowledge states of an intended audience. "[T]he traditional separation of the roles of participants in a verbal interaction into sender and receiver, speaker and addressee, is based on an illusion—namely that the message somehow 'belongs to' the speaker, that he or she is exclusively responsible for having generated it, and that the addressee is more-or-less a passive spectator to the event.... [T]he message, in the concrete and particular form it takes, is as much attributable to the existence of the addressee as it is to the existence of the speaker" (Krauss, 1987, p. 96).

These foundational studies demonstrated the importance of an addressee in message formulation, eliciting a turn to an ongoing round of research examining implications of perspective-taking and coordination for cognitive mechanisms of language use. For example, Schober (1993) demonstrated that talkers in a task involving differing spatial perspectives tended to use more egocentric frames of reference when interacting with a partner than when alone, presumably due to the potential for a live partner to provide feedback and/or to solicit clarification as needed. In more recent studies, the scope of referential communication has expanded beyond development of terms for ambiguous figures to repeated reference with more familiar items. In these studies, the set of objects used in a matching game can elicit basic level terms (such as *candle* or *shoe*) when objects span different categories, or more specific terms (such as *small candle* or *loafer*) when objects appear with category competitors (e.g., Barr & Keysar, 2002; Brennan & Clark, 1996; Keysar, Barr, Balin, & Brauner, 2000).

An overarching issue in such research has been when and how a target addressee's perspective is incorporated during utterance planning, and whether addressees likewise interpret utterances against common ground. Needless to say, there is ample evidence supporting accounts that favor an early single-stage model in which an addressee's perspective is taken into account immediately and referential terms are negotiated (e.g., Brennan, Galati, & Kuhlen, 2010), as well as those that propose a dual-process model in which primary egocentrism is supplemented by secondary perspective adjustment (e.g., Barr & Keysar, 2006). As reviewed in the previously cited papers (and in Schober & Brennan, 2003), sometimes interlocutors fail to take a partner's perspective into account (Bard *et al.*, 2000; Horton & Keysar, 1996), and sometimes what appears to be an addressee influence might be a coincidence of shared perspective (Brown & Dell, 1987; Kesyar *et al.*, 2000).

Regardless of when and how an addressee influences message formulation, this process can evoke increased similarity between interlocutors in a number of dimensions, sometimes referred to as communication accommodation (e.g., Giles, Coupland, & Coupland, 1991; Shepard, Giles, & Le Poire, 2001), or as alignment between interlocutors (e.g., Pickering & Garrod, 2004, 2007, 2013). In general, Giles' communication accommodation perspective focuses on social factors driving both convergence and divergence of talkers, while Pickering and Garrod's alignment perspective focuses on internal cognitive mechanisms that promote parity among interlocutors across all facets of language.

## Communication accommodation

Studies of interacting talkers have found fairly consistent patterns of linguistic change over the course of conversational interaction, and such changes are variously termed *coordination* (Branigan, Pickering & Cleland, 2000; Brennan & Clark, 1996; Clark, 1996), *entrainment* (Arantes & Barbosa, 2010; Brennan & Clark, 1996; Levitan & Hirschberg, 2011; McGarva & Warner, 2003), *alignment* (Branigan, Pickering, McLean, & Cleland, 2007; Pickering & Garrod, 2004, 2007, 2013), or *accommodation* (Giles *et al.*, 1991; Shepard *et al.*, 2001). Most of these proposals examine increasing similarity (convergence) of diverse aspects of interlocutor's speech, from schematic (Garrod & Doherty, 1994), to syntactic (Branigan *et al.*, 2000; Branigan *et al.*, 2007), to lexical/semantic levels (Brennan & Clark, 1996; Fusaroli, Bahrami, Olsen, Roepstorff, Rees, Frith, & Tylén, 2012; Krauss & Weinheimer, 1964; Nenkova, Gravano, & Hirschberg, 2008; Neiderhoffer & Pennebaker, 2002; Wilkes-Gibbs & Clark, 1992). Research on convergence has also included measures of acoustic attributes such as perceived accentedness, sub-vocal spectral covariation, and voice amplitude (e.g., Giles, 1973; Gregory & Webster 1999; Heldner, Edlund, & Hirschberg, 2010; Levitan & Hirschberg, 2011; Natale, 1975). Convergence in such parameters appears to be influenced by social factors that are local to communication exchanges, such as interlocutors' relative dominance or perceived prestige (Gregory, Dagan, & Webster, 1997; Gregory & Webster, 1996).

Howard Giles's *Communication Accommodation Theory* (CAT) also acknowledges the opposite pattern, accent divergence, under some circumstances (Giles *et al.*, 1991; Shepard *et al.*, 2001). Although it is tempting to attribute convergence to an automatic imitative function that increases intelligibility for the parties involved (e.g., Pickering & Garrod, 2004), divergence often does not preclude intelligibility, but serves a communicative purpose for a diverging party (Bilous & Krauss, 1988; Bourhis & Giles, 1977; Labov, 1974). It is arguable that convergence likewise serves a communicative purpose beyond intelligibility. One reason proposed for accommodation is the similarity attraction hypothesis, which claims that individuals try to be more similar to those to whom they are attracted (Byrne, 1971). Accordingly, convergence arises from a need to gain approval from an interacting partner (Street, 1982) and/or from a desire to ensure smooth conversational interaction (Gallois, Giles, Jones, Cargile, & Ota, 1995). Divergence is often interpreted as a means to accentuate individual/cultural differences or to display disdain (Bourhis & Giles, 1977; Shepard *et al.*, 2001).

According to CAT, talkers also converge or diverge along different speech dimensions as a function of their relative status or dominance in an interaction (Giles *et al.*, 1991; Jones, Gallois, Callan, & Barker, 1999; Shepard *et al.*, 2001), which is compatible with the similarity attraction hypothesis. Typically, a talker in a less dominant role will converge toward a more dominant partner's speaking style (Giles, 1973). In contrast, a talker's speech might diverge from a conversational partner's to accentuate distinctiveness, regardless of dominance (Bourhis & Giles, 1977). Finally, talkers have been found to converge on some parameters at the same time that they diverge on others (Bilous & Krauss, 1988).

In an early empirical study by Giles (1973), a group of Bristol men were interviewed by two different interviewers—one spoke with a prestigious Received Pronunciation (RP) accent, and the other was a Bristol-accented interviewer. Excerpts from the interviewees' recordings were played to separate Bristol listeners, who rated how Bristol-accented the speech sounded. When talkers interacted with the RP-accented talker, their speech was rated as less Bristol-accented than when interacting with the Bristol-accented talker. Giles interpreted this pattern as accent convergence toward the RP interviewer (i.e., upward convergence to a higher-status individual).

A second study with Welsh men found accent divergence (Bourhis & Giles, 1977). In this case, Welsh-accented talkers answered pre-recorded interview questions spoken by a single RP-accented talker. During a break in the interview, the talkers overheard the RP-accented talker make some disparaging remarks about the Welsh language (ending with, "the future of Welsh appears pretty dismal," p. 125). Then, the talkers answered a second set of questions recorded from the same RP-accented talker, and excerpts were presented to separate listeners, this time comparing ratings from the pre-insult phase to those for speech produced after the insult. In this case, talkers increased their Welsh-accentedness in the post-insult phase, displaying their Welsh status by diverging from the RP-accented talker. In one case, a Welsh speaker refused to answer questions in the post-insult phase, and responded to every question by conjugating Welsh verbs.

These early studies demonstrated both convergence and divergence in code-switching behavior that formed the basis for a fruitful line of research on social and situational modulators of communication accommodation (summarized in Giles *et al.*, 1991; Shepard *et al.*, 2001). While most studies have focused on a single speech attribute (such as speech rate or accentedness), a crucial study by Bilous and Krauss (1988) on the influence of talker sex on accommodation demonstrated that the landscape of accommodation is extremely complex. Their study assessed the so-called male dominance hypothesis with regard to convergence by comparing base-line measures collected when talkers interacted in same-sex pairs to those when they interacted in mixed-sex pairs (in a total of 60 talkers, half male). Accordingly, when interacting in mixed-sex pairs, females should converge to male speaking patterns, while males should not change their patterns when interacting with females.

Comparing across multiple measures in same- and mixed-sex pairings, Bilous and Krauss (1988) found that convergence in mixed-sex pairs was neither consistent across measures, nor were patterns of convergence explained by the male dominance hypothesis. Men and women both converged in average utterance length and frequency of short and long pauses. Women converged to men in total number of words and in frequency of interruptions, and diverged from men in frequency of back-channels and frequency of interruptions. Men converged to women in frequency of back-channels and frequency of laughter and did not diverge in any measures.

The results reported by Bilous and Krauss (1988) provided an early indication that convergence is not an all-or-none phenomenon. Talkers may converge on some attributes at the same time that they diverge on others. In this case, they measured structural attributes of conversational interaction that related to conversational dominance (holding the floor, managing turn-taking, etc.). These findings are important because they point to a major difficulty in measuring convergence. That is, no single measure can provide a comprehensive assessment of convergence in social interaction. This point will be elaborated further in a discussion of measures of phonetic convergence.

## Interactive alignment

It is clear that aspects of a social/cultural setting and relationship between interlocutors will influence the form and direction of communication accommodation. However, much of the research within the accommodation framework is mute regarding internal cognitive mechanisms that support convergence and divergence during speech production, which have only recently been taken up in the field of psycholinguistics. All such accounts of the phenomenon rest on an assumption of parity of representation between talkers and listeners. In order for a listener to converge to a talker, they must create a sufficiently detailed representation of the talker's speech. This assumption plays a central role in Pickering and Garrod's interactive alignment framework for dialogue (Pickering & Garrod, 2004, 2007, 2013).

In their mechanistic approach, Pickering and Garrod (2004) proposed a model of language use in dialogue based on a simple idea. That is, automatic priming of shared representations leads to alignment at all levels of language—semantic, syntactic, and phonological. Moreover, alignment at one level promotes alignment at other levels. On those occasions when the default automatic priming mechanism fails to yield schematic alignment (e.g., during a misunderstanding), a second more deliberate mechanism brings interlocutors into alignment. The proposed model supports inclusion of an automatic priming mechanism by citing evidence for between-talker alignment at semantic, syntactic, and phonological levels.

In their most recent paper, Pickering and Garrod (2013) draw out a critical component of the model—that language production and language comprehension processes are tightly interwoven within talkers due to processes entailed in self-monitoring of speech production. In particular, they extend concepts from theories of more general action production and perception to language. To do so, they rely heavily on the notion of efference copy, which developed in classic approaches to visual perception (in particular for accommodating the impact of eye movements on motion perception). An efference copy is a secondary signal that is generated along with a motor command (an efference). A separate monitoring system uses an efference copy signal to generate an expected sensory outcome (a forward model) that can be compared to an organism's actual sensory signal resulting from an action. Accordingly, self-monitoring during language production involves a parallel forward modeling component that uses efference copies of speech motor commands to generate simulated perceptual consequences of production. These same forward modeling processes used in speech production also drive active simulation during speech perception, which then leads to covert and sometimes overt imitation. Therefore, when a listener hears an utterance, comprehension relies on the same processes as production, leading to convergent production in a very straightforward manner. A more elaborate account of this model linking language production and perception appears in this volume (Gambi & Pickering, this volume; see also Gambi & Pickering, 2013).

Evidence in support of semantic/schematic alignment is abundant, and is to be expected, because communication could hardly be successful if interlocutors failed to use similar terms. Much of the previously discussed research on development of referential terms supports a proposal of semantic alignment. In situations investigated so far, interlocutors aligned on terms to refer to relevant properties of a task. In other less constrained situations, however, it is acknowledged that some degree of disparity in representations can be tolerated. When disparity is too great, interlocutors then must negotiate to restore mutual understanding (Clark & Wilkes-Gibbs, 1986; Wilkes-Gibbs & Clark, 1992).

A compelling study by Garrod and Doherty (1994) extended the concept of schematic alignment within an individual dyad to alignment across members of a closed community of talkers who each interacted with every other member of the community. Participants in their study played a computerized maze game in which each member of a pair moved an icon from a starting cell to an ending cell in their maze. The structure of the game required conversational collaboration, in that each interlocutor had to guide their partner to particular locations in a maze in order to complete their own movements. During task completion, interlocutors used multiple descriptions schemes for mazes that became aligned over the course of interaction. For example, one talker might use a scheme that involved referring to maze cells according to a matrix notation, while another might use a scheme involving path-like descriptions (there were four basic description schemes). Eventually, both talkers would use the same scheme on adjacent turns.

Each individual played the maze game nine times, with a limit of 10 minutes/game. A total of 20 talkers participated, either as individual pairs or as a closed community (N = 10 each). The individual pairs played the game nine times with the same partner. The closed community group played the game nine times as well, but each game was played with a different partner so that by the end, each person had played the game only once with every other person in the group. Alignment was measured as between-talker consistency in using the same description scheme from one turn to the next.

Individual pairs rapidly converged on consistent schemes in early games, and persisted in the same high level of consistency throughout the interaction. The community group started out with lower levels of consistency than individual pairs, but by the middle and late games (games 4-9), their levels of consistency were actually higher than that of individual pairs, despite the fact that they changed partners on every game. A second experiment ruled out the possibility that changing partners alone caused the increased consistency—what mattered was that the talkers formed a closed community, with each member of the community carrying the most frequent scheme from one interaction to the next. Thus, schematic alignment occurs within individual pair-wise interactions, but can also become more consistent across a community of mutually interacting talkers.

Semantic alignment is central to notions of communicative efficacy. Interlocutors do not reach mutual understanding without some degree of parity or shared meaning, which often results in shared lexical forms. However, syntax is demonstrably separate from semantics, therefore, parity in communication does not entail parity in syntactic form. With respect to a situation model, it is irrelevant whether one chooses to say, *The pirate gave the banana to the clown*, or, *The pirate gave the clown the banana*. Both structures yield the same understanding of the expressed situation, despite their syntactic differences. Although communicative efficacy does not necessarily drive syntactic parity, early corpus studies found that the likelihood of producing a particular syntactic form was greater if that same variant had been used previously in a discourse (Estival, 1985; Levelt & Kelter, 1982; Weiner & Labov, 1983). These findings, which came to be known as syntactic repetition/priming, led to a host of laboratory studies of the phenomenon, beginning with Bock (1986) and well summarized by Hartsuiker, Bernolet, Schoonbaert, and Vanderelst (2008) and Gries (2005). Overall, syntactic repetition has been found in situations without repetition of lexical items, but the findings are stronger with concurrent repetition of lexical forms and in dialogue tasks (Hartsuiker *et al.*, 2008), and are generally similar to those reported in uncontrolled corpus studies (Gries, 2005).

Two studies by Branigan and colleagues examined syntactic repetition in controlled dialogue, providing rigorous support for the proposal that interlocutors might align on syntactic forms (Branigan *et al.*, 2000; Branigan *et al.*, 2007). In these studies, a talker interacted with a confederate in an alternating picture description and matching game—one talker described a picture to be matched by their partner, and then they switched roles on the next picture. The pictures comprised line drawings of an agent holding an object next to a recipient of the object, and the describers' cards were labeled with a verb that was to be used in the description (there were six different verbs, and each verb appeared on two cards).

These materials elicit sentences with a syntactic frame that could alternate between a Prepositional Object (PO: the agent verbing the x *to* the y) or a Double Object (DO: the agent verbing the y the x). Confederates were provided with scripted sentences that controlled whether they used the PO or DO form on each trial. Of interest was whether the participants would use the same form or the alternate form on their subsequent turn describing a new picture. Note that these conversations were heavily constrained such that participants could only repeat the same descriptions if clarification was needed.

Branigan *et al.* (2000) found evidence that talkers (N = 24) were more likely to use the same syntactic frame as on a previous trial, and that the effect was stronger when the verb was also the same. Participants were 55% more likely to use the same syntactic form as the confederates when the verbs were the same, and 26% more likely when the verbs differed. Crucially, the items depicted were not the same, so the effect was not due to overlap in other forms of sentence content. For example, if the confederate described *the cowboy handing the banana to the burglar*, the participant would then describe *the pirate handing the cake to the sailor* (as opposed to *the pirate handing the sailor the cake*).

Branigan *et al.* (2007) extended these findings to investigate whether the relationship between the confederate and the participant mattered. In this case, there were three talkers, and the confederate could have been addressing the participant directly or another talker prior to the participant's description. When a participant had been the addressee of the prior description, they were 32% more likely to use the same form in their subsequent description. When they had merely been a side-participant to the confederate's utterance (they were told to check the other talker's descriptions), they were only 12% more likely to use the same form. These results lend some support to the notion that interlocutors often converge on the same syntactic forms during conversational interaction, but note that talkers often used the other form, indicating that activation of particular syntactic procedures does not fully determine the syntactic form of an utterance.

At the phonological level, a growing body of research on phonetic convergence both supports and challenges a mechanistic approach based on automatic priming. On one hand, studies have found that interacting talkers become more similar in phonetic repertoire, whether talkers come from the same or different dialect regions. On the other, observed patterns of phonetic convergence are not readily accommodated by a framework that relies so heavily on an automatic priming mechanism. To date, most of the psycholinguistic research on phonetic convergence has used non-interactive speech shadowing tasks (e.g., Goldinger, 1998), passive exposure tasks (e.g., Nielsen, 2011), or examined the impact of accumulated exposure (e.g., Pardo, Gibbons, Suppes, & Krauss, 2012; Sancier & Fowler, 1997; see review in Pardo, Jordan, Mallari, Scanlon, & Lewandowski, 2013). In studies of phonetic convergence during conversational interaction, talkers have been found to converge, but effects are subtle and uneven across multiple measures. Most importantly, as in the Branigan *et al*. (2007) study of syntactic alignment, phonetic convergence is dramatically influenced by a talker's role in a conversational setting.

## Phonetic convergence during conversational interaction

Inspired by a finding that talkers "imitated" model speech prompts in a non-interactive speech shadowing task (Goldinger, 1998), Pardo (2000, 2006) adapted the paradigm to examine phonetic convergence during conversational interaction. In addition to establishing phonetic convergence during conversational interaction, Pardo (2006) found that a talker's role in a conversation and the sex of the pair of talkers both influenced degree of phonetic convergence. Subsequent studies explored the impact of role by manipulating a talker's intention to imitate (Pardo, Cajori Jay, & Krauss, 2010) and role stability (Pardo, Cajori Jay, Hoshino, Hasbun, Sowemimo-Coker, & Krauss, 2013). In all studies, it was necessary to introduce a constrained conversational task that would guarantee between-talker repetitions of the same lexical items.

To obtain samples of naturalistic conversational speech with appropriate repetitions, the paradigm employed a modified version of the Map Task, a cooperative conversational task that was developed by the Human Communication Research Center at the University of Edinburgh, Scotland (Anderson *et al*., 1991). The Map Task comprises paired maps with labeled iconic landmarks (e.g., *walled city*, *wheat field*, *green bay*). One map in each pair, designated the Giver's map, has a path drawn from a starting point, around various labeled landmarks, to a finishing point. The corresponding map, designated the Receiver's map, has only a starting point and various labeled landmarks. The goal of the task is for a pair of talkers to communicate effectively enough that the Receiver can duplicate the path that is drawn on the Giver's map without seeing each others' maps. The Map Task is particularly useful for studying phonetic convergence during conversational interaction because talkers naturally repeat the landmark labels, and changes in an individual's phonetic repertoire can be assessed within the same lexical items by collecting recordings of the landmark label phrases before and after conversational interaction.

Measures of phonetic convergence included both acoustic analyses of speech samples and an AXB perceptual similarity test that compares a talker's pre-task and task utterances (A/B) to their partner's utterances (X) (adapted from Goldinger, 1998). This use of an AXB perceptual similarity task ensures that the apparent similarity between talkers was not coincidental—the measure reflects change in a talker's phonetic repertoire that makes them sound more similar to a model than they were prior to exposure. Measuring phonetic convergence in a perceptual task is preferable to phonetic transcriptions due to difficulties entailed in obtaining reliable transcriptions, and to ensure that the measure reflects changes that are available to ordinary listeners. Moreover, perceptual assessment provides a holistic appraisal of similarity that integrates over all acoustic-phonetic dimensions, avoiding potential pitfalls involved in committing to a single acoustic measure (Pardo, 2013; Pardo, Jordan *et al.*, 2013; Pardo & Remez, 2006). It is likely that different pairs of talkers might converge on distinct acoustic-phonetic attributes, and perceptual measures of similarity reflect patterns that would be missed when measuring acoustic attributes alone. Finally, identification of converging acoustic attributes alone leaves open the question of whether such attributes are perceptually salient and available for use during conversational interaction.

Across multiple studies of conversational interaction, talkers were found to converge in phonetic form. That is, a talker's utterances of landmark label phrases, such as *green bay* or *diamond mine*, sounded more similar to their partner's utterances during or after conversational interaction than before the talkers met. As in studies of syntactic convergence, measures of phonetic convergence indicate that the change is subtle and variable. For example, listeners selected the conversational task items as more similar to the partner's items on 65% of trials in one study (Pardo, 2006), but on only 53% of trials in another study (Pardo *et al.*, 2010). Both findings were significantly greater than chance responding (50%), but the overall range of values reported across multiple studies indicate that the phenomenon is often subtle and highly variable.

Overall, there is a great deal of variability in phonetic convergence across individual pairings that is not readily explained in these studies and merits further research. Across studies of phonetic convergence, measures for individual talkers ranged from 33% to 83% detected convergence in AXB tasks. However, in contrast with earlier findings from the communication accommodation literature, talkers in these studies did not converge in speaking rate, presumably due to an influence of role discrepancy in the Map Task (Pardo *et al.*, 2010; Pardo *et al.*, 2013). That is, Givers tended to speak faster than Receivers, maintaining a significant difference in speaking rate that Receivers neither matched nor tracked in their own speaking rates. Furthermore, item and vowel analyses have found inconsistent patterns of vowel convergence across pairs—individual pairs converged on unique acoustic-phonetic attributes that were apparent to listeners who made a global perceptual appraisal (see Pardo, Jordan *et al.*, 2013).

Taken together, these findings illustrate the complexity of phonetic convergence in conversational interaction. A glance at the literature on phonetic convergence in more constrained non-interactive settings is no less complex (see reviews in Pardo, Jordan *et al.*, 2013; Pardo, Urmanche, Wilman, & Wiener, 2017). Although it is clear that some form of representational similarity must underlie a talker's ability to converge in phonetic form to an interlocutor, it is not clear how this similarity arises and how it interfaces with a talker's appraisal of their social setting to determine their degree of convergence. Automatic priming mechanisms are not sufficient to explain simultaneous findings of global phonetic convergence and inconsistent patterns in individual acoustic attributes such as articulation rates and vowel formants. Such findings can be accommodated within the framework of entrainment, which explicitly incorporates both patterns of convergence and divergence among interacting dynamical systems.

## Entrainment

The principles of entrainment, the *maintenance tendency, superimposition*, and the *magnet effect*, were initially identified in von Holst's (1937/1973 in Gallistel, 1980) early research on endogenous rhythmicity in behavioral organization. Examining fish fin oscillations, von Holst discovered endogenous neural oscillators that likely serve as basic building blocks of complex behaviors through superimposition of coordinated patterns (Gallistel, 1981; Turvey, 1990). The *maintenance tendency* describes the observation that each oscillator prefers to operate according to its own intrinsic dynamics (frequency and amplitude). From a basic set of simple oscillators, more complex motions can be assembled through *superimposition* of oscillator dynamics. Finally, the *magnet effect* occurs when separate oscillators become coupled, and a more dominant or stable oscillator pulls a less dominant oscillator into synchrony with its rhythm.

With rigid coupling of systems with identical intrinsic dynamics (as in mechanically coupled physically identical pendulums), oscillators pull into absolute coordination or entrainment, a rare phenomenon in which both the phase relationship and the frequency of oscillation match (Schmidt & Turvey, 1989). More typical scenarios exhibit a struggle between the maintenance tendency of each oscillator's intrinsic dynamics and the magnet effect, resulting in relative coordination, in which periods of synchrony alternate with periods of asynchrony. Even in cases of apparent absolute coordination, fine-grained movement dynamics often reveal a residual latent struggle between the maintenance tendency and the magnet effect.

Although von Holst's work described complex motions of oscillating fish fins, these principles have survived scale transformation to various forms of human interaction (for example, limb movements in walking, see Turvey, 1990). Interpersonal entrainment typically exhibits only relative coordination because it both lacks rigid coupling and individuals' intrinsic dynamics are never identical. For example, when individuals of different sizes walk together, their gaits might reach moments of synchrony that persist or break down, depending on their degree of coupling and intrinsic gait differences. If individuals are closer in size, they will be more likely to maintain a frequency-locked entrainment pattern because their intrinsic dynamics are more similar. If individuals physically connect by holding hands, increased coupling might promote greater entrainment as well. In cases of relative coordination, frequencies of individual oscillators approach that of a dominant oscillator or an intermediate frequency, but any frequency match is temporary and one might observe persistent differences in phase relationships. Despite the pull to entrain to a coupled oscillator, the manifest pattern exhibits a latent influence of the original intrinsic dynamics, presumably because an external oscillator's pattern is superimposed onto an internal oscillator's pattern rather than supplanting it. Thus, an individual will never completely match the dynamics of another, but some aspects of their behavior will come into relative coordination.

The concepts of coupling, the magnet effect, and the maintenance tendency provide a ready model of the integration of internal and external forces in conversational interaction. Beek, Turvey, and Schmidt (1992) proposed that external information acts as an embedded forcing function on internal dynamics, inducing changes in the overall pattern of activity that push the activity to different values in its intrinsic range. In this way, perception can influence production through perceived external dynamics that are embedded within internal production systems, subject to degree of coupling and the maintenance tendency. Research on self-regulation of speech production, in particular the Lombard sign (Lane & Tranel, 1971) and perceptual-productive adaptation of vowel formants, speaking fundamental frequency, and consonant spectra (Houde & Jordan, 2002; Jones & Munhall, 2000; Jones & Munhall, 2003) shows that talkers can incorporate auditory feedback of their own productions to adjust subtle aspects of speech at short latencies. Perceiving the speech of other talkers might involve the same system as self-monitoring, as proposed by Gambi and Pickering (this volume). Their approach incorporates simulation through forward modeling, which is compatible with principles of entrainment (see also Fusaroli, Rczaszek-Leonardi, & Tylén, 2014).

If perception of another talker's speech yields detailed phonetic forms, such forms could influence subsequent production under circumstances, such as demands of conversational interaction, that promote coupling between talkers. At the same time, initial similarity of two talkers (e.g., whether talkers come from similar or distinct dialect backgrounds), and relative rigidity of each talker's internal dynamics will also influence degree of phonetic convergence. There is some indication that phonetic convergence during conversational interaction might be stronger for talkers from the same as opposed to different dialect regions (Kim, Horton, & Bradlow, 2011). However, there are limits to phonetic convergence which are not accounted for by similarity alone. In a study by Vallabha and Tuller (2004), talkers failed to imitate their own vowel sounds, exhibiting systematic biases in their productions that could not be accounted for by perceptual or productive processes alone.

In all studies of phonetic convergence, there were large individual differences in degree of convergence, as well as modulations by talker role and other situational factors. According to Giles' communication accommodation framework, social dominance and attractiveness influence the direction of the magnet effect pulling talkers together, but dominance is not always a straightforward consequence of talker role, and is potentially idiosyncratic to different pairs.

Finally, social dominance is irrelevant for entrainment if systems are not coupled. With looser coupling, there is likely to be less convergence, as is generally the case with informationally coupled systems, such as interacting talkers (see Schmidt & Turvey, 1989).

The notion of coupling in interpersonal coordination maps onto attention and perception, and accounts of coordination in language use will ultimately wrestle with demands of these systems. An intriguing study of coordination in rocking chairs provides direct evidence of a role for attention in rhythmic entrainment (Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007). In that study, pairs of individuals sat side-by-side in rocking chairs that had the same or different natural tempos. They were either told to rock together or at their own tempo, and they either focused on a dot in front of them or on the armrest of their partner. They hypothesized that a reliance on peripheral vision would disrupt degree of coordination of their movements, however, this was only the case in the unintentional condition. When participants intended to coordinate their rocking movements, it did not matter whether they used peripheral or focal vision to couple their movements with their partner. Thus, an intention to coordinate influenced whether perceptual information that was available peripherally was ultimately used to promote interpersonal coordination. With focal attention, on the other hand, coordination emerged regardless of intention (see also Schmidt, Richardson, Arsenault, & Galantucci, 2007).

Although coupling is a critical component for entrainment, an intention to imitate is not sufficient to induce phonetic convergence in conversational interaction. Pardo *et al*. (2010) explicitly instructed one member of each pair of 12 talkers to try to imitate their partner's speech during completion of the Map Task. As in previous studies, the role of the talker who received the instruction influenced degree of phonetic convergence. Phonetic convergence was only reliable in those pairs in which Receivers had been instructed to imitate. Instructing Givers to imitate disrupted previously observed patterns of phonetic convergence. Thus, a more elaborate account of the nature of informational coupling and its relationship to intention, attention, and perception is warranted (see Schmidt, Fitzpatrick, Caron, & Mergeche, 2011; Schmidt, Morr, Fitzpatrick, & Richardson, 2012).

Conversational interaction shares many properties with other forms of interpersonal entrainment. However, unlike studies of activities such as wrist pendulum swinging, finger tapping, or chair rocking, the intrinsic dynamics of many attributes of spoken language are extremely complex and still relatively poorly understood. At a first pass, one would at least expect to find rhythmic entrainment in speech production (Wilson & Wilson, 2005), but findings of rate entrainment in speech production during conversational interaction have been inconsistent. For example, early studies in the communication accommodation literature reported moderate interlocutor correlation in speech rate (Putman & Street, 1984; Street, 1982), but more recent investigations have failed to find consistent speech rate convergence, or found significant differences in speech rate (Pardo *et al*., 2010; Pardo, Cajori Jay *et al*., 2013). In language use, as opposed to other forms of interpersonal coordination, the situation is complicated by the fact that the surface form of most attributes is influenced by multiple linguistic and nonlinguistic (i.e., social/situational/cultural) goals simultaneously. Discovering how these multiple aims come together in acts of spoken communication across all settings is a worthwhile pursuit (see Krauss & Pardo, 2006).

## Conclusion

Conversational interaction poses a challenge for psycholinguistics, both methodologically and conceptually. One of the aims of this chapter is to demonstrate that the work entailed in overcoming some of the methodological challenges can yield useful and important contributions to an understanding of speech production and perception. As Pickering and Garrod (2004) point out, language use can be thought of as a continuum from monologue to dialogue, with some occasions of monologue simulating important aspects of dialogue, and some occasions of dialogue as little more than serial monologue. Taken together, research on conversational interaction demonstrates that parity is not the only aim—sometimes talkers express themselves in ways that lead to disparity in multiple attributes of spoken communication. Although listeners expect intelligible messages, there is plenty of latitude for disparity that can enhance mutual understanding of more than words can say.

## Author Note

Partial support for preparation of this chapter was provided by a grant from the National Science Foundation PAC program #BCS-1229033. The author thanks Helen Cairns, Eva Fernández, Chiara Gambi, Martin Pickering, and Robert Remez for their help in improving the narrative and supporting this work.

## REFERENCES

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S.,… Weinert, R. (1991). The H.C.R.C. Map Task Corpus. *Language & Speech*, 34, 351–366.

Arantes, P., & Barbosa, P. A. (2010). Production-perception entrainment in speech rhythm. *Proceedings from Speech Prosody 2010*, Chicago.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the Intelligibility of Referring Expressions in Dialogue. *Journal of Memory and Language*, 42(1), 1–22. doi: 10.1006/jmla.1999.2667

Barr, D. J., & Keysar, B. (2002). Anchoring Comprehension in Linguistic Precedents. *Journal of Memory and Language*, 46(2), 391–418. doi: 10.1006/jmla.2001.2815

Barr, D. J., & Keysar, B. (2006). Perspective taking and the coordination of meaning in language use. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of*

*Psycholinguistics*, *Second Edition* (pp. 901–938). Burlington, MA: Academic Press.

Beek, P. J., Turvey, M. T., & Schmidt, R. C. (1992). Autonomous and nonautonomous dynamics of coordinated rhythmic movements. *Ecological Psychology*, 4, 65–95.

Bilous, F. R., & Krauss, R. M. (1988). Dominance and accommodation in the conversational behaviours of same-and mixed-gender dyads. *Language and Communication*, 8(3/4), 183–194.

Bock. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, Ethnicity, & Intergroup Relations* (pp. 119–135). London: Academic Press.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75, B13–B25.

Branigan, H. P., Pickering, M. J., McLean, J. F., & Cleland, A. A. (2007). Syntactic alignment and participant role in dialogue. *Cognition*, 104(2), 163–197. doi: 10.1016/j.cognition.2006.05.006

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482–1493.

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation, Vol.* 53 (pp. 301–344). Burlington: Academic Press. doi: 10.1016/S0079-7421(10)53008-1

Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicity mention of instruments. *Cognitive Psychology*, 19, 441–472.

Byrne, D. (1971). *The Attraction Paradigm*. New York, NY: Academic Press.

Clark, H. H. (1996). *Using Language*. Cambridge, UK: Cambridge University Press.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1–39.

Estival, D. (1985). Syntactic priming of the passive in English. *Text* 5(1), 7–22.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylen, K. (2012). Coming to Terms: Quantifying the Benefits of Linguistic Coordination. *Psychological Science*, 23, 931–939. doi: 10.1177/0956797612436816

Fusaroli, R., Raczaszek-Leonardi, J., & Tylen, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, 32, 147–157.

Fussell, S. R., & Krauss, R. M. (1989). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology*, 25(3), 203–219.

Gallistel, C. R. (1980). Von Holst's coupled oscillators. In *The Organization of Action: A New Synthesis* (pp. 69–110). Hillsdale, NJ: Lawrence Erlbaum.

Gallistel, C. R. (1981). Précis of Gallistel's *The Organization of action: A new synthesis*. *Behavioral & Brain Sciences*, 4, 609–650.

Gallois, C., Giles, H., Jones, E., Cargiles, A. C., & Ota, H. (1995). Accommodating intercultural encounters: Elaboration and extensions. In R. Wiseman (Ed.), *Intercultural Communication Theory* (pp. 115–147). Newbruy Park, CA: Sage.

Gambi, C. & Pickering, M. J. (2013). Prediction and imitation in speech. *Frontiers in Psychology*, 4, 340. doi: 10.3389/fpsyg.2013.00340

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, 53(3), 181–215.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, 15, 87–109.

Giles, H., Coupland, J., & Coupland, N. (1991). Chapter 1: Accommodation theory: Communication, context, and consequence. In *Contexts of Accommodation: Developments in Applied Sociolinguistics* (pp. 1–68). Cambridge: Cambridge University Press.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279.

Gregory, S. W., Dagan, K., & Webster, S. (1997). Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior*, 21(1), 23–43.

Gregory, S. W., & Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology*, 70(6), 1231–1240.

Gries, S. T. (2005). Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research*, 34(4), 365–399. doi: 10.1007/s10936-005-6139-3

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory & Language*, 58, 214–238. doi: 10.1016/j.jml.2007.07.003

Heldner, M., Edlund, J., & Hirschberg, J. (2010). Pitch similarity in the vicinity of backchannels. *INTERSPEECH-2010*, 3054–3057.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59, 91–117.

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of speech, language, and hearing research*, 45(2), 295.

Jones, E., Gallois, C., Callan, V., & Barker, M. (1999). Strategies of accommodation: Development of a coding system for conversational interaction. *Journal of Language & Social Psychology*, 18, 123.

Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America*, 108, 1246–1251.

Jones, J. A., & Munhall, K. G. (2003). Learning to produce speech with an altered vocal tract: The role of auditory feedback. *Journal of the Acoustical Society of America*, 113(1), 532–543. doi: 10.1121/1.1529670

Keysar, B., Barr, D. J., Balin, D. J., & Brauner, J. S. (2000). Taking perspective in conversation: The use of mutual knowledge for error correction. *Psychological Science*, 11, 32–38.

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, 2(1), 125–156.

Krauss, R. M. (1987). The role of the listener: Addressee influences on message formulation. *Journal of Language and Social Psychology*, 6(2), 81–98.

Krauss, R. M., & Bricker, P. D. (1967). Effects of transmission delay and access delay on the efficiency of verbal communication. *Journal of the Acoustical Society of America*, 41, 286–292.

Krauss, R. M., Bricker, P. D., Garlock, C. M., & McMahon, L. E. (1977). The role of audible and visible back-channel responses in interpersonal communication. *Journal of Personality and Social Psychology*, 9, 295–300.

Krauss, R. M., & Chiu, C. Y. (1998). Language and social behavior. In D. Gilbert, S. Fiske, & G. Lindsey (Eds.), *The Handbook of Social Psychology, 4th Edition* (pp. 41–88). Boston: McGraw-Hill.

Krauss, R. M., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. In E. T. Higgins & A. Kruglanski (Eds.), *Social Psychology: Handbook of Basic Principles* (pp. 655–701). New York: Guilford Press.

Krauss, R. M., & Pardo, J. S. (2006). Speaker perception and social behavior: Bridging social psychology and speech science. In P. A. M. van Lange (Ed.), *Bridging Social Psychology: Benefits of Transdisciplinary Approaches* (pp. 273–278). Hillsdale, NJ: Lawrence Erlbaum.

Krauss, R. M., Vivekananathan, P. S., & Weinheimer, S. (1968). 'Inner speech' and 'external speech': Characteristics and communication effectiveness of socially and non-socially encoded messages. *Journal of Personality and Social Psychology*, 35, 523–529.

Krauss, R. M., & Weinheimer, S. (1964). Changes in the length of reference phrases as a function of social interaction: A preliminary study. *Psychonomic Science*, 1, 113–114.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in social interaction. *Journal of Personality and Social Psychology*, 4, 342–346.

Labov, W. (1974). Linguistic change as a form of communication. In A. Silverstein (Ed.), *Human Communication: Theoretical Explorations* (pp. 221–256). Hillsdale, NJ: Lawrence Erlbaum.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech & Hearing Research*, 14, 677–709.

Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology* 14, 78–106.

Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. *INTERSPEECH-2011*, Florence, Italy, 3081–3084.

McGarva, A. R., & Warner, R. M. (2003). Attraction and social coordination: Mutual entrainment of vocal activity rhythms. *Journal of Psycholinguistic Research*, 32(3), 335–354.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, 32(5), 790–804.

Nenkova, A., Gravano, A., & Hirschberg, J. (2008). *High frequency word entrainment in spoken dialogue*. Proceedings from Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers.

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, 21(4), 337–360. doi: 10.1177/026192702237953

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39(2), 132–142. doi: 10.1016/j.wocn.2010.12.007

Pardo, J. S. (2000). *Imitation and coordination in spoken communication*. Doctoral dissertation, Yale University.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119(4), 2382–2393. doi: 10.1121/1.2178720

Pardo, J. S. (2013). Measuring phonetic convergence in speech production. *Frontiers in Psychology: Cognitive Science*, 4, 559.

Pardo, J. S., Cajori Jay, I., Hoshino, R., Hasbun, S. M., Sowemimo-Coker, C., & Krauss, R. M. (2013). The influence of role-switching on phonetic convergence in conversation. *Discourse Processes*, 50, 276–300. doi: 10.1080/0163853X.2013.778168

Pardo, J. S., Cajori Jay, I., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, &*

*Psychophysics*, 72(8), 2254–2264. doi: 10.3758/APP.72.8.2254

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, 40(1), 190–197. doi: 10.1016/j.wocn.2011.10.001

Pardo, J. S., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. (2013). Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. *Journal of Memory & Language*, 69, 183–195. doi: 10.1016/j.jml.2013.06.002

Pardo, J. S., & Remez, R. E. (2006). The perception of speech. In M. Traxler & M. Gernsbacher (Eds.), *The handbook of psycholinguistics* (2nd edition, pp. 201–248). New York: Academic Press.

Pardo, J. S., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics*, 79, 637–659. doi: 10.3758/s13414-016-1226-0

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02), 169–190.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11(3), 105–110. doi: 10.1016/j.tics.2006.12.002

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral & Brain Sciences*, 36, 49–64. doi: 10.1017/S0140525X12003238

Putman, W. B., & Street, R. L. (1984). The conception and perception of noncontent speech performance: Implications for speech-accommodation theory. *International Journal of the Sociology of Language*, 46, 97–114.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C. (2007). Rocking together: dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, 26(6), 867–891. doi: 10.1016/j.humov.2007.07.002

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421–436.

Schmidt, R. C., Fitzpatrick, P., Caron, R., & Mergeche, J. (2011). Understanding social motor coordination. *Human Movement Science*, 30, 834–845. doi: 10.1016/j.humov.2010.05.014

Schmidt, R. C., Morr, S., Fitzpatrick, P., & Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *Journal of Nonverbal Behavior*, 36, 263–279.

Schmidt, R. C., Richardson, M. J., Arsenault, C., & Galantucci, B. (2007). Visual tracking and entrainment to an environmental rhythm. *Journal of Experimental Psychology: Human Perception & Performance*, 33(4), 860–870. doi: 10.1037/0096-1523.33.4.860

Schmidt, R. C., & Turvey, M. T. (1989). Absolute coordination: An ecological perspective. In S. A. Wallace (Ed.), *Perspectives on the Coordination of Movement* (pp. 123–156).

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47, 1–24.

Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of Discourse Processes* (pp. 123–164). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21(2), 211–232.

Searle, J. R. (1976). A classification of illocutionary acts. *Language in Society*, 5, 1–23.

Shepard, C. A., Giles, H., & Le Poire, B. A. (2001). Communication Accommodation Theory. In W. P. Robinson & H. GIles (Eds.), *The New Handbook of Language & Social Psychology* (pp. 33–56). New York: Wiley.

Street, R. L. (1982). Evaluation of noncontent speech accommodation. *Language & Communication*, 2, 13–31.
Turvey, M. T. (1990). Coordination. *American Psychologist*, 45(8), 938–953.
Vallabha, G. K., & Tuller, B. (2004). Perceptuomotor bias in the imitation of steady-state vowels. *Journal of the Acoustical Society of America*, 116(2), 1184–1197. doi: 10.1121/1.1764832
Weiner, E. J. & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics* 19, 29–58.
Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory & Language*, 31, 183–194.
Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, 12(6), 957–968.

# 7 Models Linking Production and Comprehension

CHIARA GAMBI[1, 2] AND
MARTIN J. PICKERING[2]

[1] Saarland University
[2] University of Edinburgh

## Introduction

Production and comprehension have traditionally been studied in separate subfields of psycholinguistics, despite the fact that many psycholinguistic tasks involve both. For example, in the picture‐word interference paradigm, participants name target pictures while ignoring auditorily presented or superimposed written distractor words (e.g., Schriefers, Meyer, & Levelt, 1990). Therefore, this paradigm measures the effect of comprehension processes on production processes. This disciplinary division notwithstanding, there is general consensus that production and comprehension are linked. However, opinions diverge with regard to the nature of this link. It is not easy to characterise scholarly disagreement on this issue, as the issue is often not discussed explicitly, and even when it is, it is rarely treated as a key element of the model being proposed.

There are, however, some exceptions to this trend. Below, we focus on models that make explicit claims about the links between production and comprehension. It is noteworthy that these models tend to emphasize what is shared between comprehension and production, rather than what is not shared. Not surprisingly, frameworks that emphasize sharing tend to be concerned with explaining language learning, acquisition, and change (see Dell & Chang, 2014; MacDonald, 2013) or dialogue (see Pickering & Garrod, 2013), rather than isolated acts of production or comprehension. In fact, the simple observation that the primary site of language use is dialogue constitutes, in itself, a strong motivation for positing links between production and comprehension (Garrod & Pickering, 2004; see section on *Dialogue*, this chapter). And theories of how linguistic representations are acquired and

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

develop over time must naturally take into account how linguistic input shapes the output of the language system (and *vice versa*; see *Frameworks that posit linked preferences*, this chapter).

Traditionally, a number of arguments are used to claim that a separation between comprehension and production should be maintained. First, dissociations between the comprehension and production abilities of patients with brain lesions appear to support the notion that the neural substrates of production and comprehension are separate. For example, Kim and Thompson (2000) reported that agrammatic patients showed intact comprehension of verbs but were impaired in verb naming. Second, asymmetries exist between the development of language comprehension and language production in infants (see Bates, 1993), and similarly, between the rate of decay of comprehension and production abilities in older adults (e.g., McKay, Abrams, & Pedroza, 1999). These findings indicate that it is not possible to fully equate production and comprehension. A comprehensive review of these and other arguments is beyond the scope of this chapter. Here, we just note that these arguments do not necessarily imply that comprehension and production are completely separate. Rather, dissociations and asymmetries are in principle compatible with some degree of sharing, as long as there are subcomponents of production that are not used in comprehension and *vice versa*.

When discussing the issue of what is shared between production and comprehension, it is useful to bear in mind the distinction between linguistic *representations* and *processes* acting on those representations. Linguistic representations are the components of memory that store information about linguistic units (e.g., phonemes, words, syntactic rules, concepts). Comprehension and production processes are the cognitive operations that can be applied to linguistic representations (e.g., retrieval, spreading activation, inhibition), as well as the operations that map from abstract representations to articulation and from acoustics to abstract representations. Processes are directional: so, for example, the process that retrieves a phonological representation given an activated semantic representation is not the same as the process that retrieves a semantic representation given an activated phonological representation. Below, we first discuss sharing with regard to representations (*Representational parity*), and then we turn to theories that posit common processes (*Linked processes*).

## Representational parity

Most theorists assume that some of the representations that are accessed during production are the same as the representations accessed during comprehension. At the single word level, the most influential theory of lexical access in production assumes conceptual (semantic), lemma (syntactic), and word-form (sound-based) representations, and proposes the network for production and the network for comprehension "coincide from the lemma level upwards" (i.e., concept and lemma nodes are shared between production and comprehension; Levelt, Roelofs, & Meyer, 1999, p. 7). Pickering and Garrod (2004) also assume shared representations (which they term the *parity hypothesis*), but extend the scope of this assumption by positing that representations used in comprehension are the same as representations used in production at all linguistic levels, from the situation model (i.e., a representation of the situation being discussed, including time, space, causal relations, intentionality, and individuals involved; Zwaan & Radvansky, 1998) to phonology and phonetics. Shared phonological representations are also part of the Node Structure Theory of MacKay (1987). Moreover, parity at the sound level is the central assumption of the Motor Theory of speech perception (Liberman & Whalen, 2000; see Galantucci, Fowler, & Turvey, 2006), and of the Episodic Theory of speech perception (Goldinger, 1998).

Parity at the semantic and lexico-syntactic levels has been confirmed by several studies. The strongest evidence comes from findings of immediate effects of comprehension on production. First, silent reading of words semantically and associatively related to the name of a target picture affects naming times for the picture (e.g., Schriefers *et al.*, 1990; Alario, Segui, & Ferrand, 2000). Second, silent reading of sentences such as *A rock star sold an undercover agent some cocaine* (double object, DO) or *A rock start sold some cocaine to an undercover agent* (prepositional object, PO) influences what sentence structure (DO or PO) is used to describe a target scene (depicting an unrelated event, such as a man reading a book to a boy; Bock, Dell, Chang, & Onishi, 2007). And, similarly, the syntactic choices of their interlocutor influence speakers' syntactic choices in dialogue (Branigan, Pickering, & Cleland, 2000; Levelt & Kelter, 1982; see Pickering & Ferreira, 2008 for a review). In addition, interlocutors align their lexical choices (Brennan & Clark, 1996; Garrod & Anderson, 1987).

As well as behavioral evidence, there is growing evidence for lexico-syntactic parity at the neural level. Two fMRI studies showed that the same neural populations are recruited during comprehension and production of sentences (Menenti, Gierhan, Segaert, & Hagoort, 2011; Segaert, Menenti, Weber, Petersson, & Hagoort, 2012). These studies identified brain areas that were activated less while producing (or comprehending) a given sentence structure, when the participant had just processed a sentence with the same structure, compared to a sentence with a different structure. This phenomenon is called repetition suppression, and it is used to localize neural areas that are sensitive to a given property of the stimulus (in this case, structure). Importantly, the areas identified were the same regardless of whether prior processing of the same structure had taken place in production or comprehension.

Regarding parity at the phonological level, it is known that silent reading of, or passive listening to, distractor words phonologically related to the name of a target picture speeds up naming times for the picture (e.g., Schriefers *et al.*, 1990; Damian & Martin, 1999). This is usually taken as evidence that comprehension of distractor words pre-activates phonological representations they share with the target, so that those representations are subsequently easier to access in production. Moreover, Kerzel and Bekkering (2000) found that participants pronounced a printed syllable more quickly while watching a video of a mouth producing the same syllable compared to when the mouth produced a different syllable (see also

Jarick & Jones, 2008). In addition, Galantucci, Fowler, and Goldstein (2009) showed that speakers are faster to produce a syllable (e.g.,/ba/) if they have just listened to the same syllable than to a syllable with a different onset (e.g.,/da/) (see also Fowler, Brown, Sabadini, & Weihing, 2003).

There is also much evidence demonstrating motor activation during the perception of speech. Such evidence suggests that speech production and speech perception make use of overlapping neural representations. Several studies have found activation of motor areas in the brain during audiovisual speech perception (e.g., Skipper, Nusbaum, & Small, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007), and also during passive listening to speech (e.g., Wilson, Saygin, Sereno, & Iacoboni, 2004). Importantly, motor activation during passive listening is articulator-specific: for example, listening to labial consonants is associated with activation of the lip representation area in motor cortex (Pulvermüller *et al.*, 2006; see Pulvermüller & Fadiga, 2010 for a review). Further, listening to speech modulates the excitability of the speech muscles that are involved in the production of the perceived sound (e.g., Fadiga, Craighero, Buccino, & Rizzolatti, 2002). Accordingly, listening to a phoneme also affects concurrent articulation of a different phoneme; for example the palatal sound/k/was produced with greater contact between the tip of the tongue and the alveolar ridge when participants were listening to the alveolar sound/t/(Yuen, Davis, Brysbaert, & Rastle, 2010).

This body of evidence supports the assumption of shared representations at the phonological level at least. However, recent findings suggest that motor activation during speech perception might occur primarily or exclusively when the task is difficult (e.g., when speech is degraded; Adank, 2012; D'Ausilio, Bufalari, Salmas, & Fadiga, 2012). Moreover, it may be that motor activation under normal listening conditions reflects listeners' tracking of the speaker's speech rate and preparation for speech in anticipation of the end of the speaker's utterance, rather than retrieval of shared content-specific phonological representations (S. K. Scott, McGettigan, & Eisner, 2009).

Figure 7.1a is intended as a schematic summary of the overview of the literature on representational parity presented in this section. It depicts what we take to be the consesus view on the issue of representational parity across levels of linguistic representations. First of all, it illustrates the fact that there is substantial consensus on parity at the semantic and syntactic level (gray *sem* and *syn* representations). Phonological representations, instead, are depicted as overlapping but not identical (partially superimposed black and white *phon* representations) to account for the fact that evidence for parity at this level might be restricted to particular tasks (as just discussed). Finally, representations at the phonetic level are labelled as speech percepts (comprehension) and speech motor commands (production) and are assumed to be separate (i.e., white speech percepts are separate from black motor command representations).

Some researchers assume substantial overlap at the phonetic level as well, but parity at this level is particularly controversial. On one hand, speakers can converge toward a model speaker at the level of low-level phonetic features that do not imply phonological distinctions (e.g., vowel duration, F0), therefore providing

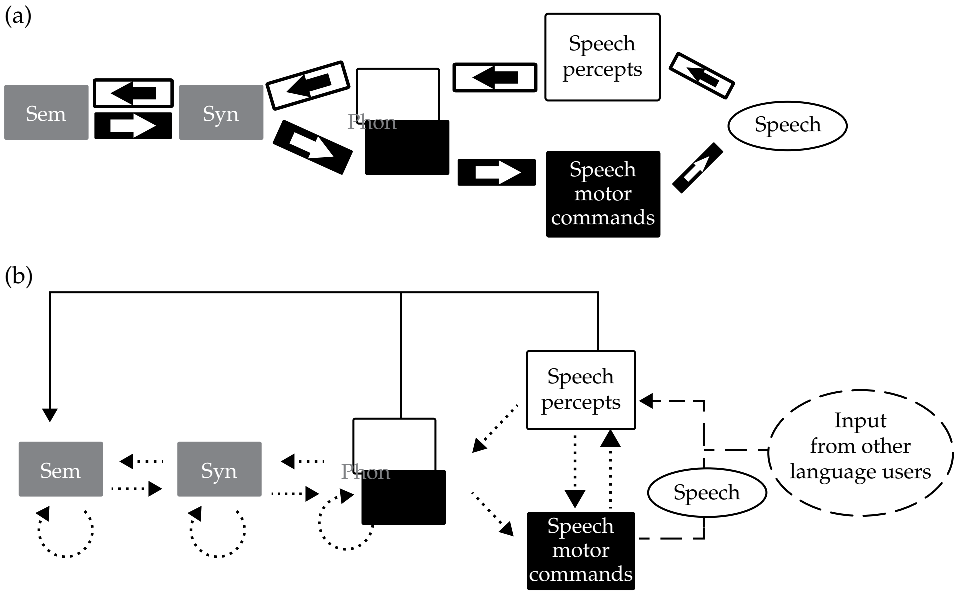



**Figure 7.1** **Panel** (a) shows production processes (white arrows on black background) and comprehension processes (black arrows on white background); **Panel** (b) shows loops: the solid line is Levelt's (1989) Perceptual Loop; the dotted arrows are Pickering and Garrod's (2013) fast loops between and within levels; the dashed arrows are external loops that include long-term effects of exposure to distributional regularities in linguistic input (see *The P-chain framework*, and *Frameworks that posit linked preferences*).

support for the hypothesis that phonetic representations are shared between comprehension and production (e.g., Goldinger, 1998; Pardo, 2006; see Chapter 6 in this volume). However, other studies failed to find evidence that listening to a phonetic variant (e.g., alveolar/r/) facilitates subsequent production of the same variant compared to listening to an alternative variant (e.g., uvular/r/) that is phonologically equivalent (Mitterer & Ernestus, 2008; Mitterer & Müsseler, 2013).

Moreover, despite some indication that adaptation to an accent in comprehension correlates with adaptation to the same accent in production at the level of an individual speaker (Evans & Iverson, 2007), there are also asymmetries between people's production and comprehension abilities in the processing of regional variability in speech. For example, Sumner and Samuel (2009) argued that listeners who have long-term perceptual experience with a dialectal phonetic variant can achieve native-like perception of that variant despite lacking native-like production representations (see also Kralijc, Brennan, & Samuel, 2008). To sum up, some findings in the literature suggest comprehension and production phonetic representations are shared or overlapping, but others support the notion of distinct comprehension and production representations at this level. Overall, the evidence for separation is stronger at the phonetic than at the phonological level. In

Figure 7.1a, we capture this by drawing separate (rather than partially overlapping) phonetic representations, but it is still unclear how much separation should be assumed at this level.

## Linked processes

Figure 7.1 does not only depict representations, but also processes. Traditionally, any process that takes place during an act of comprehension (i.e., while reading or listening) has been considered a comprehension process and, conversely, any process that takes place during an act of production (i.e., while writing or speaking), a production process. Within this tradition, the debate has focussed on whether processes that take place during production always map from semantics to syntax and from syntax to phonology (i.e., white arrows on black background, flowing from left to right in Figure 7.1a), or whether they can also map in the opposite direction (i.e., black arrows on white background, flowing from right to left in Figure 7.1a). Processes that go “backward” have been termed feedback processes and models that incorporate them (e.g., Dell, 1986) have been labelled as interactive (as opposed to purely feed-forward models; Levelt *et al.*, 1999). Similar issues have been discussed in the comprehension literature, where left-to-right processes (see Figure 7.1a) have been called top-down and right-to-left processes have been called bottom-up (e.g., Marslen-Wilson, 1987). Within the traditional view, researchers do not typically ask whether production processes can take place during comprehension or whether comprehension processes can take place during production. Instead, they tend to label any process that takes place during production as a “production process” and any process that takes place during comprehension as a “comprehension process,” regardless of the direction in which it flows.

But on an alternative view, a process is classed as a production process if it maps from representations that are higher in the linguistic hierarchy (e.g., semantics) to representations that are lower (e.g., phonology), that is from left to right in Figure 7.1a. Conversely, a process is classed as a comprehension process if it maps from lower to higher representations (e.g., from phonology to semantics), that is from right to left in Figure 7.1a. According to this definition, which we will follow in the remainder of this chapter, both production and comprehension processes could potentially be employed during any act of production, as they could during any act of comprehension (Pickering & Garrod, 2013).

Note that this difference is not merely terminological. Rather, it reflects a substantial theoretical distinction between the traditional and the alternative view, which could be tested experimentally. Should we conceptualize a processing flow that goes from phonology to syntax during an act of production as a production process (because it takes place during an act of production, as per the traditional view), or as a comprehension process (because it flows from right to left, as per the alternative view)? One way of answering this question would be by examining the neural pathways involved. For example, if we could show that the same neural pathway is involved during an act of production that implies reliance on feedback flow of information, as well as during an act of comprehension, then we would have empirical evidence for the alternative view. But if feedback during production and comprehension engage separate pathways despite a common direction in the flow of information, then the evidence would be more compatible with the traditional view.

Unfortunately, progress toward establishing how much overlap there is between production and comprehension processes has been hindered by the division between sub-disciplines within psycholinguistics. Comprehension researchers often use a production measure (e.g., word naming time) as their dependent variable, but do not interpret the production processes themselves (the "mind-in-the-mouth" assumption; Bock, 1996). Similarly, production researchers presumably realize that a task such as picture-word interference involves comprehension processes but tend to ignore them.

For example, production researchers concerned with picture-word interference have assumed that word form and phoneme representations activated by distractor words (in comprehension) send activation to related word forms and phonemes in the production network (Levelt *et al.*, 1999). But they do not ask what processes are involved in comprehension of the distractor words, focussing instead only on the end result of those processes (i.e., that some representation gets activated). For example, Damian and Martin (1999) showed that the time course of semantic and phonological effects in picture-word interference differed for visually presented distractors compared to auditorily presented distractor words, and concluded this was because of the longer presentation times used for visual distractors (in previous PWI studies). They did not consider the possibility that lexical co-activation (in comprehension) might have a different time course depending on modality (visual or auditory), and instead assumed that different presentation times led to different effects because the distractors interacted with different stages of production of the target picture name.

However, production-comprehension links have not been entirely overlooked. In fact, both feed-forward and feedback links have long been identified as a crucial component of neuro-computational theories of speech motor control (Tourville & Guenther, 2011; Hickok, 2012) and learning (Guenther & Vladusich, 2012; see also Plaut & Kello, 1999). Separately, the psycholinguistic literature on language production has also given ample consideration to this topic, albeit under the specific heading of self-monitoring. A long-standing psycholinguistic account of self-monitoring, the Perceptual Loop Theory, equates the self-monitoring system (active during acts of production) with the comprehension system (Levelt, 1983, 1989).

We first briefly present two neuro-computational models of speech motor control (Tourville & Guenther, 2011; Hickok, 2012). Then, we describe the Perceptual Loop Theory and discuss some criticisms of it. In the subsequent section, we introduce an integrated framework for language comprehension and language production (Pickering & Garrod, 2013), which includes an alternative account of self-monitoring (Pickering & Garrod, 2014), and also posits that production processes take place during acts of comprehension (and not just that comprehension processes take place during acts of production, as in self-monitoring). We then consider the P-chain framework (Dell & Chang, 2014), which makes a related proposal. Finally, we discuss proposals that place the link between production and comprehension outside specific acts of comprehension or production, in the long-term experience that speakers and listeners have with language (e.g., MacDonald, 2013).

## *Neuro-computational models of speech motor control*

In neuro-computational models of speech motor control (the Directions Into Velocities of the Articulators, or DIVA, model, see Tourville & Guenther, 2011; the Hierarchical State Feedback Control, or HSFC, model, see Hickok, 2012, 2014), forward models map from motor commands sent to the articulators to the sensory (i.e., auditory or somatosensory) consequences of executing those commands (the upward dotted vertical arrow from motor commands to speech percepts in Figure 7.1b). Forward models therefore instantiate a relatively low-level mapping between production and comprehension representations and can be considered as internal models of the language system (that is, models of the processes that cause the articulation of speech sounds during acts of production). The inverse mapping corresponds to feedback-based correction of speech movements (downward dotted vertical arrow from speech percepts to motor commands in Figure 7.1b).

Evidence that forward models are implicated in speech production comes from the finding that auditory responses to speech sounds are suppressed during speaking compared to listening (e.g., M100 suppression, as reported using magneto-encephalography in the study of Houde, Nagarajan, Sekihara, & Merzenich, 2002). This is thought to occur because forward models can be used during production to anticipate sensory stimulation and cancel it out (in a way that could be useful in distinguishing between self-generated and externally-generated sounds). In support of this claim, auditory responses are suppressed in a stimulus-specific manner: suppression occurs during covert rehearsal compared with a control task, but only when the rehearsed stimulus matches (part of) the perceived stimulus (Ylinen *et al*., 2014). In addition, enhancement rather than suppression takes place when auditory feedback is altered unexpectedly during speaking, for example by shifting pitch upwards or downward in real-time, so that the predicted stimulation ceases to match actual stimulation (e.g., Chang, Niziolek, Knight, Nagarajan, & Houde, 2013).

All of these studies are compatible with forward-model predictions operating at the level of fine-grained phonetic features. But there is also some indication that such predictions might be phonological in nature. Niziolek, Nagarjan, and Houde (2013) showed the degree of suppression is larger for sounds that are closer to a given speaker's median productions, suggesting that predictions could be computed on the basis of somewhat abstract representations. In other words, when the speaker selects a motor command to execute, the anticipated sensory consequences might correspond to an abstract phonological target (i.e., what it should sound like) rather than to a detailed phonetic target (i.e., what it is going to sound like on this particular instance).

This finding provides support for the HSFC model (Hickok, 2012, 2014). In this model, forward model predictions operate at two hierarchically organized levels: phonemes and syllables. Motor programs (corresponding to planned syllables and planned phonemes) inhibit sensory areas where perceptual targets are represented. These same areas are activated via feedback from the movements of the vocal tract and from the resulting speech output (i.e., when the speaker perceives her own productions). In addition, they can receive activation from concepts and lemmas. The discrepancy between the expected activation (propagated in the form of inhibitory connections from the motor targets) and the actual activation constitutes a prediction error, which is propagated back to the motor target areas and used for online corrections of motor programs (and learning of more accurate motor-to-sensory mappings).

The inhibitory connections therefore implement a form of prediction that maps from motor commands to expected sensory consequences of executing those commands. The excitatory backward connections, instead, implement a form of inverse correction, which maps from sensory prediction errors to changes in the motor commands needed to compensate for those errors. Evidence for a fast-cycling loop at the phonetic level, in which motor representations are rapidly mapped onto sensory representations and *vice versa*, comes from several demonstrations that speakers compensate very quickly for perturbed auditory feedback (e.g., Houde & Jordan, 1998; Jones & Munhall, 2002; Tourville, Reilly, & Guenther, 2008).

The HSFC model is closely related to the DIVA/GODIVA model proposed by Guenther and colleagues (see Tourville & Guenther, 2011). This model also incorporates the notion that somatosensory and auditory target areas are activated via forward-model predictions as well as via processing of sensory input, and that prediction errors are used for online correction (as well as for learning the mappings between movements and their sensory consequences). Importantly, both models incorporate what we might call an account of self-monitoring. They assume that a process that maps from motor areas to sensory areas (and *vice versa*) is essential for online control during speech production (see Plaut & Kello, 1999 for another computational model that instantiates this idea).

Crucially, unlike psycholinguistic theories of self-monitoring (see below), these neuro-computational models focus on sound-level representations and processes, and say very little about other linguistic levels. Hickok (2012, 2014) argued for the importance of integrating psycholinguistic theories of language production with models of speech production, and integrated lemma and conceptual representations within his HSFC. However, he did not explicitly extend the forward-model architecture to these levels. In terms of Figure 7.1b, his model assumes that dotted arrows flow in both directions between all levels, indicating that both comprehension and production processes take place during language production. But the HSFC model includes a fast-cycling within-level loop (recursive dotted arrows), which is responsible for fast error correction during production, at the phonological and phonetic levels only. No such loop is explicitly assumed at the lemma and conceptual levels.

## *An early model linking production and comprehension: The Perceptual Loop Theory of self-monitoring*

According to the Perceptual Loop Theory (Levelt, 1983, 1989), production errors are detected via comprehension of speech output (the external loop), and also via comprehension of phonological representations (the internal loop). The comparison process takes place at the level of communicative intentions (messages): If the message reconstructed by the comprehension system does not match the message originally intended, the monitor flags up an error. This comprehension-production loop is depicted in Figure 7.1b using a solid black line. This loop links comprehension representations at the phonological (phonological comprehension representations) and phonetic level (the speech percepts formed during comprehension of the speech output) to semantic representations (that are shared between comprehension and production).

Crucially, the Perceptual Loop Theory is open to criticism because it posits a relatively slow-cycling loop. First, the speech signal must be analyzed by comprehension processes to recover speech percepts (if using the external loop), or phonological representations retrieved by production processes have to be analyzed by comprehension processes to activate corresponding phonological representations in the comprehension network (if using the internal loop). Additional comprehension processes then map from sound-based comprehension representations to a semantic comprehension representation. Finally, the activated semantic representation in the comprehension network is compared to the semantic representation that was originally activated in the production network (the latter process is facilitated by shared representations at the semantic level). In addition, if a discrepancy is detected, production of the current utterance must be stopped before production of a replacement can start.

Oomen and Poostma (2002) found that having participants engage in a concurrent task while speaking caused them to stop speaking more quickly after the onset of an error (Oomen & Postma, 2002). But if the time it takes to stop were attributable to a comprehension-based loop, then one would have expected that drawing attention away from the speech signal (as in a dual task condition) would have led to longer, not shorter stopping times.

Hartsuiker and Kolk (2001) criticized the Perceptual Loop Theory assumption that production of the erroneous utterance must stop before planning of the replacement begins. Instead, they proposed that stopping the current utterance and preparing a replacement can proceed in parallel. Indeed, speakers can often resume very quickly following an interruption (in less than 100ms; Blackmer & Mitton, 1991), which suggests that they start planning the replacements before they stop articulation. More direct evidence comes from Hartsuiker, Catchpole, de Jong, and Pickering (2008; see also Gambi, Cop, & Pickering, 2015), who showed that the time it takes to stop a word depends on how difficult it is to prepare a replacement word.

However, Hartsuiker and Kolk (2001) also assume that the monitor needs to detect an error in the phonological representation before sending a signal to stop production. Interestingly, there are also psycholinguistic theories of self-monitoring that posit a purely production-based monitor (see Postma, 2000 for discussion). For example, Nozari, Dell, and Schwartz (2011) proposed that error detection is based on the amount of noise associated with production processes. One argument in favor of production-based accounts is that they allow for very rapid error detection at all levels of the linguistic hierarchy. Below, we introduce Pickering and Garrod's (2013, 2014) comprehension-based theory of self-monitoring theory, which addresses this issue by allowing the monitor to compare expected and actual comprehension representations at all linguistic levels, as soon as these representations become available.

## *Prediction during production and comprehension: The integrated theory of language production and comprehension*

Pickering and Garrod (2013) described an integrated theory of language production and comprehension that is based on the notion of forward models. This notion is derived from the motor control literature (e.g., Wolpert, 1997) and is also part of the models reviewed in *Neuro-computational models of speech motor control*. Crucially, Pickering and Garrod generalize it, by making the assumption that forward models are computed at all levels of the linguistic hierarchy, and that they are involved not only in language production but also in comprehension. Below, we first describe how forward models are implicated in self-monitoring, and then how they are implicated in prediction during comprehension.

During an act of production, the speaker forms a communicative intention (production command), which corresponds to the pre-linguistic message that the speaker intends to convey. The production command is sent to the production system (or, in Pickering and Garrod's terminology, the production implementer), and it triggers the retrieval of a set of production representations (semantics, syntax, and phonology). For example, if a speaker sees a kite and forms the intention to name this object, production processes would cause the retrieval of the corresponding concept (KITE), lemma (*kite*) and phonological form (/kaIt/). Importantly, it takes several hundred milliseconds to retrieve such representations (see Indefrey & Levelt, 2004, for estimates). Once production representations have been retrieved, they can be processed by the comprehension system (or, in Pickering and Garrod's terminology, the comprehension implementer).[1] Crucially, the theory assumes that the comprehension implementer has immediate access to production representations at all levels; so, for example, the semantic representation retrieved during production can be immediately comprehended, even before a phonological representation is built. In this respect, the proposal differs from the Perceptual Loop Theory of self-monitoring (Levelt, 1983).

In addition to retrieval of representations within the production and comprehension implementers, during an act of production a copy of the production command is sent to a forward model, which maps from the production command to the predicted comprehension representations that are about to be retrieved as a consequence of executing that production command. To return to our example of a speaker intending to name the picture of a kite, a forward model of this process could compute a prediction of aspects of the semantics (it's a flyable object), of the syntax (it's a noun),[2] and the phonology (it starts with a consonant), *before* the corresponding production representations are retrieved from memory. Therefore, predicted representations are typically ready before actual (implemented) representations. The process of self-monitoring constitutes the comparison between predicted and actual comprehension representations within the comparator (at any linguistic level). The resulting difference (the prediction error in motor control terms; Wolpert, 1997) can be used to drive online corrections (and learning), just as it can in the models described in section *Neuro-computational models of speech motor control*.

In sum, the account of self-monitoring proposed by Pickering and Garrod (2013, 2014) differs from the Perceptual Loop Theory (Levelt, 1983) in that it posits loops between production and comprehension at all levels of the linguistic hierarchy, and both within and between levels (not just from phonetics and phonology to semantics; see dotted arrows in Figure 7.1b). Moreover, such loops are faster than the loops assumed by the Perceptual Loop Theory, because they are based on comparisons between predicted and actual comprehension representations, with predicted comprehension representations being the outcome of production processes (i.e., left-to-right dotted arrows in Figure 7.1b).

With regard to acts of comprehension, Pickering and Garrod (2007) proposed that the collection of cognitive mechanisms underlying prediction during language comprehension coincides with the language production system. Federmeier (2007) made a similar proposal based on evidence that the left hemisphere is more sensitive to predictability of upcoming words than the right hemisphere (and the neural substrate for language production is predominantly left-lateralized in Broca's area), and Dell and Chang (2014) have recently reinstated this idea as the core principle of their P-chain framework (see *The P-chain framework*, this chapter).

An earlier proposal (Kempen, 2000, 2014) argued that grammatical encoding (production) and grammatical decoding (comprehension) are performed by the same processing architecture. Parallels between sentence comprehension and sentence production (e.g., similar patterns of errors occur during subject-verb agreement in both production and comprehension; Bock & Miller, 1991; Pearlmutter, Garnsey, & Bock, 1999), as well as evidence from structural priming from comprehension to production (Bock *et al*., 2007; Branigan *et al*., 2000) are consistent with this proposal. However, such evidence is indirect and can be explained by shared representations without shared processes. In addition, while it is clear how production and comprehension could share some processes (e.g., retrieving syntactic frames), it also necessary to assume some processing differences in order to explain the different start- and end-points. Note that the proposal that production processes are related to only a subset of comprehension processes, namely those involved in prediction during acts of comprehension, is not subject to this criticism.

But is there evidence for this hypothesis? Some evidence that production processes might be used during syntactic aspects of comprehension comes from a study by Kempen, Olsthoorn, and Sprenger (2012). They asked Dutch participants to paraphrase sentences from direct (e.g., *De lottowinnaar/zei:/"Ik/heb besloten/een rode auto/te kopen/voor mezelf*," The lottery winner said: "I have decided to buy a red car for myself") into indirect speech (e.g., *De lottowinnaar zei dat hij had besloten een rode auto te kopen voor zichzelf*, The lottery winner said that he had decided to buy a red car for himself) as they read them (i.e., fragment by fragment, as marked in the example). When the sentence contained an ungrammatical reflexive pronoun (e.g., the third-person reflexive pronoun in the sentence *De lottowinnaar zei: "Ik heb besloten een rode auto te kopen voor zichzelf*"), participants were faster producing the paraphrase (which contained the same third-person pronoun) than when the sentence contained a grammatical reflexive (i.e., *mezelf*), despite the fact that the ungrammaticality should have led to a processing delay in comprehension. One interpretation of these findings is that participants' expectations generated during the comprehension of the input sentences were replaced, on-line, by the expectations generated during concurrent encoding of the paraphrase (in which the same pronoun was grammatical).

Moreover, Federmeier, Kutas, and Schul (2010) showed that a late prefrontal positivity induced by plausible but unexpected nouns (which is thought to index the updating of disconfirmed predictions after an unexpected word has been encountered; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007) is greatly reduced in older adults (compared to younger adults) and, importantly, the magnitude of this component in the older group correlated with production measures of verbal fluency. Similarly, Mani and Huettig (2012) found that two-year-olds with larger production (but not comprehension) vocabularies were more likely to predict upcoming referents (as indexed by looks to corresponding pictures in the so-called visual world paradigm; cf. Altmann & Kamide, 1999). More recently, a similar correlation between verbal fluency and prediction abilities during language comprehension (again, measured using the visual world paradigm) was reported for young adults as well, but only when listeners could preview pictures in the visual display (and presumably started retrieving their names; Hintz, Meyer, & Huettig, 2014). These studies suggest that the ability to predict during language comprehension is correlated with language production abilities at least in some task contexts.

In accordance with this and related evidence, Pickering and Garrod (2013) proposed that production processes underline comprehenders' ability to predict what another is about to produce. This route to comprehension is termed prediction-by-simulation. It starts with the comprehender covertly imitating the producer: this means that, based on the initial part of the producer's utterance, or on contextual information (i.e., what he assumes about the producer from previous interactions, or from background knowledge), the comprehender recovers the most likely intention (production command) underlying the utterance at time *t*. He can then run this command through the production implementer. If he does that, he will end up imitating the producer. This mechanism therefore explains alignment (Pickering & Garrod, 2004; see *Representational parity*).

In addition, the comprehender can run ahead the command that he recovered at time *t*, thus predicting what he would be likely to utter next if he were in the producer's shoes. If he runs this new command through the production implementer, he will be able to complete the producer's utterance (see *Dialogue*). If he runs this new command through the forward production model, the comprehender may generate the predicted semantics, syntax, and phonology at time $t+1$. When the producer continues his utterance, the comprehender builds comprehension representations for the actual utterance at $t+1$ and can compare them to the representations he had predicted. He can then use the resulting discrepancy to adjust the recovered production command, thus revising his understanding of the intention underlying the producer's utterance.

In addition to the correlational evidence cited above, one study established a causal link between production processes and prediction during language comprehension (Lesage, Morgan, Olson, Meyer, & Miall, 2012). This study applied repetitive Transcranial Magnetic Stimulation (rTMS) to the right cerebellum. In rTMS, as in other types of TMS, a magnetic coil is used to induce small electric currents in a particular area of the brain; in particular, several pulses are delivered at a low frequency, which is known to suppress neural activity for some time after stimulation has ended. Importantly, evidence suggests that forward model computations related to motor execution take place in the cerebellum (e.g., Wolpert, Miall, & Kawato, 1998), and some have linked it to the computation of internal models in general (e.g., Ito, 2008). Disrupting activity in the right cerebellum caused participants in this study to delay their eye-movements to predictable visual referents during sentence comprehension (but not to unpredictable visual referents). Crucially, other conditions with no stimulation or stimulation to a control site did not show the same selective effect. Therefore, this study suggests that forward model computations might support prediction during comprehension.

Note that, because the forward model is functionally distinct from the production implementer, Pickering and Garrod's (2013) account does not claim that full activation of the production implementer will be observed whenever prediction-by-simulation is used. For example, the account does not predict that activation of language production areas in the brain will be always observed during language comprehension. Rather, such activation is most likely to occur under conditions in which production is relied upon more. According to Pickering and Garrod, this is the case when comprehension is difficult (e.g., in a noisy environment; see Adank, 2012).

But in addition, there is another route to prediction available in comprehension, which they termed prediction-by-association. This route does not involve production processes (i.e., forward models or the production implementer), and could be used whenever covert imitation fails. This route to prediction in comprehension makes use of regularities in the input to the process of comprehension. Unlike prediction-by-simulation, it does not rely on knowledge of how we produce language. Instead, it relies on our ability to learn regular patterns of perceptual events, which applies equally to domains in which we have the ability to generate the patterns through action and domains in which we lack this ability (e.g., predicting the sound of leaves moved by the wind).

In sum, Pickering and Garrod's (2013) account allows for the existence of processes that are not (usually) shared between acts of comprehension and acts of production. In other words, there are some production processes that do not always operate during acts of comprehension (e.g., processes involved in retrieving articulatory programs; prediction-by-simulation via forward model computations), and there are comprehension processes that do not operate during acts of production (e.g., the prediction-by-association route).

## *The P-chain framework*

The P-chain framework (Dell & Chang, 2014) claims that the process responsible for prediction during acts of comprehension is the same process that is used during acts of production. Therefore, although it does not claim that all processes are shared between comprehension and production, it posits that there is a single cognitive architecture subserving both tasks.

This assumption stems from the architecture of Chang, Dell, and Bock's (2006) Dual Path model of sentence acquisition (in children) and structural priming (in children and adults). This model is a recurrent neural network; during training, the model learns to predict the next word in the input (as in Elman, 1990). Prediction errors are generated by comparing the predicted with the actual comprehended input, and are used to change the weights between units in the network, so that the model learns to correctly predict grammatical word sequences in the future. Sometimes the model uses meaning (inferred from context) to help in this prediction process and the ability to do prediction from meaning is the same mechanism that the model uses for production.[3]

Dell and Chang (2014) proposed that the model's architecture instantiates a set of principles that govern the functioning of the cognitive system for language. They termed this collection of principles the *P-chain framework*. The framework highlights the tight links existing between language comprehension (which they term *processing*) and *production*. Such links are organized in a chain, or loop, of cause-effect relations. In a nutshell, the use of production-based prediction mechanisms during language comprehension generates prediction errors, which in turn drive changes in the language system (i.e., they make it more likely to generate predictions in line with previous input), thus providing an explanation for structural priming effects, that is for the fact that comprehending a given sentence structure makes it more likely that the same structure will be selected in a subsequent act of production. When these changes build up over time, they can explain acquisition of structural representations for different languages (English: Twomey, Chang, & Ambridge, 2014; Japanese: Chang, 2009; German: Chang, Baumann, Pappert, & Fitz, 2015). Finally, input regularities, on which comprehension is tuned, are themselves the output of the mechanism responsible for language production (in other speakers); therefore, there is also a slow-cycling loop through which the production processes of other speakers provide the input that trains comprehension processes. This final link is inspired by the Production-Distribution-Comprehension account (MacDonald, 2013) that is examined in the next section.

## *Frameworks that posit linked preferences*

Several theorists have appealed to the idea that regular patterns in language use emerge as a consequence of the constraints imposed by communication, and that such patterns in turn affect how speakers and listeners process language. This approach corresponds to positing a long-term loop (dashed arrow in Figure 7.1b) that is external to the language production and comprehension architecture (i.e., outside the mind/brain of individual language users).

One version of this idea is that speaker choices in production are constrained by ease-of-comprehension principles. For example, the Hyper- And Hypo-Articulation model of speech production (Lindblom, 1990) claims that speakers' tendency to reduce articulation (and therefore their own effort) is constrained by the necessity for listeners to recover the intended message. This leads speakers to counteract the tendency to hypo-articulate (i.e., to produce forms that are reduced in duration and/or intensity), precisely in those contexts in which the listener cannot draw on other sources of information (i.e., other than the speech signal) to infer meaning. There are other versions of this proposal in phonetics (e.g., the Smooth Signal Redundancy hypothesis, Aylett, & Turk, 2004).

Similarly, the Uniform Information Density hypothesis (UID; Levy & Jaeger, 2007; Jaeger, 2010) claims that producers strive to keep information transfer rate within the range of the comprehender's processing rate (i.e., channel capacity); in this way, they avoid conveying too much information or too little information per unit (word, phoneme, etc.). This hypothesis can also be phrased in terms of surprisal, which is the predictability of a unit in context. High surprisal means that a unit is unpredictable based on previous context, and therefore adds information in that context, whereas low surprisal means that a unit is highly predictable and adds little information. UID claims that when given a choice speakers will tend to use a structure that keeps surprisal relatively constant across units. Production preferences are therefore explained as stemming from the limitations of the comprehension system. Interestingly, these preferences may be shaped by learning: Jaeger and Snider (2013) provided evidence that the higher the surprisal of a structural alternative in comprehension, the more likely that alternative is to be subsequently preferred in production.

The idea that production preferences can be explained with reference to comprehension is also related to the Audience Design Hypothesis: the notion that producers take into account their addressee's knowledge when planning their utterances (e.g., Brown-Schmidt, 2009; Clark, 1996). A review of this literature is beyond the scope of this chapter. However, we note that experimental evidence for the extent to which audience design affects production is mixed. In particular, there is controversy over the rapidity with which the addressee's knowledge can affect production. Keysar and colleagues have argued that producers are egocentric (e.g., Horton & Keysar, 1996), and take into account what their addressee can or cannot know only when given sufficient time and during a relatively late stage of production. In addition, it has been suggested that while speakers might adapt at the level of lexical choices, they in fact do not do so at the phonetic level (Bard *et al.*, 2000; but see Galati & Brennan, 2010 for criticism, and Arnold, Kahn, & Pancani, 2012 for some evidence that speakers might adapt at least at the level of production speed). Finally, addressees are clearly facilitated when speakers adopt previously established referential labels, but we do not know whether such facilitation occurs because listeners are sensitive to mutual knowledge between them and the speaker (Brown-Schmidt, 2009) or simply because repetition increases availability (Barr & Keysar, 2002).

Another related view is the Production-Distribution-Comprehension account (MacDonald, 2013), which assumes that the relevant constraints shaping patterns of language use over time relate to production rather than to comprehension. In particular, it assumes three principles: (i) Easy First: Words that are more easily retrieved are produced first (e.g., Bock & Warren, 1985); (ii) Plan Reuse: Utterance plans that have been recently used tend to be reused (e.g., Bock, 1986); and (iii) Reduce Interference: Elements that are more similar to one another (and therefore tend to interfere in memory) are placed farther apart (Gennari, Mirković, & MacDonald, 2012 for semantic similarity; Jaeger, Furth, & Hilliard, 2012 for phonological similarity). The claim, then, is that these constraints on production lead to distributional regularities to which comprehenders adapt as they accumulate linguistic experience. Therefore, linguistic forms that are easier to produce become easier to comprehend as well. This account thus explains both why certain structures are easier to comprehend than others and cross-linguistic patterns of language variation (i.e., typology). We refer the reader to Chapter 3 in this volume for an in-depth discussion of the evidence in favor and against the frameworks briefly introduced in this section.

## Dialogue

As mentioned in the Introduction, the nature of language use in dialogue (i.e., conversation) is one key motivation for positing links between comprehension and production. First, each participant in a dialogue regularly has to switch between acts of production and acts of comprehension. Such switches do not only occur between dialogue turns, but also within a turn, as listeners produce backchannels (e.g., *Yes, OK*, or *eh?*) to provide continuous feedback to the speaker. Moreover, such switches occur rapidly, as long intervals between turns are rare (Stivers *et al.*, 2009). Finally, such switches can occur at any point within an utterance, with listeners taking over from speakers even after single words or incomplete constituents, and sometimes producing grammatical and pragmatically appropriate completions to these fragments (e.g., Clark & Wilkes-Gibbs, 1986; Lerner, 1991).

These phenomena demonstrate that the output of comprehension processes can rapidly affect production processes. For example, understanding the speaker's utterance leads to rapid backchannel responses from the listener and such backchannels can be quickly acted upon by the speaker. This suggests that loops must exist between comprehension processes and production processes, and that these loops must operate at a relatively fast rate. Moreover, the phenomenon of collaborative turn completions also suggests continuity between production and comprehension processes. Take the excerpt below (from Kurtić, Brown, & Wells, 2013, p. 726). B produces his utterance as a completion to the first part of A's utterance (*so I'm not sure whether they'll still be so willing to volunteer but I'll*) and it is so well-timed that it overlaps with the end of A's own turn (brackets indicate the start and end of speech produced in overlap by A and B). In turn, A shows evidence of having understood (and accepting) B's completion, despite it overlapping with her own turn; in fact, she goes as far as repeating B's utterance word by word.

A: …so I'm not sure whether they'll still be so willing to volunteer but I'll [send them an email and ask]
B: [tell them about the free lunch]
A: I'll tell them about the free lunch

Finally, inter-turn intervals tend to cluster around a value that varies between 0 and 200ms (across languages; Stivers *et al.*, 2009), with both long gaps and long overlaps being comparatively rare. This further suggests close links between production and comprehension, and has prompted the suggestion that comprehenders might be able to anticipate turn ends (de Ruiter, Mitterer, & Enfield, 2006; Magyari & de Ruiter, 2012).

## Summary

In this chapter, we have described several models that posit explicit links between production and comprehension. It is generally agreed that information about concepts, lemmas, and syntactic frames is shared between production and comprehension processes. But opinions diverge on the degree of sharing of phonological information, and most theorists assume that phonetic representations are separate (despite some dissent). Both the language production and the language comprehension literatures have internal debates about the directionality of processes, but such debates have not been framed as debates about the sharing of processes between production and comprehension until quite recently.

In Figure 7.1a, we assumed that any process that maps from "higher" linguistic levels (semantics) to "lower" linguistic levels (phonology) should be named a production process (left-to-right arrows), and that every process that maps in the inverse direction (right-to-left arrows) should be named a comprehension process. Based on this definition, we identified three accounts that assume comprehension processes take place during acts of production, in the form of self-monitoring: Levelt's (1989) Perceptual Loop theory (solid lines in Figure 7.1b), neuro-computational models of speech motor control (such as Hickok, 2012), and Pickering and Garrod's (2013) integrated account of language production and comprehension (dotted lines in Figure 7.1b). Such accounts differ in terms of the nature and speed of the loops they assume exist between production and comprehension.

We presented two accounts that assume that production processes take place during acts of comprehension to support prediction: Pickering and Garrod's integrated account, and the P-chain framework. In addition, we briefly described a number of frameworks that posit slower-cycling loops between production and comprehension; that is, loops that are mediated by long-term experience with language and that can explain the development of linguistic preferences (dashed lines in Figure 7.1b). Overall, many researchers assume some degree of sharing of processes, but the range of views on this issue is far wider than on the issue of shared representations. Finally, we noted that the assumption of links between language production and language comprehension is also motivated by language use in dialogue. We believe that more explicit theorizing on the relations between production and comprehension processes, in both monologue and dialogue, would benefit the field.

## Acknowledgments

We would like to thank Franklin Chang and an anonymous reviewer for useful discussion.

## NOTES

1 Note that the description provided by Pickering and Garrod (2013), and summarized here, appears to imply separate production and comprehension representations at all levels. In practice, the theory is consistent with the idea that the same representations are accessed during production as well as during comprehension. The two sets of representations they assume correspond to the output stage of production and comprehension processes respectively, but such processes may have access to the same pool of representations.

2 In the example, we focus on single word retrieval. However, Pickering and Garrod (2013) have also discussed this process in relation to constituent ordering (p. 339).

3 Note that within this single cognitive architecture, the Dual path model incorporates separate weights for the word-to-syntax ("comprehension") and syntax-to-word ("production") directions, but both of these representations are hypothesized to be used in both comprehension and production tasks.

## REFERENCES

Adank, P. (2012). The neural bases of difficult speech comprehension and speech production: Two Activation Likelihood Estimation (ALE) meta-analyses. *Brain and Language*, 122(1), 42–54. doi: 10.1016/j.bandl.2012.04.014

Alario, F.X., Segui, J., & Ferrand, L. (2000). Semantic and associative priming in picture naming. *The Quarterly Journal of Experimental Psychology*, 53A(3), 741–764. doi: 10.1080/027249800410535

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. doi: 10.1016/S0010-0277(99)00059-1

Arnold, J. E., Kahn, J. M., & Pancani, G. C. (2012). Audience design affects acoustic reduction via production facilitation.

*Psychonomic Bulletin & Review*, 19(3), 505–512. doi: 10.3758/s13423-012-0233-y

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47(1), 31–56. doi: 10.1177/00238309040470010201

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42(1), 1–22. doi: 10.1006/jmla.1999.2667

Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, 46(2), 391–418. doi: 10.1006/jmla.2001.2815

Bates, E. (1993). Comprehension and production in early language development: Comments on Savage-Rumbaugh et al. *Monographs of the Society for Research in Child Development, Serial No. 233*, 58*(3–4)*, 222–242. doi: 10.1111/j.1540-5834.1993.tb00403.x

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, 39(3), 173–194. doi: 10.1016/0010-0277(91)90052-6

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, 104(3), 437–458. doi: 10.1016/j.cognition.2006.07.003

Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review*, 3(4), 395–421. doi: 10.3758/BF03214545

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, 104(3), 437–458. doi: 10.1016/j.cognition.2006.07.003

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23(1), 45–93. doi: 10.1016/0010-0285(91)90003-7

Bock, K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47–67. doi: 10.1016/0010-0277(85)90023-X

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75(2), B13–B25. doi: 10.1016/S0010-0277(99)00081-5

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482–1493. doi: 10.1037/0278-7393.22.6.1482

Brown-Schmidt, S. (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, 61(2), 171–190. doi: 10.1016/j.jml.2009.04.003

Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proceedings of the National Academy of Sciences*, 110(7), 2653–2658. doi: 10.1073/pnas.1216827110

Chang, F. (2009). Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language*, 61*(3)*, 374–397. doi: 10.1016/j.jml.2009.07.006

Chang, F., Baumann, M., Pappert, S., & Fitz, H. (2015). Do lemmas speak German? A verb position effect in German structural priming, *Cognitive Science*, 39*(5)*, 1113–1130. doi: 10.1111/cogs.12184

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272. doi: 10.1037/0033-295X.113.2.234

Clark, H. H. (1996). *Using language*. Cambridge University Press. Cambridge, U.K.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1–39. doi: 10.1017/cbo9780511620539

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(2), 345–361. doi: 10.1037/0278-7393.25.2.345

D'Ausilio, A., Bufalari, I., Salmas, P., & Fadiga, L. (2012). The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*, 48(7), 882–887. doi: 10.1016/j.cortex.2011.05.017

Dell, G.S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93 (3), 283–321. doi: 10.1037/0033-295X.93.3.283

Dell, G. S., & Chang, F. (2014). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 20120394. doi: 10.1098/rstb.2012.0394

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, 82*(3)*, 515–535. doi: 10.1353/lan.2006.0130

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211. doi: 10.1207/s15516709cog1402_1

Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the Acoustical Society of America*, 121(6), 3814–3826. doi: 10.1121/1.2722209

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, 15(2), 399–402. doi: 10.1046/j.0953-816x.2001.01874.x

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505. doi: 10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, 115(3), 149–161. doi: 10.1016/j.bandl.2010.07.006

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75–84. doi: 10.1016/j.brainres.2006.06.101

Fowler, C. A., Brown, J. M., Sabadini, L., & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language*, 49(3), 396–413. doi: 10.1016/S0749-596X(03)00072-X

Galantucci, B., Fowler, C. A., & Goldstein, L. (2009). Perceptuomotor compatibility effects in speech. *Attention, Perception, & Psychophysics*, 71(5), 1138–1149. doi: 10.3758/APP.71.5.1138

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13(3), 361–377. doi: 10.3758/BF03193857

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, 62(1), 35–51.

Gambi, C., Cop, U., & Pickering, M. J. (2015). How do speakers coordinate? Evidence for prediction in a joint word-replacement task. *Cortex*, 68, 111–128, doi: 10.1016/j.cortex.2014.09.009

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27(2), 181–218. doi: 10.1016/0010-0277(87)90018-7

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8(1), 8–11. doi: 10.1016/j.tics.2003.10.016

Gennari, S. P., Mirković, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive*

*Psychology*, 65(2), 141–176. doi: 10.1016/j.cogpsych.2012.03.002

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279. doi: 10.1037/0033-295X.105.2.251

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, 25(5), 408–422. doi: 10.1016/j.jneuroling.2009.08.006

Hartsuiker, R. J., & Kolk, H. H. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, 42(2), 113–157. doi: 10.1006/cogp.2000.0744

Hartsuiker, R. J., Catchpole, C. M., de Jong, N. H., & Pickering, M. J. (2008). Concurrent processing of words and their replacements during speech. *Cognition*, 108(3), 601–607. doi: 10.1016/j.cognition.2008.04.005

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, 13(2), 135–145. doi: 10.1038/nrn3158

Hickok, G. (2014). Towards an integrated psycholinguistic, neurolinguistic, sensorimotor framework for speech production. *Language, Cognition and Neuroscience*, 29(1), 52–59. doi: 10.1080/01690965.2013.852907

Hintz, F., Meyer, A.S., & Huettig, F. (2014). The influence of verb-specific featural restrictions, word associations, and production-based mechanisms on language-mediated anticipatory eye-movements. *Proceedings of the 27th Annual CUNY conference on human sentence processing*. Ohio State University. Columbus, OH.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground?. *Cognition*, 59(1), 91–117. doi: 10.1016/0010-0277(96)81418-1

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, 279(5354), 1213–1216. doi: 10.1126/science.279.5354.1213

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *Journal of Cognitive Neuroscience*, 14(8), 1125–1138. doi: 10.1162/089892902760807140

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92(1), 101–144. doi: 10.1016/j.cognition.2002.06.001

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9(4), 304–313. doi: 10.1038/nrn2332

Jarick, M., & Jones, J. A. (2008). Observation of static gestures influences speech production. *Experimental Brain Research*, 189(2), 221–228. doi: 10.1007/s00221-008-1416-7

Jaeger, F.T. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61(1), 23–62. doi: 10.1016/j.cogpsych.2010.02.002

Jaeger, T. F., Furth, K., & Hilliard, C. (2012). Incremental phonological encoding during unscripted sentence production. *Frontiers in Psychology*, 3, doi: 10.3389/fpsyg.2012.00481

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, 127(1), 57–83. doi: 10.1016/j.cognition.2012.10.013

Jones, J. A., & Munhall, K. G. (2002). The role of auditory feedback during phonation: studies of Mandarin tone production. *Journal of Phonetics*, 30(3), 303–320. doi: 10.1006/jpho.2001.0160

Kempen, G. (2000). Could grammatical encoding and grammatical decoding be subserved by the same processing module? *Behavioral and Brain Sciences*, 23(1), 38–39. doi: 10.1017/S0140525X00402396

Kempen, G. (2014). Prolegomena to a neurocomputational architecture for human grammatical encoding and decoding. *Neuroinformatics*, 12(1), 111–142. doi: 10.1007/s12021-013-9191-4

Kempen, G., Olsthoorn, N., & Sprenger, S. (2012). Grammatical workspace sharing during language production and language comprehension: Evidence from grammatical multitasking. *Language and Cognitive Processes*, 27(3), 345–380. doi: 10.1080/01690965.2010.544583

Kerzel, D., & Bekkering, H. (2000). Motor activation from visible speech: evidence from stimulus response compatibility. *Journal of Experimental Psychology: Human Perception and Performance*, 26(2), 634–647. doi: 10.1037/0096-1523.26.2.634

Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language*, 74(1), 1–25. doi: 10.1006/brln.2000.2315

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition*, 107(1), 54–81. doi: 10.1016/j.cognition.2007.07.013

Kurtić, E., Brown, G. J., & Wells, B. (2013). Resources for turn competition in overlapping talk. *Speech Communication*, 55(5), 721–743. doi: 10.1016/j.specom.2012.10.002

Lerner, G. H. (1991). On the syntax of sentences-in-progress. *Language in Society*, 20(03), 441–458. doi: 10.1017/S0047404500016572

Lesage, E., Morgan, B. E., Olson, A. C., Meyer, A. S., & Miall, R. C. (2012). Cerebellar rTMS disrupts predictive language processing. *Current Biology*, 22(18), R794–R795. doi: 10.1016/j.cub.2012.07.006

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41–104. doi: 10.1016/0010-0277(83)90026-4

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, 14(1), 78–106. doi: 10.1016/0010-0285(82)90005-6

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–38. doi: 10.1017/S0140525X99001776

Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In B. Schöllkopf, J. Platt, and T. Hoffman (Eds.). *Advances in Neural Information Processing Systems*, Vol. 19 (pp. 849–856), Cambridge, MA: MIT Press.

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, 4(5), 187–196. doi: 10.1016/S1364-6613(00)01471-6

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W.J. Hardcastle and A. Marchal (Eds.). *Speech Production and Speech Modelling* (pp. 403–439). Springer Netherlands.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4, doi:10.3389/fpsyg.2013.00226

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.

MacKay, D. G., Abrams, L., & Pedroza, M. J. (1999). Aging on the input versus output side: Theoretical implications of age-linked asymmetries between detecting versus retrieving orthographic information. *Psychology and Aging*, 14(1), 3–17. doi: 10.1037/0882-7974.14.1.3

Magyari, L., & De Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in Psychology*, 3, doi:10.3389/fpsyg.2012.00376

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers.

*Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 843–847. doi: 10.1037/a0029284

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1), 71–102. doi: 10.1016/0010-0277(87)90005-9

Menenti, L., Gierhan, S. M., Segaert, K., & Hagoort, P. (2011). Shared language overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science*, 22(9), 1173–1182. doi: 10.1177/0956797611418347

Mitterer, H., & Ernestus, M. (2008). The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition*, 109(1), 168–173. doi: 10.1016/j.cognition.2008.08.002

Mitterer, H., & Müsseler, J. (2013). Regional accent variation in the shadowing task: Evidence for a loose perception–action coupling in speech. *Attention, Perception, & Psychophysics*, 75(3), 557–575. doi: 10.3758/s13414-012-0407-8

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What does motor efference copy represent? Evidence from speech production. *The Journal of Neuroscience*, 33(41), 16110–16116. doi: 10.1523/JNEUROSCI.2137-13.2013

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, 63(1), 1–33. doi: 10.1016/j.cogpsych.2011.05.001

Oomen, C. C., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes*, 17(2), 163–184. doi: 10.1080/01690960143000010

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4), 2382–2393. doi: 10.1121/1.2178720

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41(3), 427–456. doi: 10.1006/jmla.1999.2653

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, 134(3), 427–459. doi: 10.1037/0033-2909.134.3.427

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02), 169–190. doi: 10.1017/S0140525X04000056

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension?. *Trends in Cognitive Sciences*, 11(3), 105–110. doi: 10.1016/j.tics.2006.12.002

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(04), 329–347. doi: 10.1017/S0140525X12001495

Pickering, M. J., & Garrod, S. (2014). Self-, other-, and joint monitoring using forward models. *Frontiers in Human Neuroscience*, 8, doi: 10.3389/fnhum.2014.00132

Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.). *The Emergence of Language* (pp. 381–415). Mahwah, NJ: Taylor and Francis.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, 77(2), 97–132. doi: 10.1016/S0010-0277(00)00090-1

Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11(5), 351–360. doi: 10.1038/nrn2811

Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, 103(20), 7865–7870. doi: 10.1073/pnas.0509989103

Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29(1), 86–102. doi: 10.1016/0749-596X(90)90011-N

Scott, S. K., McGettigan, C., & Eisner, F. (2009). A little more conversation, a little less action—candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience*, 10(4), 295–302. doi: 10.1038/nrn2603

Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension—an fMRI study. *Cerebral Cortex*, 22(7), 1662–1670. doi: 10.1093/cercor/bhr249

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *Neuroimage*, 25(1), 76–89. doi: 10.1016/j.neuroimage.2004.11.006

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, 17(10), 2387–2399. doi: 10.1093/cercor/bhl147

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26), 10587–10592. doi: 10.1073/pnas.0903616106

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, 60(4), 487–501. doi: 10.1016/j.jml.2009.01.001

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952–981. doi: 10.1080/01690960903498424

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, 39(3), 1429–1443. doi: 10.1016/j.neuroimage.2007.09.054

Twomey, K., Chang, F., & Ambridge, B. (2014). Do as I say, not as I do: A lexical distributional account of English locative verb class acquisition. *Cognitive Psychology*, 73, 41–71. doi: 10.1016/j.cogpsych.2014.05.001

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7(7), 701–702. doi: 10.1038/nn1263

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, 1(6), 209–216. doi: 10.1016/S1364-6613(97)01070-X

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9), 338–347. doi: 10.1016/S1364-6613(98)01221-2

Ylinen, S., Nora, A., Leminen, A., Hakala, T., Huotilainen, M., Shtyrov, Y., & Mäkelä, J. P. (2014). Two distinct auditory-motor circuits for monitoring speech production as revealed by content-specific suppression of auditory cortex. *Cerebral Cortex*, 25(6), 1576–1586. doi: 10.1093/cercor/bht351

Yuen, I., Davis, M. H., Brysbaert, M., & Rastle, K. (2010). Activation of articulatory information in speech perception. *Proceedings of the National Academy of Sciences*, 107(2), 592–597. doi: 10.1073/pnas.0904774107

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, 123(2), 162–185. doi: 10.1037/0033-2909.123.2.162

# Part II Comprehension

# 8 Overview

EVA M. FERNÁNDEZ AND
HELEN SMITH CAIRNS

Queens College and the Graduate Center, City University of New York

Research on language comprehension builds upon a basic model that assumes comprehenders convert the linguistic signal (spoken, signed, or written) into a set of units used to activate lexical material. These recovered lexical representations in turn trigger the retrieval or projection of the syntactic relations among words, relations which, along with lexical semantics, become the information used by the system to recover meaning and integrate it into ongoing discourse processing. It is with such a model in mind that Part 2 of this volume is organized. The sequence of chapters roughly follows the processing levels required for comprehension at the word-, sentence-, and discourse-level, beginning with speech perception, continuing through lexical and morphological processing, and working up to sentence processing (parsing). Subsequent chapters address higher-level semantic and pragmatic processing. Topics in bilingualism are addressed throughout, and a later chapter deals with comprehension in special populations. Part 2 closes with a chapter on neurolinguistics, describing research on the neural underpinnings at every stage of comprehension, and pulling all the various levels together into a comprehensive description of the brain functions that support comprehension. In this overview, we describe the contents of these chapters, relating them to each other and to existing trends in the literature.

## Decoding the signal

Speech perception is the extraordinarily rapid and automatic decoding of a physical acoustic signal into the mental units ultimately used to retrieve words from the lexicon. David Pisoni's chapter opens with discussion of this intrinsic property of speech perception: its robustness in the presence of variability

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

(Pisoni, this volume, p. 193). Pisoni describes early research on speech perception as providing a solid empirical foundation identifying the acoustic attributes of speech and their impact on intelligibility. Contemporary models of speech perception devote more attention to broader problems, and the predominant models of lexical identification rely heavily on memory representations of previously heard words. Variables such as speaker identification, acoustic/phonetic variation, and context—considered obstacles to speech perception in earlier theories—become part of the lexical identification process. Phenomena such as perceptual constancy arise from computational processes at the time of retrieval from memory representations. Pisoni describes the Neighborhood Activation Model (NAM) for lexical recognition, developed in his lab over the past two decades (Luce & Pisoni, 1998). This model posits that words are recognized in terms of their opposition and contrasts with other phonologically similar words. A particularly valuable aspect of Pisoni's chapter is its attention to hearers with cochlear implants, a topic not covered elsewhere in this volume.

Within research on speech perception, a fertile area is how non-native speech perception operates, as well as how the perceptual abilities of second language (L2) learners develop. These issues are the focus of Ocke-Schwen Bohn's chapter. The perception of non-native or L2 speech relies on the mapping of L2 phonetics to those of the first language (L1). Bohn demonstrates the best way to evaluate this mapping and shows how it accounts for both perceptual ease/difficulty and the difficulty/ease of learning L2. Bohn presents in some detail three models that relate to L1/L2 mapping, which have served to formulate a great deal of the research conducted on non-native and L2 speech perception since the 1990s. One is the Perceptual Assimilation Model (PAM) dealing with non-native speech perception (see also Best, this volume) and PAM-L2, its extension to deal with phonological learning of a second language (Best & Tyler, 2007). The third is Flege's Speech Learning Model (SLM; Flege, 1995). Bohn describes perceptual universals that contribute to both perception and learning and argues that studies of L2 phonological acquisition provide powerful evidence regarding the malleability of the human mind. He points out how this evidence base disconfirms proposals about a decrease in plasticity of the brain, proposals that have their origins in Lenneberg's critical period hypothesis (Lenneberg, 1967). The evidence reviewed by Bohn instead indicates that speech perception remains malleable over the life span.

## Processing words

A major theoretical issue in lexical retrieval is the interplay between lexicon-based and learning-based models of word knowledge. Petar Milin, Eva Smolka, and Laurie Feldman offer a comprehensive chapter about the representation and retrieval of morphologically complex words, with a particularly valuable cross-linguistic perspective. They survey research in both the combinatorial tradition and in the learning-based tradition, describing empirical evidence on word recognition for morphologically complex words, encompassing both inflectionally and derivationally complex morphology. Milin and colleagues offer valuable methodological information, with descriptions of the lexical decision task and priming (masked and overt), and evaluative explanations of which aspects of word recognition each responds to. A major distinction between lexicon-based and learning-based models is that the former assume the explicit representation of morphemes in word structure, while the latter assume no explicit representation of morphemes and account for morphological effects by appeal to patterns of form and meaning that are processed simultaneously.

Smolka and colleagues describe dual mechanism and single mechanism accounts of morphological processing. The single mechanism models they discuss include parallel-distributed connectionist models. Their chapter incorporates a presentation of information-theoretic models and the Naive Discrimination Learning model. A section on lexical access discusses how priming techniques can determine when semantic effects arise during lexical access, and how EEG effects match behavioral effects. Further, neighborhood density affects responses to primes.

David Braze and Tao Gong's contribution to this volume begins with the observation that reading has a parasitic relationship with speech (Braze & Gong, this volume, p. 269). Their comprehensive chapter outlines a taxonomy for writing systems, and goes on to describe models of word recognition and reading comprehension. Braze and Gong describe writing systems in which graphemes relate to phonological units or syllables, suggesting implications for word recognition for each. With writing systems as different as those of English and Chinese, at issue is the ease of mapping from symbols to linguistic units, a question encoded into the Orthographic Depth Hypothesis, which proposes that the more transparent and consistent that mapping, the easier a writing system will be to learn and use.

Braze and Gong present a variety of word recognition models, beginning with Morton's logogen model (Morton, 1969). Braze and Gong take the position that all word recognition results in the retrieval of relevant information from the word's lexical representation; after retrieval, psycholinguistic processes associated with sentence and higher level comprehension take over. Turning to the processing of printed text, Braze and Gong introduce the E-Z Reader framework (Reichle, Warren, & McConnell, 2009), which accounts for how eye movements are guided by the interaction of visual processing, attention and lexical processing in the reading of text. As they turn to reading comprehension, Braze and Gong invoke the Simple View of Reading model (Gough & Tunmer, 1996) and neurolinguistic evidence that comprehension processes are the same whether the original modality is visual or acoustic. Thus, differences between the two modalities are identifiable only for initial lexical processing.

Judith Kroll and Fengyang Ma's chapter on the bilingual lexicon explores in great depth the counter-intuitive recent findings that both of a bilingual's languages are continually active even when one language alone is required. They examine this research from the perspective of two models of the bilingual lexicon, the Revised Hierarchical Model (RHM; Kroll & Stewart, 1994) and the Bilingual Interactive Model + (BIA+; Dijkstra & Van Heuven, 2002). (For additional discussion of the RHM, in the context of translation, see Paolieri, Morales, & Bajo, this volume.) Kroll and Ma describe research indicating that the simultaneous activation of both languages is independent of language proficiency and of similarity between the languages.

Kroll and Ma explore research demonstrating simultaneous activation in both comprehension and production, as well as mechanisms to account for the bilingual's ability to inhibit the active yet at the moment not required language. Neurolinguistic studies that employ ERP and MRI techniques have made it possible to explore the time course and the location, respectively, of activation. A large set of variables have been examined in the literature (and are discussed in the chapter), including effects of context, of cognates, of language dominance, and of cross-language lexical neighbors. The common finding is that cross-language interactions are observed relatively early in processing, indicating that they are the result of bottom-up phonetic processes, rather than top-down processes.

Finally, Kroll and Ma address the learning of new words by monolinguals and bilinguals and explore the recurring finding that bilinguals are more proficient word learners than are monolinguals. They conclude that this advantage is specific to a bilingual's language experience and not a general manifestation of enhanced executive function.

## Sentences and discourse

A central component of language comprehension is the processing of individual sentences, which at its core involves determining the hierarchical relations between linearly-ordered words. Perceptual and lexical information is used by the parser—constrained by the grammar—to compute the basic meaning of a sentence. We know that the parser has extra-grammatical strategies that it employs for structure-building, for instance (in English) canonical order and gap-filling strategies. Gap filling, structure building, and referent identification never violate grammatical constraints; however, extra-grammatical considerations guide decisions on which the grammar is silent. For instance, a single meaning of an ambiguous sentence, which is not resolvable by the grammar, will be selected by parsing preferences. The basic sentence meaning is, in turn, operated on by adding inferences and higher-order conceptual and semantic processing to create a richer sentential representation for use in text or conversation. The chapter by Matthew Traxler, Liv Hoversten, and Trevor Brothers addresses all of these basic processes in sentence processing.

A classic problem in psycholinguistics is determining how the comprehender's grammar is employed in this process. Traxler and colleagues address the use of syntax in processing, which they describe as an interface system that mediates between lexical, sentence, and discourse semantics. They also address instances when syntax is marginalized in sentence interpretation, as in dual streams and good-enough parsing.

Traxler and colleagues review neurolinguistic evidence for stages of parsing in monolingual speakers and reactions to semantic and syntactic violations. With this information as a baseline, they explore neurolinguistic information about sentence comprehension in bilinguals. They describe evidence on the effects of age of acquisition of L2 as it may bear upon the critical period hypothesis for second language learning, which points to a complex interaction among proficiency, age of acquisition, and the relationship between L1 and L2 with respect to neurolinguistic measures of comprehension and reactions to structural violations.

The chapter by Janet Nicol and Andrew Barss also deals with comprehension of individual sentences, through the lens of how anaphors and subject-verb agreement are processed. The on-line assignment of referents to reflexives and third person pronouns provides an elegant demonstration of the obedience of the comprehension mechanism to grammatical principles (A and B of Binding Theory) and clausal structure. Anaphor resolution is completely determined by the relationship of the anaphor to its referent within the phrase structure of the sentence (assigned by the parser), and linear order is irrelevant. Subject-verb agreement in comprehension is of interest because of its comparison to agreement in production (see also Franck, this volume). The comprehension of agreement is, like anaphor resolution, determined by the phrase structure of the sentence being processed and depends on the application of grammatical principles during on-line sentence comprehension. Like anaphor resolution, and unlike agreement in production, intervening material does not affect the assignment of subject-verb agreement. This conclusion follows from the accumulated empirical evidence, some with conflicting results, which Nicol and Barss review with great care.

Most early sentence processing research employed techniques involving silent reading. Along with improvements on technologies for reproducing recorded speech we have witnessed tremendous growth in the interest to study how the rhythm and intonation of speech interacts with sentence-level processing, as well as the development of models of sentence processing that integrate both explicit and implicit prosodic considerations. Elizabeth Pratt's chapter begins with an introduction to the features of prosodic structure, from which are derived variations in the rhythm and intonation of speech. Pratt describes research on how the speech stream is parsed, and on how prosody contributes to the recovery of syntactic structure, research that has drawn significantly from ambiguous strings that can be disambiguated prosodically (e.g., *Old men and women sat on the bench*; Lehiste, 1973). Pratt's chapter also offers an overview of investigations on how prosody interacts with memory and reading, including silent reading and reading in a second language. Some of this research links oral reading fluency to comprehension (e.g., Good & Jefferson, 1983; Anema, 2008); other branches of this research have examined the role that implicit prosody plays in comprehension (Frazier & Gibson, 2016).

Petra Schumacher's chapter on semantic-pragmatic processing takes us beyond the basic meaning of individual sentences. While anaphor resolution is a completely form-based process in individual sentences, reference assignment across sentences is a different sort of operation. Reference for pronouns and definite noun phrases depends upon a number of factors, including information status (given or new), referential prominence, and whether the referent must be inferred. Reference location varies in processing complexity, which can be measured by both behavioral and neurolinguistic techniques. Speakers select referential forms and prosodic cues to ease the processing load of their interlocutors. Addressing "speaker meaning" as distinct from "sentence meaning," Shumacher shows that the ability to make inferences about the speaker's meaning is a critical aspect of determining speaker intent. She shows how Gricean implicatures play a role in inferring speaker meaning. She also discusses instances of meaning transfer (metonymy), which are highly context sensitive. Complement coercion is another kind of communicative device that involves the intersection of selectional restrictions and aspect of the verb and real world knowledge. Gricean principles also come into play. All in all, comprehension of strings of sentences require grammatical and lexical knowledge, plus a great deal of extra-linguistic cognitive processing.

## Putting it all together

Our understanding of comprehension processes would not be complete without some notion of what happens when things go wrong. Jet Vonk, Eve Higby, and Loraine Obler discuss comprehension impairments and their origins, in healthy aging, aphasia, and dementia. In healthy aging there is not a problem with linguistic architecture, but the sensory and cognitive underpinnings of comprehension may account for comprehension deficits. For instance, executive function and working memory are known to decline with age. Automatic processing shows no age-related decline, but controlled processing does. Other abilities may be unimpaired, but are slowed. Some studies show impairment in tasks involving complex syntax and ambiguity.

While comprehension was once believed to have been spared in Broca's (agrammatic) aphasia, research shows that this is the case only if real world knowledge can provide sentence meaning. Required to rely solely on grammatical structure, agrammatic aphasics perform at chance levels, while still appearing to perform better than chance when asked to make grammaticality judgments. Difficulty with language comprehension has typically been the hallmark of Wernicke's aphasia. While people with this type of aphasia have a number of comprehension-related deficits, they differ depending upon the exact site of the lesion. Finally, linguistic deficits in the dementias are probably primarily the result of underlying cognitive deficits. Since dementia is progressive, deficits will shift depending on what stage of the disease is encountered.

Neurolinguistic data are presented throughout this volume, but the chapter by Michael Skeide and Angela Friederici is the only one exclusively devoted to the neurolinguistic organization and processing of language. An important theoretical question about language processing has been its modularity, in particular if some processors are informationally encapsulated. Neurolinguistic studies of sentence comprehension suggest that there are aspects of processing that are distinct both temporally and spatially in the brain. Brain areas dedicated to language are distinct from domain-general areas dealing with general cognition.

Skeide and Friederici explore the neural basis of many aspects of sentence processing that psycholinguists have identified over the years with behavioral methods, for instance, syntactic versus semantic processing and bottom-up versus top-down processing. Models of neurolinguistic sentence comprehension trace operations from perception of phonological elements through lexical processing, lower- and higher-level syntactic processing, and high-level semantic integration. Skeide and Friederici identify the brain regions known to be associated with each process, as well as their temporal relationships. Finally, they show how neurolinguistics identifies the two characteristics of language that have been hypothesized to confer upon it an overwhelming evolutionary advantage: language is an efficient medium for communication as well as a unique tool for the conceptual representation of the world.

## Summing up

The chapters comprising Part 2 of this volume do not exhaustively cover research on language comprehension, but they do address foundational processes in the comprehension of speech and writing, in the retrieval of words, in the recovery of syntactic information, and in the assignment of semantic meaning within and between sentences in discourse. Comprehension has always been a particularly prolific area in the study of language processing, perhaps because it provides straightforward ways to test for how the various components of knowledge of language are put to use in real time. Comprehension also offers windows into the organization of the linguistic cognitive architecture, a matter addressed by the research from multiple perspectives, including bilingualism, normal aging, and neurocognition.

## REFERENCES

Anema, Inge. (2008). *The Relationship between Fluency-Based Suprasegmentals and Comprehension in Oral and Silent Reading in Dutch Speakers of English*. Doctoral dissertation, City University of New York.

Best, C. T. & Tyler, M. D. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5, 175–197. doi: 10.1017/S1366728902003012

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and*

*Linguistic Experience: Issues in Cross-language Research* (pp. 233–277). Timonium, MD: York Press.

Frazier, L. & Gibson, E. (Eds.) (2016). *Explicit and implicit prosody in sentence processing*. New York, NY: Springer.

Good, R. H., & Jefferson, G. (1998). Contemporary perspectives on Curriculum-Based Measurement validity. In M. R. Shinn (Ed.), *Advanced applications of Curriculum-Based Measurement* (pp. 61–88). New York: Guilford.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 1–13). Mahwah, NJ: Lawrence Erlbaum Assoc

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilinguals memory representations. *Journal of Memory and Language*, 33, 149–174. doi: 10.1006/jmla.1994.1008

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa*, 7, 107–122.

Lenneberg, E. (1967). *Biological foundations of language*. New York, NY: Wiley.

Luce, P. A. & Pisoni, D. B. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear & Hearing*, 19, 1–36.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76(2), 165–178. doi: 10.1037/h0027366

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychon Bull Rev*, 16(1), 1–21.

# 9 Speech Perception: Research, Theory, and Clinical Application

DAVID B. PISONI

Indiana University

## Introduction

The most important property of human speech perception is its robustness in the face of diverse acoustic stimulation and degradation in the sensory properties of the acoustic signal. Speech perception is without doubt one of the most flexible and highly adaptive information processing skills that humans have developed in the course of evolution. Human listeners adapt and compensate very quickly and effortlessly to large acoustic changes in the vocal sound source and numerous sources of acoustic degradation in the speech signal without any significant loss of speech intelligibility. Moreover, human listeners are able to recognize and successfully understand speech from unfamiliar talkers, non-native speakers, as well as speech synthesis-by-rule systems, under a wide range of adverse and challenging conditions.

In this chapter, I first provide a brief summary of the historical foundations of speech perception going back to the 1940s during World War II. Then I move to the modern era of speech science and consider the set of core theoretical assumptions that framed the traditional approach to speech perception. Next I describe several new empirical and theoretical developments that have had a significant influence on current thinking, research and clinical application. These new developments have been quite dramatic in nature and have led to numerous fundamental changes in the direction of research on speech perception and a substantial reconceptualization of the principle theoretical issues that defined the field from the early 1950s. Finally, I briefly review some recent clinical findings on speech perception in prelingually deaf children who have received cochlear implants as a medical treatment for profound deafness. Research on this unique clinical

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

population, the modern-day equivalent of the so-called "forbidden experiment" in language acquisition, has provided additional new insights about the basic underlying sensory and neurocognitive processes used in speech perception in normal-hearing listeners and has raised a host of new questions about the interactions between sensory coding of highly-degraded underspecified speech signals at the auditory periphery and the role of cognition, learning and memory in speech perception and language acquisition.

## Historical foundations of speech perception

The field of human speech perception has a long history going back to the end of the nineteenth century when hearing scientists first began to use electrically recorded audio signals to assess hearing loss (Fletcher, 1953; Miller, 1951). Most of what we currently know about the basic acoustical and sensory foundations of speech perception comes from the pioneering research carried out by telephone engineers at Bell Telephone Laboratories (Flanagan, 1965; Fletcher, 1953). This extensive body of applied research on speech and the assessment of telephone communication equipment established the necessary and sufficient acoustical conditions needed for effective and highly reliable speech transmission and reception over conventional telephone circuits and provided an enormous body of empirical data on human hearing and speech communication (Flanagan, 1965).

All of the quantitative experimental methods used today for assessing speech intelligibility can be traced directly back to the early studies carried out at Bell Labs. These studies were designed to evaluate the effectiveness of various components used to transmit speech over the telephone (Fletcher, 1953). The primary focus of this research program was on speech intelligibility—the recognition of speech sounds, isolated words, and short meaningful sentences (Allen, 1994). Further pioneering research on the acoustics of speech, especially speech communication in noise, was carried out at the Psycho-Acoustic Laboratory at Harvard University during World War II. Summaries of the most important research findings were written by Hudgins *et al.*, (1947) and Rosenzweig and Stone (1948). A comprehensive review chapter was also published after the war by Licklider and Miller (1951).

Although the two applied research programs on hearing and speech communication at Bell Laboratories and Harvard University provided much of the basic foundational knowledge that we have about hearing and human speech communication, almost all of these investigations were primarily empirical in nature focusing research almost exclusively on practical real-world telephone and military-related communications problems. One of the general principles that emerged from this research was the discovery that the most important acoustical attributes of speech intelligibility were represented by changes in the time-frequency-intensity pattern of the speech signal and how these spectral relations change over time (Licklider & Miller, 1951). Little, if any, effort was devoted to broader theoretical and conceptual issues in speech perception or spoken language processing.

After World War II ended, speech and hearing scientists, acoustical engineers and experimental psychologists began to focus their efforts more broadly on the human listener as a central component of the speech communication process (Miller, 1951; Hirsh, 1952). Scientists at Bell Labs and Harvard continued their applied research started during the war. Other programs on speech pursued new lines of basic research investigating the fundamental nature of speech and speech perception (Borden *et al.*, 2011). Efforts were also begun at this time to develop methods for speech synthesis-by-rule that could be used in automated reading machines for blind veterans returning from the war (see Allen, Hunnicutt & Klatt, 1987; Cooper, Delattre, Liberman, Borst, & Gerstman, 1952).

## Traditional approach to speech perception: Speech cues and phonetic perception

The early studies on the acoustics of speech and the theoretical efforts at developing quantitative models of speech production at MIT along with the behavioral experiments on the perception of speech cues at Haskins Laboratories were directly responsible for identifying a small number of core theoretical problems in the field of human speech perception and spoken language processing (see Liberman, 1996; Pisoni, 1978; Stevens & House, 1972; Studdert-Kennedy, 1976). A critical review and theoretical discussion of the foundational problems in speech perception and word recognition was undertaken by Klatt (1979). Other more recent reviews of the field of can be found in chapters in Pisoni and Remez (2005) and Gaskell (2007).

The field of human speech perception has changed quite dramatically over the last 40 years since the publication of Michael Studdert-Kennedy's seminal review chapter on speech perception in 1976. Studdert-Kennedy reviewed a large body of literature on the perception of speech sounds and proposed an integrated conceptual framework that motivated an enormous number of novel experimental studies on speech perception in the late 1970s and early 1980s. Studdert-Kennedy attempted to provide a unifying theoretical framework to bring a seemingly diverse set of empirical findings together in one place. One centerpiece was Studdert-Kennedy's detailed discussion of three foundational issues about the nature of speech and the speech perception process—linearity, acoustic-phonetic invariance, and segmentation. These three issues have remained major challenges in the field for more than 50 years.

## Linearity, invariance, and segmentation

Chomsky and Miller (1963) argued that one of the most important problems in speech perception is that the speech signal fails to meet the conditions of linearity and acoustic-phonetic invariance. As a consequence, the basic recognition problem in speech perception becomes very challenging for both humans and machines. The linearity condition assumes that for each phoneme there must be a particular stretch of sound in the utterance, and if phoneme X is to the left of phoneme Y in the phonemic representation, the stretch of sound associated with X must precede the stretch of sound associated with Y in the physical signal. The acoustic-phonetic invariance condition assumes that for each phoneme X, there must be a specific set of criterial acoustic attributes or "defining features" associated with phoneme X in all contexts. These acoustic attributes must be present whenever X or some variant of X occurs, and they must be absent whenever some other phoneme occurs in the representation.

Acoustic-phonetic research since the early 1950s has shown that it is difficult, if not impossible, to reliably identify acoustic attributes that match the perceived phonemes in an utterance independently of the surrounding context (Fant, 1962; Liberman & Studdert-Kennedy, 1977). The massive coarticulation effects present in speech production create a great deal of context-conditioned acoustic variability in the speech signal. Often a single acoustic segment contains information about several neighboring linguistic units; and, conversely, the same linguistic unit is often represented acoustically in quite different ways, depending on the surrounding phonetic context, speaking rate, talker and regional dialect (Liberman *et al*., 1967). The context-conditioned variability between acoustic signal and phoneme resulting from coarticulation also poses enormous problems for the segmentation of speech into context-free units. Because of the failure to meet the linearity and acoustic-phonetic invariance conditions, it has been difficult to reliably identify boundaries between adjacent words, syllables and phonemes in the speech waveform. Although some gross segmentation is possible according to strictly acoustic criteria (see Fant, 1962), the number of acoustic segments that can be identified in the speech waveform is typically greater than the number of phonemes in the utterance and, moreover, no simple first-order invariant mapping has been found between these acoustic attributes and perceived phonemes in the linguistic message.

The lack of acoustic-phonetic invariance and segmentation problems identified by Chomsky and Miller suggested something unique about speech as an acoustic signal (Liberman, 1996). Relations between segments of the acoustic signal and units of linguistic analysis are very complex and this, in turn, places strong constraints on the types of perceptual theories that might be proposed for speech perception. For example, filter or template-matching theories are generally considered as poor candidates for human or machine speech recognition because linguistic segments cannot be defined exclusively by simple acoustic attributes in the speech signal. As Chomsky and Miller (1963) observed, if both the acoustic-phonetic invariance and linearity conditions were met, the task of building machines capable of automatically recognizing segmental phonemes in human speech would be greatly simplified (Moore, 2007a,b). It would just be a matter of arranging an appropriate set of filters in a network. Although numerous attempts have been made following these suggestions in the past, the results have been uniformly unsuccessful because of the inherent variability in the acoustic signal (Lindgren, 1967). Passive theories of speech recognition involving template matching and filtering are no longer considered as potential models of human speech perception (see Klatt, 1979).

Until Michael Studdert-Kennedy's review chapter, research on speech perception was "intellectually isolated" from several closely related allied fields such as cognitive psychology, cognitive neuroscience, and computer science. As a result, almost all of the theoretical and conceptual developments in the field relied heavily on formalist linguistic descriptions of speech. One of the core assumptions of linguistic analysis is that the process of human speech perception results in the construction of a sequence of abstract, idealized, context-free symbols such as phonemes arranged in a strictly linear order like beads on a string. A nice description of the traditional abstractionist linguistic view of speech is summarized in Hockett's (1955) well-known Easter Egg Analogy:

> Imagine a row of Easter eggs carried along a moving belt; the eggs are of various sizes, and variously colored, but not boiled. At a certain point, the belt carries the row of eggs between the two rollers of a wringer, which quite effectively smash them and rub them more or less into each other. The flow of eggs before the wringer represents the series of impulses from the phoneme source; the mess that emerges from the wringer represents the output of the speech transmitter. At a subsequent point, we have an inspector whose task it is to examine the passing mess and decide, on the basis of the broken and unbroken yolks, the variously spread out albumen, and the variously colored bits of shell, the nature of the flow of eggs which previously arrived at the wringer [p. 210].

Although this is a good characterization of the acoustical structure of speech, our understanding of the relations between segmental phonemes in the symbolic linguistic message and the acoustic correlates of abstract linguistic units and lexical contrast has not progressed very much over the last 50 years. This narrow segmental "symbol-processing" view of the process of human speech perception has been reconsidered recently as more empirical studies demonstrate that the original principles proposed by Chomsky and Miller (1963) may have been too superficial given what we now know about how the peripheral and central auditory system and brain encodes and stores complex time-varying signals like speech (Hickok & Poeppel, 2007; Obleser *et al*., 2007; Scott & Wise, 2003).

Many recent studies have shown that acoustic variability and fine phonetic and indexical details of speech are encoded by listeners and affect early perceptual processing, word recognition, lexical access and spoken language comprehension (Klatt, 1986). Understanding how human listeners compensate and adapt to different sources of variability in speech, how they encode, store and process the fine acoustic-phonetic details of speech signals and how they manage to reliably recover the talker's intended linguistic message under a wide variety of listening conditions has now become one of the major challenges in the field (see Pisoni & Levi, 2007).

Investigating how listeners encode and process speech variability is the new "holy grail" of the field. Research on human speech perception has dramatically changed its emphasis from a narrow focus on the study of speech cues and segmental phonetic perception to a much broader set of issues related to spoken word recognition and comprehension. Speech variability is no longer viewed as an undesirable source of noise in the acoustic signal that needs to be reduced or eliminated. If anything, there are now strong reasons for believing that variability in speech actually serves a very important function not only in early word learning and language development but also in speech in noise and other capacity demanding information processing operations such as selective attention, working memory and episodic memory, processes that all require the storage, maintenance and manipulation of robust highly-detailed lexical representations of spoken words (Pichora-Fuller *et al.*, 1995; Nahum *et al.*, 2008).

This is a very exciting time to be working in the field of speech perception. Many of the long-standing traditional beliefs and fundamental assumptions about speech and spoken language processing are currently being reexamined and revised in light of new empirical findings and theoretical developments in closely related disciplines, such as cognitive psychology and cognitive science (Moore, 2007a,b). Until fairly recently, the field of speech perception was also significantly delayed in reaping the benefits from new research findings and theoretical developments in neurobiology and neuroscience (Hickok & Poeppel, 2007). Speech scientists have always relied very heavily on theoretical concepts and dogma embodied in conventional linguistic analyses and descriptions of the structure and function of speech and spoken language thereby isolating themselves from researchers in other fields who were working on very similar problems (Lindgren, 1967).

## Fundamental problems in speech perception

The fundamental problem in speech perception is to understand how the human listener recovers the talker's intended linguistic message from information encoded in the time-varying acoustic signal (Moore, 2005). This general problem has been traditionally broken down into a smaller set of more specific questions such as: What stages of perceptual analysis intervene between presentation of the speech signal and recognition of the talker's message? What types of cognitive and linguistic processing operations occur at each stage? What are the primary processing units in speech perception and spoken word recognition and what is the nature, content, and specificity of the neural representations of speech in the auditory system? Finally, what sensory, perceptual, and neural mechanisms are used in speech perception and spoken language processing? These traditional questions are now being addressed in new ways that raise additional issues about the nature of phonetic, phonological, and lexical knowledge, encoding, and processing of acoustic variability by the brain, perceptual learning, neural plasticity, and adaptation. Other issues focus on the study of individual differences, a topic we will touch on briefly in the final section.

## Self-organization in speech perception

Over the last few years, speech scientists realized that many of the fundamental properties of speech that are responsible for its perceptual robustness such as speed and fluency, automaticity, perceptual learning, adaptation, and "graceful" error recovery are not unique to speech or spoken language processing but reflect more general properties shared by other self-organizing systems in physics, biology, and neuroscience (Grossberg & Myers, 2000; McNellis & Blumstein, 2001; Sporns, 1998, 2003). One novel approach to speech perception and spoken word recognition has also emerged recently from independent developments in categorization in cognitive psychology (Nosofsky, 1986; Kruschke, 1992) and frequency-based phonology in linguistics (Pierrehumbert, 2001; Bybee, 2001). These two perspectives offer fresh ideas and new insights into how to solve many of the foundational problems in speech perception related to linearity, invariance, and segmentation (see Nygaard & Pisoni, 1995; Pisoni & Luce, 1986).

## Variability and perceptual constancy in speech perception

One of the major methodological developments has been a conscious and deliberate effort to directly study variability in speech perception (Pisoni, 1997). We have been investigating variability in speech from different talkers, speaking rates, and speaking modes to determine how these changes in the acoustic signal affect speech perception and spoken word recognition (Pisoni & Levi, 2007). Our findings suggest that many long-standing theoretical assumptions about the basic perceptual units in speech perception such as features, phonemes, and syllables need to be substantially revised. In particular, the long-standing foundational assumption that speech perception depends on the recognition of abstract, idealized, context-free symbolic representations encouraged research that was designed to identify simple first-order acoustic invariants and simply ignored problems of acoustic-phonetic and indexical variability.

The conventional "abstractionist" or "analytic" approach to speech perception treated variability as a source of "noise" in the speech signal that needed to be eliminated to uncover the "hidden" abstract underlying symbolic linguistic message (Elman & McClelland, 1986). During the early days of speech research in the 1950s and 1960s, many factors known to produce variability in speech were deliberately eliminated in traditional phonetic perception experiments. To take one example, almost all of the basic research findings in the field of speech perception have used a very small number of talkers—often a single male talker producing "lab speech" under highly controlled conditions in the laboratory (see Port, 2007). Research on speech perception using multiple talkers from different geographical regions speaking non-standard dialects was rarely carried out until only a few years ago (see Clopper & Pisoni, 2004a,b). These new studies on talker variability and regional dialect perception have shown that properties of the vocal sound source—the speaker's voice, are closely linked to the perception of spoken words (Nygaard, Sommers, & Pisoni, 1994; also see Pardo, this volume, for a review of research on the complex interplay of speech perception and speech production in conversation).

## Symbolic versus nonanalytic approaches to speech perception

Our research on speech variability provides support for an alternative view of the speech perception process that is also compatible with a large and growing body of research and theory in cognitive science that deals with "exemplar" or "episodic" models of categorization and "multiple-trace" models of memory (Hintzman, 1986; Kruschke, 1992; Nosofsky, 1986). The non-analytic approach to problems in cognition emphasizes the encoding of specific instances of perceptual events. In contrast with the conventional symbol-processing approach to speech, this alternative view assumes that speech perception and spoken word recognition make use of highly detailed information in the speech signal about the content of the linguistic message as well as the episodic context and that both sources of information are encoded, processed, and stored by the listener, becoming part of a very rich and highly detailed memory representation of speech in long-term memory (Goldinger, 1998).

A critical foundational assumption underlying the nonanalytic approach to speech perception and spoken word recognition is that variability in the speech signal is "lawful" and "informative" to the listener (Elman & McClelland, 1986; Pisoni, 1997). According to this account, listeners encode and store "particulars"—specific instances of events and associated episodic contexts, rather than generalities, abstractions, or idealized prototypes (Nosofsky, 1986; Kruschke, 1992). Abstraction and perceptual constancy occurs but they emerge from "computational processes" at the time of retrieval from very rich and highly detailed memory representations rather than at the time of the initial perceptual analysis.

A large number of studies carried out in our lab over the last 25 years have demonstrated that "indexical" information about a talker's voice as well as very detailed information about regional dialect, speaking rate, speaking style, and other episodic properties of speech are encoded and processed and become part of the long-term representational knowledge base that a listener has about the words of his/her language (Pisoni, 1997). Our research and the work of others has shown that the human perceptual system encodes and retains very fine episodic details of speech. The encoding and processing of variability in speech not only play a significant role in challenging listening environments such as the perception of speech in noise or multi-talker babble by normal hearing listeners, but disturbances in the encoding and processing of speech variability and episodic context also underlie the difficulties observed in hearing-impaired listeners who use hearing aids and cochlear implants, which provide highly degraded underspecified representations of the early sensory properties of speech (Nahum *et al.*, 2008).

## Spoken word recognition and lexical neighborhoods

One proposed solution to the traditional problems in speech perception has been to reframe the long-standing acoustic-phonetic invariance issue by proposing that the primary function of speech perception is word recognition and lexical access rather than phonetic perception. In collaboration with Paul Luce, we developed the Neighborhood Activation Model (NAM) of spoken word recognition that approached the acoustic-phonetic invariance problem directly by assuming that a listener recognizes spoken words "relationally" in terms of oppositions and contrasts with other phonologically similar words in memory (Luce & Pisoni, 1998). Thus, word recognition and lexical selection are carried out by processes that involve discrimination among potential lexical candidates in memory rather than by the conventional bottom-up approach, which assumes that listeners first recognize the individual component phonemes of words and then carry out a search process to locate an abstract idealized representation of a word form in their mental lexicon (see Frederiksen, 1971; Forster, 1976; Morton, 1979).

This approach to speech perception avoids the long-standing problem of having to explicitly recognize individual component phonemes of words directly by locating invariant acoustic-phonetic properties in the speech waveform. As Klatt suggested more than 25 years ago, if speech perception is viewed as the recognition of "lexical candidates" from acoustic information encoded in the speech waveform then there is no need to assume any "intermediate" segmental representations of the input signal. Klatt argued further that the construction of intermediate phonetic representations of speech actually discards detailed acoustic information in the speech waveform that could be potentially useful in word recognition and lexical access. By retaining information for errorful recovery and delaying a final representational commitment until the discarded acoustic information in the signal is no longer needed, the speech perception process becomes extremely robust to large acoustic changes in the signal (see Klatt, 1979).

Our work on the role of the mental lexicon in speech perception began with a series of novel computational analyses of the sound patterns of words using phonetic transcriptions obtained from a computer-readable dictionary (Pisoni *et al.*, 1985). These analyses revealed that spoken words could be organized into "similarity spaces" or "lexical neighborhoods" based on simple metrics of phonological similarity (Treisman,1978a,b). Additional analyses showed that words could be organized by frequency and density within a specific lexical neighborhood and that activation and competition among phonetically similar lexical candidates was an important elementary foundational component of the process of recognizing spoken words in isolation and in sentences (Bell & Wilson, 2001).

Behavioral studies with normal hearing listeners demonstrated that word frequency, lexical density, and neighborhood frequency all affect spoken word recognition performance (Luce & Pisoni, 1998). In NAM, words are assumed to be recognized "relationally" in the context of other phonetically similar words in lexical memory. The speed and accuracy of recognizing a particular word can be accounted for by both the density and frequency of the lexical neighbors activated in memory by the target word. Over the last 20 years, NAM has provided a very powerful theoretical framework for a wide variety of novel studies on the role of the lexicon in speech perception. Consistent with several other current models, NAM assumes that spoken words are organized into similarity spaces in lexical memory and are recognized by processes of activation and competition rather than search or sophisticated guessing strategies (Gaskell & Marslen-Wilson, 2002).

## Indexical and linguistic channels in speech perception

Speech is a biologically significant acoustic signal that simultaneously encodes two parallel channels of information generated by a human talker (Pisoni, 1997). On the one hand, the time-varying speech waveform carries the talker's intended symbolic linguistic message—the words, sentences, and prosodic patterns of a natural language. On the other hand, the speech waveform also encodes and transmits reliable information about the vocal sound source—the talker's physical, social, and mental states. This parallel "complementary channel" of information in speech provides important attributes about the talker's voice quality, which is an additional source of episodic context that the listener encodes and uses to recognize speech, especially under degraded and impoverished listening conditions (Creelman, 1957). Evidence from a long series of studies in our lab going back to the early 1980s has shown that normal-hearing listeners automatically and unconsciously encode, store, and process information about the talker's voice in parallel with the linguistic message (Pisoni, 1997).

The linguistic and indexical channels of speech are encoded in parallel in the speech waveform and are inseparable in both speech production and perception although this "duplex" nature of speech has not been widely recognized or fully appreciated by most speech and hearing scientists. Most speech scientists know about the symbolic linguistic properties of speech—the well-known "speech cues" that underlie and support segmental phonetic perception and speech intelligibility. Until recently, however, little was known about the contribution of the indexical properties of speech and the role these complementary attributes play in speech perception and spoken word recognition (Van Lancker & Kreiman, 1987).

The major indexical properties of speech include the talker's: gender, regional dialect, speaking rate, speaking mode, physical and emotional states, age, height, weight and other idiosyncratic features (Abercrombie, 1967). In past theoretical accounts of speech perception, theorists maintained a strict dissociation between the *linguistic properties* of speech that carry the speaker's intended message and the *indexical attributes* that provide information about the talker's voice (Studdert-Kennedy, 1974; Kreiman & Sidtis, 2011). The dissociation between the form and content of speech has a long history in descriptive linguistics and phonetics, which has been carried over to theoretical accounts of speech perception despite the fact that both channels of information are encoded and carried simultaneously by the same acoustic waveform.

An example of the functional parallelism of the indexical and linguistic attributes of speech is shown in Figure 9.1 from Hirahara and Kato (1992) which displays the encoding of speech at the auditory periphery. The *absolute*

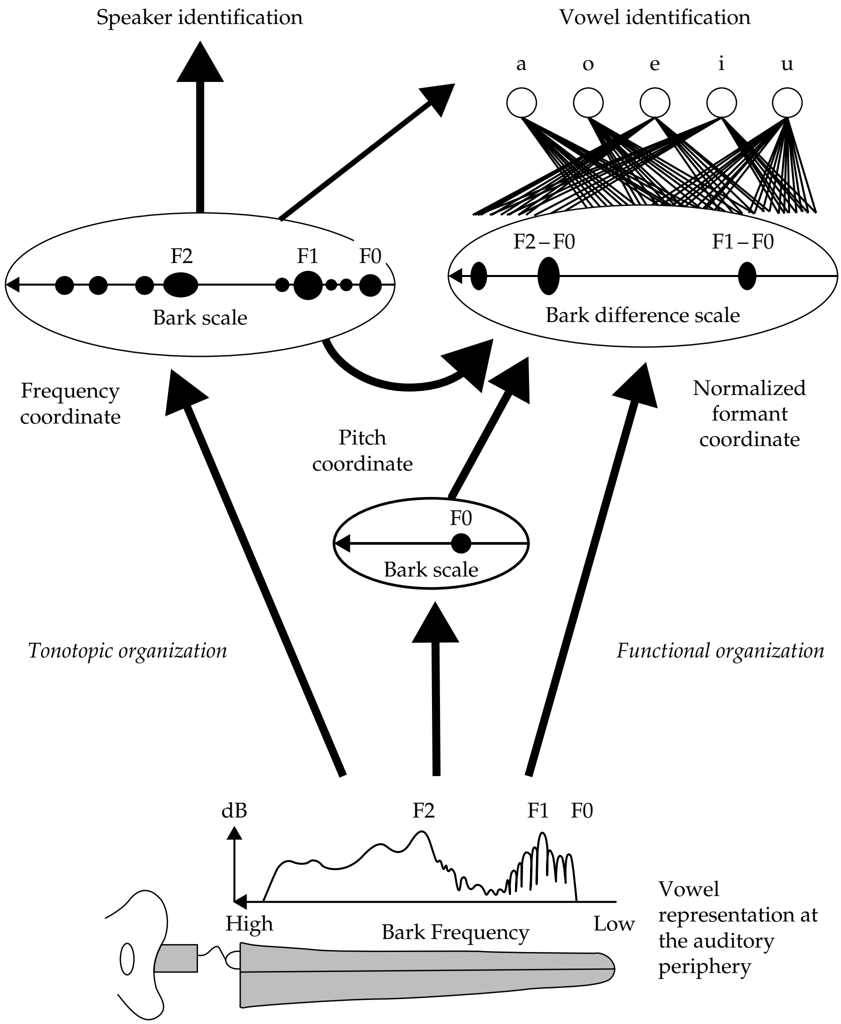



**Figure 9.1** Vowel projections at the auditory periphery reveal that information for speaker identification (Panel A) and perception of vowel quality (Panel B) is carried simultaneously and in parallel by the same acoustic signal. The tonotopic organization of the absolute frequencies using a bark scale provides reliable cues to speaker identification, whereas the relations among the formant (F1, F2 and F3) patterns in terms of difference from F0 in barks provide reliable cues to vowel identification. (Reprinted from Hirahara & Kato, The Effect of F0 on Vowel Identification in Speech Perception, Production and Linguistic Structure, p. 109, Copyright 1992, with permission from IOS Press).

*frequencies* of the vowel formants shown by peaks in the spectrum envelope provide important cues to speaker identification (Panel A), whereas the *relative differences* among the formants specify information for vowel identification (Panel B). Both sets of attributes are carried in parallel by the same acoustic signal and both channels of information are automatically encoded by the peripheral and central auditory mechanisms used in recognizing speech and understanding spoken language.

Numerous recent studies have suggested that the duplex property of speech is one of the core "defining features" of the speech signal that is responsible for the robustness of human speech perception in noise and under adverse listening conditions. While this may seem at first glance to be a novel and unconventional view of speech perception, I have suggested that any principled theoretically motivated account of human speech perception and spoken language processing will also have to be compatible with what we currently know about human information processing and how human memory and learning works, especially perceptual learning and episodic memory (Pisoni, 2000). Speech perception and spoken language processing are not carried out in a vacuum isolated from the rest of cognition and the brain; the core sensory and perceptual processes and cognitive infrastructure used in speech perception and spoken word recognition are inseparable from memory, learning, and cognitive control processes that reflect the operation of many separate neurobiological components working together as an integrated system.

## Research on deaf children with cochlear implants: Inferring function from dysfunction

We have been studying speech perception in prelingually deaf children who have received cochlear implants as a medical intervention for a profound sensorineural hearing loss (Pisoni, 2005). Research on this clinical population has provided new knowledge about perceptual learning, neural plasticity and the development of speech perception (Pisoni *et al.*, 2000). Although cochlear implants work well for many profoundly deaf children, they do not always provide optimal benefits to all deaf children who receive them. Some children do extremely well following implantation and display age-appropriate speech and language skills on a battery of traditional clinical tests administered under quiet listening conditions. In contrast, many other deaf children often struggle for long periods of time after they receive their cochlear implants and frequently fail to achieve comparable levels of speech and language performance. If we can identify the reasons why good cochlear implants users are doing so well, we should be able to use these findings to help low-performing children improve their speech and language skills and reach their potential to derive optimal benefits from their cochlear implants. We have also been studying individual differences in these children to assess the "representational specificity" of their phonological and lexical representations of speech, which are often weak and coarsely coded.

## What is a cochlear implant and how does it work?

A cochlear implant is a surgically implanted electronic device that functions as an auditory prosthesis for hearing-impaired children and adults who have severe-to-profound sensorineural hearing loss (see Figure 9.2). Cochlear implants provide direct electrical stimulation to the surviving spiral ganglion cells of the auditory nerve bypassing the damaged hair cells of the inner ear to restore the sense of hearing and provide access to sound and acoustic stimulation to higher centers of the brain. While cochlear implants provide many deaf adults and children with access to sound, they do not restore all aspects of normal hearing. The temporal fine structure of complex acoustic signals like speech and music are poorly encoded by the current generation of processing strategies used in cochlear implants.

All cochlear implants have two components: an internal receiver that is connected to an electrode array, which is inserted in the cochlea, and an external signal processing unit that is located behind the external ear that transmits a radio signal to the internal receiver and electrodes. The external processing unit consists of a microphone that picks up sound from the environment and a signal processor

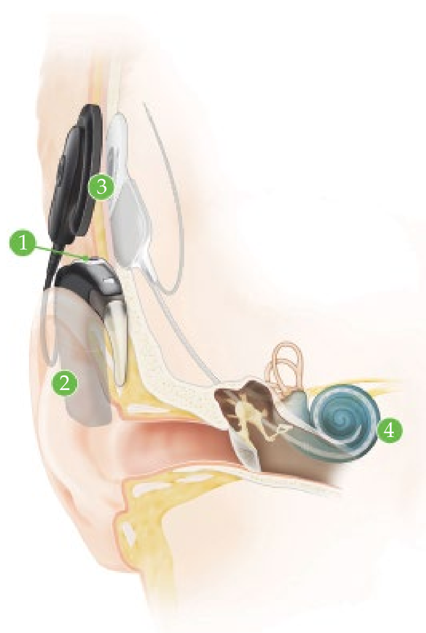



**Figure 9.2** Internal and external components of a multichannel cochlear implant system, consisting of a microphone (1) that picks up sound from the environment, a speech processor (2) that converts sound into electrical signals, a surgically implanted receiver (3), and an internal electrode array inserted in the cochlea (4) that sends electrical pulses to each electrode tonotopically. (Image courtesy of Cochlear Americas, © 2016)

that codes frequency, amplitude, and time and compresses the signal to match the dynamic range of the ear. Cochlear implants do a good job at encoding temporal and amplitude information in speech (Shannon *et al.*, 1995). Depending on the manufacturer, different place coding schemes are used to encode, represent, and transmit frequency information.

For postlingually profoundly deaf adults, a cochlear implant provides a transformed spectrally degraded electrical signal to an already fully developed auditory system and mature language processing system. In the case of a congenitally deaf child, however, a cochlear implant provides novel electrical stimulation through the auditory sensory modality. The electrical stimulation received from a cochlear implants provides access to sound and an opportunity to perceive speech and develop spoken language for the first time after a period of auditory deprivation often beginning prenatally. Although their brain and nervous system continue to develop in the absence of normal auditory input and stimulation, numerous studies have shown that substantial cortical reorganization has already taken place during the early period of sensory deprivation even before implantation (Kral, 2013). As a result, many elementary speech and language processing skills are delayed and often develop in an "atypical" manner after implantation.

## Age of implantation

Age of Implantation is one of the major demographic variables that has been found to be associated with almost all conventional outcome measures. Children who receive cochlear implants at a young age, often under three years, consistently perform better on a wide range of conventional behavioral speech and language measures than children who are implanted at older ages (Kirk, Pisoni, & Miyamoto, 2000; Niparko *et al.*, 2010). Duration of deafness prior to implantation is also related to outcome and benefit. Children who have been deaf for shorter periods of time before implantation do much better on a variety of clinical measures than children who have been deaf for longer periods of time. Both findings, age of implantation and duration of deafness, demonstrate the contribution of neural plasticity and the role of sensitive periods in sensory, perceptual and linguistic development (Konishi, 1985; Marler & Peters, 1988).

## Experience- and activity-dependent learning

The nature of the early social and linguistic environment that a deaf child experiences following implantation has also been found to affect performance on a range of speech and language outcome measures. Children who are immersed in "auditory-oral" language-learning environments that emphasize the use of oral communication and the development of speaking and listening skills consistently do better on a wide range of assessments of speech perception, word recognition, and spoken language comprehension than children placed in "total" communication environments who are exposed to Signed Exact English (Geers, Brenner, &

Davidson, 2003). Total communication children typically do more poorly on receptive and expressive language tests that rely on highly automatized rapid phonological coding skills than children who are educated using auditory-oral methods (Geers *et al.*, 2011). Very few deaf children with cochlear implants learn American Sign Language along with spoken language. Children who use auditory-oral communication modes also produce more intelligible speech and show significant gains in expressive vocabulary and other measures of spoken language processing (Tobey *et al.*, 2003).

Our recent findings also suggest that many deaf children with cochlear implants have comorbid disturbances and delays in several elementary cognitive processes. A period of auditory deprivation during early development followed by exposure to highly degraded underspecified acoustic information from a cochlear implant affects cognitive and linguistic development in a variety of ways. Differences resulting from both prelingual deafness and subsequent neural reorganization of multiple brain systems may be responsible for the enormous variability observed in speech and language outcomes following implantation.

## Theoretical implications of research on deaf children with cochlear implants

Research on cochlear implants also has several important implications for research and theory in cognitive and linguistic development. Above and beyond the immediate clinical issues of helping deaf children learn to speak and listen like normal-hearing typically developing children and reach their full intellectual potential, research on deaf children with cochlear implants also serves as a unique "model system" to study the effects of brain plasticity, experience- and activity-dependent learning and neural development in speech perception and language development.

Clinical assessments of outcomes and benefits following implantation have routinely used measures of speech perception and spoken word recognition. Performance on open-set spoken word recognition tasks is the "gold standard" of clinical outcomes in the field and is frequently used to track benefit over time in both children and adults after implantation (Kirk & Choi, 2009). In open-set word recognition tests, no response alternatives are provided, and a listener must use their entire lexicon for recognition. In our studies of individual differences in deaf children with cochlear implants, we have found that open-set spoken word recognition performance is strongly correlated with a large number of other speech and language outcome measures, such as discrimination of minimal pairs of words, vocabulary knowledge, sentence comprehension as well as speech intelligibility (Pisoni *et al.*, 2000; Pisoni, 2005). Our findings also suggest that many deaf children with cochlear implants, especially low-functioning children, may have significant comorbid delays and deficits in several other domains of learning, memory, and neurocognitive functioning, specifically episodic memory processes related to the encoding, storage and retrieval of the fine-grained acoustic-phonetic and indexical details of speech.

## Summary and Conclusions

In this chapter, I have argued that speech perception and spoken word recognition draw heavily on highly detailed episodic encoding of the early sensory properties of speech. Sensory processing and early encoding are critical elementary components supporting robust speech perception and word recognition, but speech perception and language processing require more than just hearing and sensory encoding of speech signals at the auditory periphery. Speech perception and language comprehension also require that multiple components are connected to each other and work together in synchrony as a functionally integrated information processing system, not merely a collection of separate autonomous processing modules. One of the reasons why speech perception is highly robust is because human listeners are able make optimal use of multiple sources of information encoded in the speech signal—the traditional symbolic-linguistic pathway that encodes fine-grained highly detailed acoustic-phonetic information specifying the talker's symbolic linguistic message and the indexical pathway, which encodes contextual attributes specifying the talker's gender, regional dialect, and mental and physical state. We are now coming to realize that the traditional abstractionist symbol-processing view of speech perception is incomplete and only half the story. Variability in speech is lawful and highly informative and is an extremely valuable source of episodic context that listeners routinely make use of in recognizing speech in noise and under adverse listening conditions. Research on processing of variability in speech has also provided several new promising directions for understanding the enormous individual differences in speech and language outcomes routinely found in profoundly deaf children who have received cochlear implants. The individual differences observed in this clinical population are not mysterious, anomalous or idiopathic but reflect how the compromised auditory system and brain adapts and accommodates to highly degraded underspecified acoustic-phonetic and indexical information transmitted by cochlear implants.

## REFERENCES


Abercrombie, D. (1967). *Elements of general phonetics*. Edinburg: Edinburgh University.

Allen, J. B. (1994). How do humans process and recognize speech? *IEEE Transactions on Speech Audio*, 2(4), 567–577. doi: 10.1109/89.326615

Allen, J. B., Hunnicutt, M. S., & Klatt, D. (1987). *From text to speech: The MITalk system*. Cambridge, UK: Cambridge University Press.

Bell, T. S. & Wilson, R. H. (2001). Sentence recognition materials based on frequency of word use and lexical confusability. *Journal of the American Academy of Audiology*, 12, 514–522.

Borden, G. J., Harris, K. S., & Raphael, L. J. (2011). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech 6th edition*. Baltimore: Wolters Kluwer/ Lippincott Williams & Wilkins.

Bybee, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Chomsky, N. & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. Duncan Luce, Robert R. Bush, and Eugene Galanter (Eds.) *Handbook of mathematical psychology* (pp. 269–321). New York: Wiley.

Clopper, C. G. & Pisoni, D. B. (2004a). Homebodies and Army brats: Some effects of early linguistic experience and residential history on dialect categorization. Language Variation and Change, 16, 31–48. PMID: 21533011; PMCID: PMC3083069. doi: 10.1017/S0954394504161036

Clopper, C. G. & Pisoni, D. B. (2004b). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32, 111–140. PMID: 21451736; PMCID: PMC3065110. doi: 10.1016/S0095-4470(03)00009-3

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., & Gerstman,L. J. (1952). Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America*, 24, 597–606. doi: 10.1121/1.1906940

Creelman, C. D. (1957). Case of the unknown talker. *Journal of the Acoustical Society of America*, 29, 655. doi: 10.1121/1.1909003

Elman, J. L. & McClelland, J. L. (1986). Exploiting lawful variability in the speech waveform. In J.S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processing* (pp. 360–385). Hillsdale, NJ: Erlbaum.

Fant, C. G. M. (1962). Descriptive analysis of the acoustics of speech. *Logos*, 5, 3–17.

Flanagan, J. L. (1965). *Speech analysis synthesis and perception*. Heidelberg: Springer-Verlag.

Fletcher, H. (1953). *Speech and hearing in communication*. Huntington, New York: Krieger.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New Approaches to Language Mechanisms* (pp. 257–287). Amsterdam: North Holland.

Frederiksen, J. R. (1971). Statistical decision model for auditory word recognition. *Psychological Review*, 78(5), 409–419. doi: 10.1037/h0031326

Gaskell, M. G. (2007). *The oxford handbook of psycholinguistics*. New York: Oxford University Press.

Gaskell, M. G. & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology*, 45, 220–266. doi: 10.1016/S0010-0285(02)00003-8

Geers, A., Brenner, C., & Davidson, L. (2003). Factors associated with development of speech perception skills in children implanted by age five. *Ear and Hearing*, 24, 24S–35S. doi: 10.1097/01.AUD.0000051687.99218.0F

Geers, A. E., Tobey, E. A., & Moog, J. S. (2011). Long-term outcomes of cochlear implantation in early childhood. *Ear and Hearing*, Vol. 31, No. 1, Supplement 2011. doi: 10.1097/AUD.0b013e3181ffd5dc

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251–279. doi: 10.1037/0033-295X.105.2.251

Grossberg, S., & Myers, C. W. (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review*, 107, 735–767. doi: 10.1037/0033-295X.107.4.735

Hickok, G. & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402. doi: 10.1038/nrn2113

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, 93, 411–428. doi: 10.1037/0033-295X.93.4.411

Hirahara, T. & Kato, H. (1992). The effect of F0 on vowel identification. In Y. Tohkura, E. Vatikiotis-Bateson & Y Sagisaka (Eds.) *Speech perception, production and linguistic structure* (Pp. 89–112). Tokyo: IOS Press.

Hirsh, I. J. (1952). The intelligibility of speech. In I. J. Hirsch *The Measurement of Hearing*. (ch. 5). New York: McGraw-Hill Book Company, INC.

Hockett, C. F. (1955). *A manual of phonology*. Baltimore: Waverly Press.

Hudgins, C. V., Hawkins, J. E., Karlin, J. E., & Stevens, S. S. (1947). The development of recorded auditory tests for measuring hearing loss for speech. *The Laryngoscope*. 57(1), 57–89. doi: 10.1288/00005537-194701000-00005

Kirk, K. I. & Choi, S. (2009). Clinical Investigations of Cochlear Implant Performance. In J. K. Niparko (Ed.), *Cochlear Implants: Principles and Practice, Second Edition*, (pp. 191–222). Philadelphia: Lippincott Williams & Wilkins.

Kirk, K. I., Pisoni, D. B., & Miyamoto, R. T. (2000). Lexical discrimination by children with cochlear implants: Effects of age at implantation and communication mode. In S.B. Waltzman and N.L. Cohen (Eds.), *Cochlear Implants* (pp. 252–254). New York: Thieme Medical Publishers.

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7, 279–312.

Klatt, D. H. (1986). The problem of variability in speech recognition and in models of speech perception. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and Variability in Speech Processing* (pp. 300–319). Hillsdale, NJ: Erlbaum.

Konishi, M. (1985). Birdsong: From behavior to neuron. *Annual Review of Neuroscience*, 8, 125–170. doi: 10.1146/annurev.ne.08.030185.001013

Kral, A. (2013). Auditory critical periods: A review from system's perspective. *Neuroscience*, 247, 117–133. doi: 10.1016/j.neuroscience.2013.05.021

Kreiman, J. & Sidtis, D. (2011). *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Oxford, UK: Wiley-Blackwell. doi: 10.1002/9781444395068

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99, 22–44. doi: 10.1037/0033-295X.99.1.22

Liberman, A. M. (1996). *Speech: A Special Code*. Cambridge, Massachusettes: The MIT Press.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431–461. doi: 10.1037/h0020279

Liberman, A. M. & Studdert-Kennedy, M. (1977). Phonetic Perception. In R. Held, H. Leibowitz & H. L. Teuber *Handbook of Sensory Physiology*, Vol VIII, (pp. 145–174). Heidelberg: Springer-Verlag, Inc.

Licklider, J. C. R. & Miller, G. A. (1951). The perception of speech. In Stevens, S. S. (Ed), (1951). *Handbook of experimental psychology* (pp. 1040–1074). Oxford, England: Wiley.

Lindgren, N. (1967). Speech-Man's natural communication. *IEEE Spectrum*, 4, 75–86. doi: 10.1109/MSPEC.1967.5215808

Luce, P. A. & Pisoni, D. B. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear & Hearing*, 19, 1–36. doi: 10.1097/00003446-199802000-00001

Marler, P. & Peters, S. (1988). Sensitive periods for song acquisition from tape recordings and live tutors in the swamp sparrow, Melospiza georgiana. *Ethology*, 77, 76–84. doi: 10.1111/j.1439-0310.1988.tb00193.x

McNellis, M. G. & Blumstein, S. E. (2001). Self-organizing dynamics of lexical access in normal and aphasics. *Journal of Cognitive Neuroscience*, 13, 151–170. doi: 10.1162/089892901564216

Miller, G. A. (1951). The Perception of Speech. In G. A. Miller *Language and Communication (ch. 3)*. New York: McGraw-Hill Book Company, Inc. doi: 10.1037/11135-003

Moore, R. K. (2005). Towards a unified theory of spoken language processing. *Proceedings of the 4th IEEE International Conference on Cognitive Informatics*, Irvine, USA, 8–10 August 2005. doi: 10.1109/coginf.2005.1532629

Moore, R. K. (2007a). PRESENCE: A Human-inspired architecture for speech-based human-machine interaction. *IEEE Transactions on Computers*, 56(9), 1176–1188. doi: 10.1109/TC.2007.1080

Moore, R. K. (2007b). Spoken language processing: Piecing together the puzzle. *Speech Communication*, 49, 418–435. doi: 10.1016/j.specom.2007.01.011

Morton, J. (1979). Word recognition. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics 2: Structures and Processes* (pp. 107–156). Cambridge: MIT Press.

Nahum, M., Nelken, I., & Ahissar, M. (2008). Low-Level Information and High-Level Perception: The Case of Speech in Noise. *PLOS Biology*, 6(5), doi: 10.1371/journal.pbio.0060126.

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N. Y., Quittner, A. L., & Fink, N. E. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association*, 303(15), 1498–1506. doi: 10.1001/jama.2010.451

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 115, 39–57. doi: 10.1037/0096-3445.115.1.39

Nygaard, L. C. & Pisoni, D. B. (1995). Speech perception: New directions in research and theory. In J.L. Miller and P.D. Eimas (Eds.), Handbook of Perception and Cognition. Volume II: Speech, Language and Communication (pp. 63–96). New York: Academic Press.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5, 42–46. PMID: 21526138; PMCID: PMC3081685. doi: 10.1111/j.1467-9280.1994.tb00612.x

Obleser, J., Wise, R. J., Alex Dresner, M., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *The Journal of Neuroscience*, 27(9), 2283–2289. doi: 10.1523/JNEUROSCI.4663-06.2007

Pichora-Fuller, M. K., Schneider, B. A. & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustic Society of America*, 91(1), 593–608. doi: 10.1121/1.412282

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure* (pp. 137–158). Amsterdam: John Benjamins. doi: 10.1075/tsl.45.08pie

Pisoni, D. B. (1978). Speech Perception. In W. K. Estes (Ed.), *Handbook of Learning and Cognitive Processes: Volume 6* (pp. 167–233). Hillsdale, NJ: Erlbaum Associates.

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson & J.W. Mullennix (Eds.), Talker Variability in Speech Processing (pp. 9–32). San Diego: Academic Press.

Pisoni, D. B. (2000). Cognitive factors and cochlear implants: Some thoughts on perception, learning, and memory in speech perception. *Ear & Hearing*, 21, 70–78. doi: 10.1097/00003446-200002000-00010

Pisoni, D. B. (2005). Speech perception in deaf children with cochlear implants. In D.B. Pisoni & R.E. Remez (Eds.), *Handbook of speech perception*. (pp. 494–523). Oxford: Blackwell Publishers. doi: 10.1002/9780470757024.ch20

Pisoni, D. B. & Levi, S. V. (2007). Representations and representational specificity in speech perception and spoken word recognition. In M.G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 3–18). Oxford University Press: UK.

Pisoni, D. B. & Luce, P. A. (1986). Speech perception: Research, theory and the principle issues. In E. C. Schwab and H. C. Nusbaum (Eds.), *Pattern Recognition by Humans and Machines: Speech Perception*. Volume 1 (pp. 1–50). New York: Academic Press. doi: 10.1016/B978-0-12-631403-8.50006-0

Pisoni, D.B. & Remez, R.E. (Eds.). (2005). The Handbook of Speech Perception. Malden, MA: Blackwell Publishing. doi: 10.1002/9780470757024

Pisoni, D. B., Cleary, M., Geers, A. & Tobey, E. A. (2000). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *Volta Review*, 101, 111–164.

Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, 4, 75–95. doi: 10.1016/0167-6393(85)90037-8

Port, R. F. (2007). How are words stored in memory? Beyond phones and phonemes. (2007). *New Ideas in Psychology*, 25, 143–170. doi: 10.1016/j.newideapsych.2007.02.001

Rosenzweig, M. R., & Stone, G. (1948). Wartime research in psycho-acoustics. *Review of Educational Research Special Edition: Psychological Research in the Armed Forces*, 18(6), 642–654. doi: 10.2307/1168189

Scott, S. K. & Wise, R. J. S. (2003). Functional imaging and language: A critical guide to methodology and analysis. *Speech Communication*, 41, 23–34. doi: 10.1016/S0167-6393(02)00089-4

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 5234, 303–304. doi: 10.1126/science.270.5234.303

Sporns, O. (1998). Biological variability and brain function. In J. Cornwell (Ed.), *Consciousness and Human Identity* (pp. 38–56). Oxford: Oxford University Press.

Sporns, O. (2003). Network analysis, complexity, and brain function. *Complexity*, 8, 56–60. doi: 10.1002/cplx.10047

Stevens, K. N. & House, A. S. (1972). Speech Perception. In J. V. Tobias (Ed.), *Foundations of Modern Auditory Theory*, (pp. 3–62). New York: Academic Press.

Studdert-Kennedy, M. (1974). The perception of speech. In T. A. Sebeok, (Ed.), *Current trends in linguistics Volume XII* (pp. 2349–2385). The Hague: Mouton.

Studdert-Kennedy, M. (1976). Speech Perception. In N. J. Lass, (Ed.), *Contemporary Issues in Experimental Phonetics* (pp. 243–293). New York: Academic Press, Inc.

Tobey, E. A., Geers, A. E., Brenner, C., Altuna, D., & Gabbert, G. (2003). Factors associated with development of speech production skills in children implanted by age five. *Ear & Hearing*, 24, 1S, 36S–45S. doi: 10.1097/01.AUD.0000051688.48224.A6

Treisman, M. (1978a). Space or lexicon? The word frequency effect and the error response frequency effect. *Journal of Verbal Learning and Verbal Behavior*, 17(1), 37–59. doi: 10.1016/S0022-5371(78)90518-2

Treisman, M. (1978b). A theory of the identification of complex stimuli with an application to word recognition. *Psychological Review*, 85(6), 525–570. doi: 10.1037/0033-295X.85.6.525

Van Lancker D. & Kreiman J. (1987). Voice discrimination and recognition are separate abilities. *Neuropsychologia*, 25, 829–834. doi: 10.1016/0028-3932(87)90120-5

# 10 Cross-Language and Second Language Speech Perception

OCKE-SCHWEN BOHN

Aarhus University

## Introduction

What happens when naive listeners encounter the sounds of a foreign language for the first time, and what happens when naive listeners learn an additional sound system in the course of second language acquisition? This chapter deals with cross-language speech perception (the naive perception of the sounds of an unfamiliar language), and with the related dynamic that unfolds when the sound systems of the native language (L1) and a non-native language (L2)[1] have to coexist in the mind of a second language learner.

Research on cross-language and L2 speech perception is primarily motivated by three overlapping interests, which each are most closely associated a) with basic issues in speech perception research (see Chapter 9), b) with core questions about the malleability of the human mind and how it copes with two or more languages (see also Chapters 4, 13, 14, 23, 28, 29), and c) with important practical, methodological, and theoretical aspects of L2 acquisition (see also Chapters 4 and 39).

Regarding speech perception research, studies of monolingual listeners perceiving the sounds of their L1 can only to a very limited extent address the question of which aspects of speech perception are language-specific, and which aspects reflect universal ways of how humans perceive speech (see also Chapter 9). This chapter will present several instances of universal phenomena of speech perception that could only have been discovered through cross-language research. With respect to the malleability of the human mind, a pervasive question in general psychological research and especially in L2 acquisition research concerns the extent to which experience can change well-established perceptual patterns past infancy. This chapter will show how L2 speech perception research has contributed importantly to an understanding of the factors which favor perceptual learning past infancy and which make it possible for adults to reshape their speech

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

perception abilities to accommodate the sounds of the L2. Regarding L2 acquisition in general, its major aim has been to predict learning problems, and to discover how the learner (not the theoretical linguist or psychologist) analyzes linguistic material. In order to meet these challenges, L2 acquisition research needs to use appropriate methodological tools guided by empirically motivated theoretical approaches. This chapter will present an overview of research inspired by current models of cross-language and L2 speech perception, which has resulted in a fairly good understanding of the causes of perceptual problems, and of the levels of analysis used by naive listeners and by L2 learners.

The one overarching question in all research on cross-language and L2 speech perception is: "What makes non-native sounds difficult to perceive?" This is the kind of question to which everyone, including the general public, has an answer: "It's the L1 and the age of the learner." Just about any L1 English speaker knows that L1 Japanese speakers have a problem with the English contrast between/r/ and/l/(as in *rock* and *lock*), whereas other non-natives can easily tell a rock from a lock. And just about anyone who knows an immigrant family can come up with an example which shows that, even after years of residence in the L2 community, the parents are still struggling with a new sound contrast, whereas the children are not.

However, the true complexity of the original question becomes obvious once we ask what it is about the L1, or more specifically about the mapping of the L2 onto the L1, that makes non-native sounds difficult to perceive, and what precisely is meant by "age." Concerning the age factor in non-native speech perception, research has presented very clear evidence that the differences between more successful younger and less successful older learners are not biologically grounded. We can be quite certain that these differences do not exist because of neurological maturation, which causes the flexible child brain to change into an inflexible adult brain, as posited by Lenneberg's (1967) Critical Period Hypothesis. Rather, the adult brain retains the capacity to change, with age differences in non-native speech perception reflecting experience-based differences in the development of phonetic systems. Concerning L2 to L1 mapping, the research on non-native speech perception which will be reviewed in this chapter has shown on which level this occurs, and has come up with a fairly good account of why some of these mappings cause perceptual problems, while others do not.

This chapter provides an overview of cross-language and L2 perception by focusing on how theoretical models have addressed, and what empirical evidence has contributed to, these three issues:

- the perceptual relationship of sounds of the native and the non-native language in the mind of the native listener and the L2 learner—the mapping issue,
- how this relationship may or may not cause perceptual and learning difficulty—the perceptual and learning difficulty/ease issue, and
- whether and how experience with the non-native language affects the perceptual organization of speech sounds in the mind of L2 learners—the plasticity issue.

A handbook chapter cannot do full justice to the insights that have accumulated over the past ca. 30 years of intensive research on cross-language and L2 perception, or to the entire history of the field, which spans back at least 130 years (Hale, 1885; Boas, 1889; Polivanov, 1931). This chapter will present the most important theoretical approaches and some of the evidence (dis-)confirming predictions based on these models by drawing on studies that have examined the cross-language and L2 perception of segments (consonants and vowels) and of lexical tones. For an overview of the much less explored area of intonation, see the special issue of *Studies in Second Language Acquisition* edited by Mennen and de Leeuw (2014). For more comprehensive overviews of the field of cross-language and L2 perception, see the edited volumes by Strange (1995) and Bohn & Munro (2007), and the special issue of *Journal of Phonetics* edited by Best *et al*. (2011).

## The mapping issue

The development of infant speech perception over the first year of life can to some extent be characterized as a process by which the infant constructs a multidimensional perceptual space inhabited by the sounds of the ambient language(s). The nature of this space and its construction is subject to innate auditory constraints (Dooling & Hulse, 2014) and is also shaped by prenatal experience (e.g., DeCasper & Spence, 1986; Moon, Lagercrantz & Kuhl, 2013). For those growing up and staying in a monolingual environment, this perceptual space becomes fine-tuned to the ambient language throughout infancy (e.g., Kuhl *et al*. 2006; Ko, Soderstrom, & Morgan 2009), childhood (e.g., Hazan & Barrett, 2000; Sundara, Polka & Genesse 2006), up to and including adulthood (Butcher, 1976; Burnham & Mattock, 2007), resulting in greater perceptual resolution for important phonetic dimensions in the native language, while attenuating phonetic dimensions that are nonfunctional in the L1.

In cross-language and in L2 perception, listeners with a well-established and functionally fine-tuned perceptual space for the L1 have to deal with a non-native sound system for which the existing space may be quite ill-designed (Iverson *et al*., 2003). The function of the native language perceptual space in cross-language and L2 perception has been described as a "phonological sieve" (Trubetzkoy, 1939), the "lens of native phonetics" (Bohn *et al*., 2011), or simply as the "native language filter" (Ryan, 1983). These metaphors (sieve, lens, or filter) are conceptually equivalent, but a very important difference exists between Trubetzkoy's original assumption that the mapping of a non-native to the native language happens on the *phonological* level, and later clear evidence (summarized below) that *phonetic* detail of native and non-native speech sounds strongly influences listeners in cross-language and early L2 perception.

In research on cross-language and L2 perception, a number of methods have been used to examine how the sounds of an L2 are perceptually assimilated to native categories (perceptual assimilation is also referred to as interlingual identification or as mapping, and these three terms will be used interchangeably unless the context requires differentiation). Even though the predictions of models of cross-language and L2 perception cannot be tested without clear and valid insights on the L2-to-L1 mapping, surprisingly few studies have compared these methods for accuracy and validity, the most important exceptions being a series of studies by Strange and her collaborators (e.g., Strange *et al.*, 1998, 2004, 2005, 2007; see also Rochet, 1995; Strange, 2007; Bohn, 2002).

The easiest method, but also the method that is least valid, for assessing L2-to-L1 mapping, is to compare the phonetic or phonemic symbols used for transcribing the sounds of interest of the two languages. This method is typically used as a first heuristic aimed at identifying potential problems for L2 learners (e.g., Briere, 1968; Collins & Mees, 1984). However, there are at least four reasons why this method should not be used: First, the choice of comparing either phonetic or phonemic symbols across two languages preempts the empirical question of whether L2-to-L1 mapping occurs at the phonetic or the phonological level. Second, phonetically (and phonologically) identical sounds may be transcribed with different symbols for historical reasons, such as the vowel in *bet*, which is phonetically [ɛ] in the standards accents of both American and Southern British English, but is incorrectly transcribed as [e] in traditional descriptions of Southern British English (Schmitt, 2007). Depending on whether one chooses [ɛ] or [e] as the English vowel to be mapped onto a Spanish vowel space (which has only [e], not [ɛ]), the result would either be a perfect match or a mismatch. Third, and conversely, acoustic cross-language studies often reveal for both consonants and vowels what Disner concluded from her comparison of the vowels of Germanic languages, namely, that "vowels transcribed with the same symbol do not necessarily have identical phonetic quality" (Disner, 1978: 21). Fourth, it would take very detailed and thus unwieldy transcriptions to capture the occasionally large context-induced variation for segments in question. A case in point is the [ɪ] vowel of English and of German, which would appear to be a perfect cross-language match based on the comparison of symbols used to transcribe the vowel in English *kit* and in German *Kitt*. However, in a study comparing the acoustic and perceptual similarity of German and American English vowels, Strange *et al.* (2005) showed that the alleged 1:1 mapping suggested by the shared symbol does not do justice to the language-specific context-dependent acoustic variation and resulting complex perceptual relationships between the realization of vowels of these two languages (see also Harnsberger, 2000, 2001; Levy & Strange, 2008).

If the comparison of phonetic or phonemic symbols yields a misleading picture of how listeners map the sounds of a non-native to their native language, acoustic comparisons of the sounds of the two languages might appear to be a feasible alternative. A number of cross-language and L2 studies have indeed used this method, either as a first heuristic or as the only means to predict L2-to-L1 mapping (e.g., Bohn & Flege, 1990; Flege, 1991, 1992; Polka, 1991, 1995; Wu, Munro, & Wang, 2014). It is likely that cross-language comparisons of some phonetically relevant acoustic dimensions may shed light on this mapping, such as comparisons of pitch (F0) contours for the perception of tonal contrasts (Wu *et al.*, 2014) or of voice onset time (VOT) for the perception of voicing contrasts (Bond & Fokes, 1991). However, the acoustic comparison method for determining perceptual L2-to-L1 mapping is problematic even for these fairly straightforward dimensions and especially for some segment types (such as vowels) for two main reasons. The first of these is methodological: Acoustic comparisons only make sense if a number of variables which may affect the realization of speech sounds are controlled when recording speakers from different language groups, for example, speaking rate and phonetic and prosodic context. Other variables, such as vocal tract size differences between speakers of two languages may be impossible to control and will have to be treated with appropriate normalization procedures (e.g., for vowels, see Adank, Smits, & van Hout, 2004). An additional methodological problem derives from the fact that it may be difficult to know beforehand which acoustic parameters are actually perceptually relevant for listeners' L2-to-L1 mapping. For example, is it sufficient to compare vowels across languages based on measurements of vowel duration and the formant frequencies near the vowel midpoint, or should comparisons include the lack of, or presence of vowel inherent spectral change (e.g., Andruski & Nearey, 1992) and the different language-specific dynamics of the transitions at the syllable margins (e.g., Strange & Bohn, 1998)?

Even if these methodological problems can be satisfactorily addressed, there is a second problem with acoustic comparisons, namely that of validity: How well do these comparisons reflect listeners' perception of the relation of native and non-native sounds? Strange and her collaborators have addressed this question in a series of studies comparing the cross-language perception of vowels to their acoustic similarity (e.g., Strange *et al.*, 2004, 2005, 2007), and the result of these studies is not encouraging: There were "marked discrepancies … between acoustic comparisons and direct perceptual comparisons of L1/L2 similarity patterns" (Strange, 2007, p. 54; for similar results see Levy, 2009 and Nishi *et al.*, 2008). The main reason for this discrepancy is that the production (and hence acoustic properties) of vowels varies systematically as a function of prosodic and phonetic context, that the kind and degree of this variation is specific to each language, and that listeners are guided by L1-specific expectations regarding this variability. For instance, Bohn and Steinlen (2003) reported that L1 Danish listeners mapped the English [ʌ] vowel (as in *hut*) unto the Danish back vowel [ɔ] if it occurred in a [dʌt] syllable (*dut*), but to the Danish front vowel [ø] if it occurred in a [gʌk] syllable (*guk*), probably because Danish vowels vary much less across different phonetic contexts than English vowels (Bohn, 2004).

In conclusion, neither the comparison of phonetic symbols across languages, nor comparisons of the acoustic properties of speech sounds yield valid information on how naive listeners and L2 learners map the sounds of the non-native to the native language. The consequence is that "if researchers want to know how … listeners are perceptually assimilating L2 segments to L1 phonological categories, direct measures of those perceptual assimilation patterns are probably the most reliable indicators of L1/L2 perceptual relationships" (Strange, 2007, p. 54; see also Rochet, 1995; Bohn, 2002).

Two approaches can be identified in research which has heeded this advice: In one of these, listeners are asked to provide a discrete native language orthographic label for a non-native stimulus, thus providing information on the *ecphoric*[2] mapping of a stimulus to an episodic memory trace. In the other approach, listeners are presented with two auditory stimuli and asked to provide graded ratings of the *perceptual* similarity (Tulving, 1981) of a native and a non-native speech sound. This approach, in which listeners compare specific productions of L1 and L2 sounds and rate them on a scale ranging from "very similar" to "very dissimilar," has not been widely used (but see Flege, Munro, & Fox, 1994), probably for two reasons. First, because of the theoretical problem that the relation between the specific L1 sound that the listener hears and the listener's internal representation of that sound is difficult to assess, and second, because of the practical problem that participants may have to provide ratings of perceptual similarity on a very large number of stimulus pairs. For example, the Flege *et al*. (1994) study paired three (of five) Spanish vowels with seven (of 12) English vowels in one phonetic context, resulting in 405 vowel pairs (including foils) which the participants had to rate. Ratings on the full set of vowels, produced in different phonetic contexts, would require thousands of vowels pairs. It might be difficult to find volunteers for such a study.

Ecphoric mapping tasks have been used in a large number of studies, either presenting listeners with a free choice of orthographic (or IPA) labels (as in Best, McRoberts, & Goodell, 2001; Bohn & Best, 2012) or a closed set of force choice alternatives (e.g., Bohn & Flege, 1992; Polka & Bohn, 1996; Guion *et al*., 2000; Best *et al*., 2003; Bohn & Steinlen, 2003, Strange *et al*., 2004, 2005). In many of these studies, listeners do not just provide an L1 response to an L2 stimulus, but also rate the L1-L2 match on a Likert scale for goodness of fit. Some studies (e.g., Guion *et al*., 2000) combine the identification and the goodness-of-fit data into a single metric called fit index, thus providing a weighted measure of the L1-to-L2-mapping. For example, in the Guion *et al*. (2000) study, L1 Japanese listeners identified English [b] stimuli with Japanese/b/in 84% of the instances, and these mappings received a mean rating of 5.3 on a scale ranging from 1 (bad example) to 7 (very good example). The proportion of/b/identifications (0.84), multiplied by the mean goodness rating, yields a fit index of 4.5 of English/b/to Japanese/b/. The fit index looks like a straightforward quantification of L1-to-L2 mappings, but, as Strange (2007) pointed out, it rests on the assumption that the rating scale is an interval scale. Strange (2007) recommends a more conservative approach which assumes an ordinal level of quantification for Likert-scale judgments of goodness of fit (e.g., Strange *et al*., 1998, 2004, 2005).

A general problem with ecphoric mapping tasks (with or without additional goodness-of-fit ratings) is related to the potential ambiguity of the labels used by the listeners. If the listeners have a free choice of response labels, their responses may be hard to interpret (e.g., Bohn *et al*., 2011). If listeners are presented with a set of forced choice alternatives, they may not use the labels consistently, especially if they have to use orthographic labels of a language like English with irregular correspondences between sounds and letters.

In spite of these problems and in spite of the fact that it is unknown whether ecphoric or perceptual tasks provide a more valid picture of L2-to-L1 mapping (see Bohn & Polka, 2009), studies using either approach have provided much more reliable and valid bases for evaluating the claims of models of cross-language and L2 perception (see below) than acoustic comparisons or comparisons of phonetic symbols. A very important contribution of these studies to speech perception research in general is the solid support for the claim that the basis for interlingual identification is "substantial … rather than formal" (Catford, 1968, p. 164). Just about any study of L1-to-L2 mapping has confirmed that cross-language perception occurs at the level of context-sensitive perceptual units (allophones), not at the abstract level of the phoneme (see also Johansson, 1975a,b; Logan, Lively, & Pisoni 1991, 1993; Harnsberger 2000; Strange *et al.*, 2007). Finally, it should be mentioned that L2-to-L1 mapping has so far exclusively been studied in the auditory domain. However, several studies of cross-language perception and of perceptual training have examined the contribution of the visual channel to speech perception (e.g., Werker, Frost, & McGurk, 1992; Hardison 1999, 2003, 2005; Hazan *et al.*, 2005; Navarra & Soto-Faraco, 2007; Wang, Behne, & Jiang, 2008). It would be interesting to assess how visual information contributes to L2-to-L1 mapping, for instance, whether visual information on lip rounding in English/ʃ/(as in *show*) makes this segment appear less similar to the unrounded Danish/ç/(as in *sjov*) and thus, according to the models presented in the next section, more learnable. An additional potential channel of information that could guide or misguide L2 learners who are literate is orthography (e.g., Escudero, Hayes-Harb, & Mitterer, 2008).

## The perceptual and learning difficulty/ease issue

An important basic finding of research on non-native speech perception is that any perceptual problem that non-native listeners might have exists at the central linguistic level of phonetic or phonemic processing, not at the peripheral auditory level. For instance, Miyawaki *et al.* (1975) reported that L1 Japanese listeners discriminated stimuli from a [ɹɑ]-[ɫɑ] continuum continuously and at a generally lower level than L1 American English listeners, who showed a clear discrimination peak for stimuli from this native contrast near the phoneme boundary. However, the two listener groups did not differ when they discriminated nonspeech stimuli that consisted only of the most important acoustic cue to the [ɹɑ]-[ɫɑ] contrast, namely isolated F3 contours. Further studies (e.g., Werker & Tees, 1984 on consonants; Kewley-Port, Bohn, & Nishi, 2005 on vowels) clearly indicate that adult listeners have not lost peripheral auditory sensitivities to acoustic cues differentiating non-native contrast, but that perceptual problems exist at a central linguistic level. In other words, any perceptual or learning difficulty encountered by non-native listeners is not due to sensory loss (and therefore irreversible) but to attentional preferences (and therefore perhaps malleable).

The ease or difficulty with which listeners perceive and learn a particular non-native sound or sound contrast varies as a function of the native language. For example, Bohn and Best (2012) reported that L1 Danish listeners identified and discriminated stimuli from an English *rock-lock* continuum much like L1 English listeners, whereas L1 Japanese listeners' identification function was less categorical, and their discrimination performance was much lower than the L1 English listeners'. For listeners who share an L1, the ease or difficulty with which they perceive and learn non-native sounds differs for different sounds and sound contrasts of the L2. For example, Flege, Bohn, and Jang (1997) examined the use of spectral and temporal cues in the identification of stimuli from two English vowel continua, *beat-bit* and *bet-bat*. They reported that L1 Spanish listeners identified the *bet-bat* continuum much like the L1 English listeners, (using the spectral cue much more than the duration cue), but differed greatly from the L1 English listeners in their identification of the *beat-bit* continuum (relying much more on the duration cue than the L1 English listeners). This variation in perceptual difficulty has not just been observed for consonants (e.g., Polka, 1992; Bohn *et al.*, 2011) and vowels (e.g., Polka, 1995; Polka & Bohn, 1996; Best *et al.*, 2003), but also for tones (Hallé, Chang, & Best, 2004; Burnham & Mattock, 2007; So & Best, 2010, 2014).

The literature on cross-language and L2 perception is replete with examples like these, which show that non-native speech sounds and sound contrasts vary in terms of perceptual difficulty, both initially and in terms of ease of learning. Over the past ca. 30 years, several models of cross-language perception and L2 speech development have been developed to account for why some non-native sounds and contrast are easy to perceive, while others provide difficulties. Because of space limitations, this section will focus exclusively on two of these models, Best's (1995) Perceptual Assimilation Model (PAM) and its extension to L2 learning, PAM-L2 (Best & Tyler, 2007), and Flege's (1995) Speech Learning Model (SLM). Both PAM and SLM have been enormously influential. These models have been around long enough to be tested in a large number of cross-language and L2 perception studies on many different segment types (e.g., vowels, stop consonants, approximants). For other more recently developed models of cross-language perception and L2 speech, which have not yet been tested on as large a range of phenomena as SLM and PAM/PAM-L2, see Strange (2007, 2011) on the Automatic Selective Perception model (ASP), Kuhl *et al.* (2008) on the Native Language Magnet theory-expanded (NLM-e), Escudero and Boersma (2004) on the Second Language Linguistic Perception Model (L2LP), and Major and Kim (1996) on the Similarity Differential Rate hypothesis (SDRH).

## *The Perceptual Assimilation Model (PAM) and PAM-L2*

PAM was developed to explain cross-language speech perception by naive listeners. Specifically, it aims to account for why the degree of perceptual difficulty varies considerably across non-native contrasts. PAM assumes that listeners assimilate non-native phones to native phonological categories that are phonetically "closest" in terms of phonetic similarity.[3] Categorization of a non-native phone as a realization of a native phoneme may result in a (near-) perfect match, as when L1 Danish listeners assimilate English [$p^h$] to Danish/p/in 100% of all instance, and rate the match at 7.4 on a 9-point scale (Horslund, Ellegaard, & Bohn, 2015). Categorization may also result in a less than perfect match, as when English [ʃ] is assimilated to Danish/ç/in 82.5% of the instances with a mean rating of 4.3 (Horslund *et al.*, 2015). However, non-native phones are not necessarily "Categorized" (in PAM terminology): If a non-native phone is moderately similar to more than one native phoneme, PAM assumes that it will be perceived as an "Uncategorized" speech sound. In the Horslund *et al.*, 2015 study, this was the case for English [ʤ], which was assimilated to three Danish categories,/dj/,/tj/, and/j/, with a modal/dj/response in just 61% of the instances.[4] Finally, there may be rare cases in which the phonetic characteristics of the non-native phone are quite unlike those of any native sound, as for click sounds produced with an ingressive velaric airstream. The non-native phone, which is then not assimilated to any of the phones in the native phonological system, is considered to be "Non-assimilable" in PAM terminology (see Best, McRoberts, & Goodell, 2001).

PAM predicts that difficulty or ease of cross-language speech perception, as reflected in discrimination levels for non-native contrasts, depends on how contrasting phones are assimilated. The most important assimilation types in PAM, their assimilation to native categories, and the predicted discrimination levels are shown in Table 10.1.

As shown in Table 10.1, there are two contrasts types that should be very easy to discriminate in cross-language perception: The Two Category (TC)

**Table 10.1** Assimilation types, examples of their perception, and predicted discrimination levels according to the Perceptual Assimilation Model (Best, 1995).

| *Assimilation of non-native phonetic contrasts* | | |
|---|---|---|
| *Assimilation Type* | *Perceived as:* | *Discrimination* |
| Two Category TC | Exemplars of two different categories, e.g., English [w]-[ɫ] to German /v/-/l/ | Very good/ excellent |
| Uncategorized-Categorized UC | Uncategorizable speech vs. native exemplar, e.g., Australian English [ɜ]-[ʉ] to Japanese | Very good/ excellent |
| Non-Assimilable NA | Nonspeech sounds (not assimilated to native phonetic space) e.g., Zulu click contrasts to English | Good/excellent |
| Category Goodness CG | Exemplars of single native category differing in goodness of fit e.g., Norwegian [i]-[y][5] to French /i/ | Good/very good |
| Uncategorized-Uncategorized UU | Uncategorizable speech sounds (in between native categories) e.g., Australian English [əʉ]-[o] to Japanese | Fair/good |
| Single Category SC | Exemplars of a single native category (equivalent goodness of fit) e.g., Xhosa [ɓ]-[ᵐb] to Spanish /b/ | Poor/fair |

assimilation type, in which two non-native phones are assimilated to two different L1 categories, as in the case of English [w]-[ɫ] to German/v/and/l/, respectively (e.g., Bohn & Best, 2012), and the Uncategorized-Categorized (UC) assimilation type, in which one of the non-native phones has no clear counterpart in the L1 (as English [ɜ] in Japanese), and the other part of the contrast is categorized as an L1 exemplar (as English [ʉ] to Japanese/ɯ/(as in Bundgaard-Nielsen, Best, & Tyler, 2011a). If the contrasting phones are both assimilated to the same L1 category, discriminability will depend on each non-native phone's similarity to the phonetic properties of the native phoneme. If one of the contrasting phones is a phonetically better example of the native category than the other, this Category Goodness (CG) difference will result in fairly good discrimination, as when Norwegian [i] and [y][5] are both assimilated, with a sizeable difference in goodness of fit, to French/i/(Best *et al.*, 2003). However, if the contrasting phones are equivalent in degree of phonetic fit to the native category (Single Category assimilation—SC), discrimination will be poor, as when Xhosa [ɓ]-[ᵐb] are both assimilated to Spanish /b/, with no difference in goodness of fit (Calderon & Best, 1996). The sounds in a contrast may both be Uncategorized (UU assimilation type), for which PAM predicts fair to good discrimination, as for L1 Japanese perception of Australian English [əʉ]-[o] (Bundgaard-Nielsen *et al.*, 2011a). Finally, the members of a non-native contrast may both be perceived as nonspeech sounds, that is, not assimilated to native phonetic space, in which case discrimination is predicted to be good or excellent (as for the Zulu unaspirated voiceless lateral-apical click contrast in Best, McRoberts, & Sithole, 1988). Thus, PAM predicts the following hierarchy of discriminability of non-native contrasts, based on assimilation patterns, from easiest to most difficult: TC = UC > NA ≥ CG > UU > SC.

PAM, which is a model of cross-language perception by naive monolingual listeners on first contact with non-native sounds, has been extended by Best and Tyler (2007) to PAM-L2, which aims to account for the perceptual changes that occur in L2 learning. An important difference between PAM and PAM-L2 is that PAM predicts how well the sounds of non-native contrasts are discriminated based on how they are assimilated to L1 categories, whereas PAM-L2 considers the L2 learner's aim to learn the higher order invariants (phonemes) of the L2, which are the building blocks of the L2 lexicon. PAM-L2 adds the predictions that L2 learners will have difficulties forming new L2 categories for both SC and TC contrasts. Contrasts assimilated as SC will provide a learning problem because the phonetically similar members of the L2 contrast are "just" variants in the L1 phonological system, leading to the perception of SC-assimilated contrasts as homophones. For TC contrasts, PAM-L2 hypothesizes that L2 learners do not feel any lexical pressure to form new categories because the two phones of the L2 are each similar to two different phonemes of the L1. However, PAM-L2 predicts that L2 learners should easily establish a new L2 category for the member of a CG assimilation that is a less-good example of the L1 category. For the UU and UC assimilation types, PAM-L2 predicts that learning success depends, in addition to lexical pressure, on how similar the Uncategorized L2 phones are perceived to be to L1 phones that approximate them in phonological space, leading to either successful assimilation of the Uncategorized phone to an L1 category (if similar), or to the formation of a novel category (if dissimilar). Finally, with respect to the NA contrast type, Best and Tyler (2007: 31) present various scenarios depending on whether the originally non-assimilable L2 sounds will remain outside the phonological space of the L2 learner, or whether they will become perceptually integrated into that space, most likely as Uncategorized sounds.

## *The Speech Learning Model*

Flege's (1995) Speech Learning Model (SLM) differs from PAM primarily in that the SLM is originally a model of L2 speech production, which focuses on L2 learning of individual phones, whereas PAM is a model of the perception of non-native contrasts by naive monolingual listeners. Even though SLM is an L2 production model, it can fairly easily be extended to L2 perception, not the least because of the SLM assumption that speech production is guided by perceptual representations of L2 speech sounds. SLM generates predictions about how the L2-to-L1 relationship of individual phonetic categories affects the accuracy with which the sounds of the L2 will be produced by learners who are highly experienced with the L2. Table 10.2 illustrates how, according to SLM, the relationship of L2 to L1 sounds in perceptual mapping predicts learning success (production accuracy, and, by extension, perception accuracy) as a function of age of learning (AOL). For a correct understanding of the SLM it is very important to note that L1-L2 phonetic relationships exist on a continuum from "identical" over "similar" to "new" (to make this point clear, Table 10.2 also contains columns labeled "very similar" and "quite dissimilar"), and that even though age differences are predicted by SLM, these differences exist because category formation for L2 sounds becomes less likely through childhood as representations for neighboring L1 sounds develop. That is, the SLM rejects a maturational account of age differences and proposes instead an interactive account (Flege, Yeni-Komshian, & Liu, 1999). Like PAM-L2, SLM assumes that "the mechanisms and processes used in learning the L1 sound system … remain intact over the life span, and can be applied to L2 learning" (Flege, 1995, p. 239).

Table 10.2 is a simplification of SLM predictions because it lists discrete points on the "L1-L2 relation" and the "age of learning" axes where SLM works with continua. The table assumes that all learners, irrespective of AOL, have had sufficient time and experience with the L2 to realize their learning potential. As shown in Table 10.2, identical sounds, such as English and Korean initial [m] (Schmidt, 1996) trivially do not present a learning problem. The SLM prediction for very similar sounds is that only young learners (in SLM terms, "early bilinguals") will successfully perceive and produce L2 sounds which differ only slightly, such as English [$t^h$] - Danish [$t^{sh}$], which differ in that the Danish [$t^{sh}$] has a longer VOT and a more intense frication portion than English [$t^h$] (Horslund *et al.*, 2015; Garibaldi & Bohn, 2015). SLM makes this prediction because "the likelihood of phonetic differences between L1 and L2 sounds … being discerned decreases as AOL increases," and because "category formation for an L2 sound may be blocked

**Table 10.2** SLM predictions of L2 learners' perception (and production) accuracy as a function of age of learning and L1-L2 sound relation (with examples). A "+" indicates successful learning, a "−" incorrect perception and accented production, and "−/+" and "+/−" intermediate degrees of accuracy. Note that the labels on both axes of Table 10.2 are points along continua.

| | *L1-L2 relation* | | | | |
|---|---|---|---|---|---|
| | *Identical* | *Very similar* | *Similar* | *Very dissimilar* | *New* |
| | *English and Korean [m]* | *English [tʰ] - Danish [tˢʰ]* | *English[i],[ɪ]- German [iː], [ɪ]* | *English [ʃ] – Danish [ɕ]* | *English [ɹ] - Japanese* |
| Child | + | + | + | + | + |
| Adolescent | + | − | −/+ | +/− | + |
| Adult | + | − | − | −/+ | + |

because of equivalence classification" (Flege, 1995, p. 239). Adolescent and adult learners perceive similar phones of the L2 to be more or less deviant exemplars of L1 categories, whereas child learners with their less developed L1 sound system are more likely to evade equivalence classification, which would block L2 category formation. SLM claims that for older learners ("late" bilinguals), equivalence classification leads to merged categories, which are used to process perceptually linked L1 and L2 sounds. Table 10.2 illustrates the graded nature of the "L1-L2 relation" and the AOL dimensions, and the resulting graded predictions, by listing the English [i] and [ɪ] and the German [iː] and [ɪ] under "similar." Bohn and Flege (1990, 1992) showed that that even though these English vowels and their German equivalents differ slightly both in terms of their temporal and spectral properties, L1 German listeners identify English [i] with German [iː] and English [ɪ] with German [ɪ]. SLM predicts that younger learners, perhaps even up to adolescence because of arguable greater phonetic difference between these German and English vowels than between the Danish and English /t/ realizations, will form new categories for these vowels, whereas adult learners would end up with one merged category for English [i] and German [iː], and another merged category for English [ɪ] and German [ɪ]. In other words, adult learners will approximate, but not fully master the similar L2 categories. This prediction was confirmed for both production and perception in the Bohn and Flege (1990, 1992) studies with adult L1 German learners.

For sounds of the L2 which are somewhat more dissimilar, such as the more posterior, lip-rounded English [ʃ] with a relatively low peak frequency and the more anterior, unrounded Danish [ɕ] with a relatively high peak frequency (Højen, 2002; Bohn, 2013), SLM predicts that because of the relatively great dissimilarity, adolescents are quite likely, and adults somewhat likely, to discern this difference and establish a new category. Finally, SLM predicts that irrespective of AOL, L2 learners will successfully perceive and produce new sounds which are phonetically very different from any existing L1 sound, given sufficient L2 experience, as reported by MacKain, Best, and Strange (1981) for the perception of the English [ɹ] – [ɫ] contrast by L1 Japanese listeners.

A large proportion of the studies of cross-language and L2 perception which have been conducted since the early 1990s have either been inspired by PAM(-L2) and SLM, or have been designed specifically to test the predictions of these models. These studies have provided considerable support for the hypotheses generated by PAM(-L2) and by SLM, based on experiments examining the perception of a broad range of segment types, such as different consonants and consonants contrasts (e.g., Flege, 1991b; Best & Strange, 1992; Hallé, Best, & Levitt, 1999; Guion *et al*., 2000; Best *et al*., 2001; Harnsberger, 2001; Bohn & Best, 2012; Bohn *et al*., 2011), vowels and vowel contrasts (e.g., Flege, 1987; Bohn & Flege, 1990; Flege *et al*., 1997; Flege, MacKay, & Meador, 1999; Bundgaard-Nielsen *et al*., 2011a; Bohn & Garibaldi, 2017), and tones (e.g., Hallé *et al*., 2004; Sereno & Wang, 2007; So & Best, 2010, 2014).

## *Perceptual and learning difficulty: Beyond L2-to-L1 mapping of segments and contrasts*

All cross-language and L2 perception studies have confirmed the basic assumption of SLM and PAM(-L2) (and other models) that cross-language perception and L2 speech learning are strongly influenced by how L2 sounds are mapped unto L1 categories. PAM-L2 and especially SLM also emphasize the important function of L2 experience, the age factor and language use patterns (e.g., Flege, Schmidt, & Wharton, 1996) and specifically in the case of PAM-L2, the role of lexical factors (familiarity and size, see Flege, Takagi, & Mann, 1996; Bundgaard-Nielsen *et al*., 2011a,b).[6]

However, the very well documented strong influence of L2-to-L1 mapping of segments (SLM) or contrasts (PAM, PAM-L2) on cross-language and L2 perception cannot account for phenomena that are due to either universal perceptual biases or to higher-order generalizations which may shape perception. Regarding the influence of higher-order generalizations, Bohn and Best (2012) reported that discrimination of an English [w]- [j] continuum by listeners with either Danish, German, or French as L1 was at or near ceiling and substantially better than L1 English listener's discrimination, even though only French has an English-like [w]- [j] contrast, whereas German has a [v]-[j] and Danish a [ʋ]-[j] contrast, which both should compromise accurate perception of English [w]-[j] because of phonetic mismatches. The superior discrimination by the non-native listeners was not predicted by PAM, SLM, or any other model of cross-language perception. Bohn and Best suggested that this surprising result was due to a systemic characteristic shared by French, Danish, and German, namely, the contrast between rounded and unrounded front vowels. Thus, the non-native listeners applied the L1-general sensitivity to lip rounding in vowels to the non-native [w]-[j] contrast, which is importantly differentiated through lip rounding. Further support for the idea that current models need to be amended to allow for the influence of higher-order characteristics comes from a recent study by Pajak and Levy (2014), which showed that native experience with duration to distinguish vowel categories sensitizes listeners to non-native duration contrasts for consonants. Pajak and Levy propose a hierarchical inductive inference framework which accounts for how non-native speech perception is affected by higher order generalizations about linguistic structures in addition to L2-to-L1 mappings of non-native segments and contrasts as described by SLM and PAM(-L2).

These mappings, which so importantly shape cross-language and L2 perception, have been found to be subject to universal perceptual biases, which non-native listeners bring to the task of cross-language perception. One of these biases forms the basis of Bohn's (1995) Desensitization Hypothesis for non-native vowel perception, which states that "whenever spectral differences are insufficient to differentiate vowel contrasts because previous linguistic experience did not sensitize listeners to these spectral differences, duration differences will be used to differentiate the non-native vowel contrast" (Bohn, 1995, p. 294), irrespective of whether the duration cue is phonologically relevant in the listener's L1. The Desensitization Hypothesis was originally formulated to account for the surprising finding that L1 Spanish listeners, whose L1 has an/i/vowel as in *beat*, but no/ɪ/vowel as in *bit*, perceptually differentiated an English *beat-bit* continuum by relying heavily on duration, which is not used in Spanish to differentiate vowels, and which L1 English listeners use only as a secondary cue, if at all (Flege & Bohn, 1989, Flege *et al*., 1997). The predictions of the Desensitization Hypothesis have subsequently been confirmed for a variety of L1s, which do not use duration to differentiate vowel contrasts (e.g., Spanish, Russian, Mandarin, Portuguese, Polish, Catalan, Turkish) and L2s, which have a more densely populated vowel space than the respective L1s (English, German, Dutch), as in Kondaurova and Francis, 2008; Rauber *et al*., 2005; Bogacka, 2004; Cebrian, 2006; Escudero, Benders, and Lipski 2009; Darcy and Krüger, 2012.

Another universal bias was discovered by Polka and Bohn in a series of infant vowel perception studies (Polka & Bohn, 1996; Bohn & Polka, 2001). In these studies, which were conducted using an infant-appropriate version of the change/no change paradigm (the headturn procedure), infants' discrimination ability depended strongly on the direction in which a contrast was presented to them. For instance, a change from a background [y] to a foreground [u] was much easier to discriminate than a change from a background [u] to a foreground [y]. Polka and Bohn observed asymmetries like this one for a number of vowel contrasts irrespective of the ambient language of the infants. What all the asymmetries had in common was that it was always a change from the more peripheral vowel (in the articulatory/acoustic vowel space) to the more central vowel that was harder to discriminate than a change in the other direction (from more central to more peripheral), as when both English-learning and German-learning infants could easily discriminate a change from [ɛ] to [æ], but had a harder time discriminating a change from [æ] to [ɛ]. To account for these and other findings (summarized in Polka & Bohn, 2003), Polka and Bohn (2011) proposed the Natural Referent Vowel

(NRV) framework, which aims to account for the privileged role of peripheral vowels in perception by infants and non-native listeners. Importantly, the NRV framework predicts that asymmetries observed with young infants are "lost" if the infant experiences the sound contrast in the ambient language (like German-learning children experience an [u]-[y] contrast), whereas the bias favoring peripheral vowels is maintained in the absence of specific contrast experience, even for adults (as for L1 English listeners' perception of [u]-[y]). The prediction that peripheral vowels retain their special status in non-native perception has been confirmed in several studies of adult cross-language perception (summarized in Polka & Bohn, 2011).

To summarize, the most important influence on difficulty and ease in cross-language and L2 perception derives from how sounds of the L2 are mapped unto L1 categories. However, non-native speech perception may also be affected by listeners' higher order generalizations about the role of, for example, lip rounding (Bohn & Best, 2012) or segment duration (Pajak & Levy, 2014) across different types of segments. In addition, L2-to-L1 mapping is shaped by universal biases which favor certain acoustic dimensions (e.g., duration, Bohn, 1995) or segments with certain acoustic/articulatory properties (as described in the NRV framework) prior to specific linguistic experience.

## The plasticity issue

The two models of L2 perception presented in some detail in this chapter, SLM and PAM-L2, both assume that perceptual learning is possible at all ages but will be influenced by the entire language learning history of the individual. This claim of life-long perceptual plasticity contrasts with the extension of Lenneberg's (1967) Critical Period Hypothesis to L2 learning, which assumes a developmental loss of neural plasticity. This loss would make it very hard, if not impossible, for adults to establish new perceptual categories for sounds of the L2, or to change already established categories to accommodate sounds of the L2.

Evidence for evaluating these competing claims comes either from studies of naturalistic L2 learning, in which participants acquire their L2 through exposure to and use of the L2 in their daily lives, or from perceptual training studies, in which typically naive listeners receive laboratory training on specific non-native contrasts, usually with feedback on their performance, which is assumed to shape their perception over several training sessions.

The vast majority of L2 perception studies with naturalistic learners supports the view that speech perception remains malleable over the life span, given that learners have had sufficient time and experience to realize their learning potential. A very strong argument in favor of plasticity comes from those studies of perceptual learning which have shown that sounds of the L2 differ in learnability for the same learner group (in terms of L1, experience with the L2, etc.). If perceptual learning were limited by maturation, this limitation should affect the full range of L2 sounds, i.e., in terms of SLM, both similar and new sounds. No study has shown such an across-the-board lack of plasticity, on the contrary (e.g., Flege *et al.*, 1997).

However, there are also studies that suggest limits on perceptual plasticity. Some of these studies have reached this conclusion because they confounded of age of learning and exposure to/experience with the L2 (e.g., Yamada & Tohkura, 1991a,b; but see Yamada, 1995). Comparing child learners who have resided in the L2 speech community for several years with adult learners who had much less experience does not address the plasticity issue because learning takes time. However, some studies which have avoided this confound have been interpreted to demonstrate a "striking lack of behavioral plasticity" (Pallier, Bosch, & Sebastian-Galles, 1997, p. B9). The evidence against plasticity comes primarily from studies of Spanish-dominant Catalan-Spanish bilinguals with extensive and early exposure to L2 Catalan. Interestingly, both the Pallier *et al.* 1997 study and a study on Catalan-Spanish bilinguals by Sebastian-Galles and Soto-Faraco (1999) showed great variability in the Spanish-dominant group, with some participants performing like the L1 Catalan participants. This suggests that the " … severe limitations to the malleability of the initially acquired L1 phonemic categories" (Sebastian-Galles & Soto-Faraco, 1999, p. 120) are not severe enough to affect all individuals. It would be of great interest to know how the Spanish-dominant Catalan-Spanish bilinguals who performed like the L1 Catalan participants differed in terms of learner characteristics from those bilinguals who performed worse. Finally, a series of studies by Dupoux and colleagues (Dupoux *et al.*, 1997; Dupoux, Peperkamp, & Sebastian-Galles, 2001; Dupoux *et al.*, 2008) could lead one to believe that they present evidence against perceptual plasticity in L2 learners. The titles of these three studies, which examined L2 stress perception, all contain the word "deafness," which suggests the irreversible loss of an ability. However, none of these studies was conducted with highly experienced L2 learners at or near the level of ultimate attainment, which makes their results irrelevant for any discussion of perceptual plasticity.

Training studies have greatly contributed to the plasticity issue by examining which procedural and stimulus characteristics contribute to perceptual learning. Typically, these studies involve at least two groups whose language background characteristics are equivalent. One group serves as the control (no treatment) group, which receives no training and participates only in a pretest (to establish the baseline) and a posttest (to establish whether simple retesting affected perceptual performance). The experimental group receives training between pre- and posttest, and this training may consist of either identification or discrimination of either synthetic or natural stimuli, produced by one or several talkers, in one or several contexts, with or without feedback in ca. 10 to 45 sessions which typically last 20-30 minutes each. Perceptual training studies have now been conducted for more than 30 years (since McClaskey, Pisoni, & Carrell, 1983; Strange & Dittman, 1984), and they have shown quite clearly which conditions lead to robust learning, i.e., learning that is not specific to the trained items or talkers, which is maintained over time after training, and which extends to production (for methodological aspects of perceptual training studies, see Logan & Pruitt, 1995): Robust learning is best achieved through a high variability training technique in which trainees are exposed to the full range of variability within each category, thus emulating what learners will encounter naturalistically in a new language environment. The bottom line of a large number of studies which have employed high variability training is nicely summarized in a review of studies on the behavioral and neurological effects of lexical tone training conducted by Wang and colleagues: "the adult brain retains a high degree of plasticity" (Sereno & Wang, 2007, p. 257; see also Bradlow *et al.*, 1997; Bradlow *et al.*, 1999; Lively, Logan, & Pisoni, 1993; Lively *et al.*, 1994; Logan, Lively, & Pisoni, 1991, 1993; Pisoni & Lively, 1995; Trapp & Bohn, 2002; Iverson, Hazan, & Bannister, 2005; Jongman & Wade, 2007, Rato, 2014).

To summarize, both studies of naturalistic learning and of laboratory training support the claim of current models of non-native speech perception that perceptual learning is possible at all ages which have been studied so far. This claim is further strengthened by studies of changes in speech production, no matter whether these changes occurred after short or long periods of exposure, and no matter whether they are due to perceptual adaptations to a non-native language or to changes in the native language variety (e.g., Sancier & Fowler, 1997; Chang, 2012; Flege *et al.*, 1997; Harrington, Palethorpe, & Watson, 2000). An important lacuna in research on plasticity in speech perception is the age group of mature adults (over the age of 40) because evidence supporting plasticity stems mostly from younger adults. The optimistic and realistic claim regarding intact learning mechanisms and processes over the life span should be further tested with older age groups, which have so far not been included in research on cross-language and L2 perception.

## Conclusion

The introduction to this chapter presented as the one overarching question in all research on cross-language and L2 speech perception, "What makes non-native sounds difficult to perceive?" The sections of this chapter presented overviews of how research has dealt with three issues that need to be resolved in order to address this question.

Research on the mapping issue—the perceptual relationship of sounds of the native and the non-native language—has shown that L2-to-L1 mapping can only be studied directly, through perceptual experiments examining the assimilation/mapping/interlingual identification of the sounds of the L2 and the L1. Other methods, such as the comparison of phonetic symbols or acoustic comparisons of the sounds of the L1 and the L2, are likely to yield results that are not valid. Research addressing the mapping issue has also contributed to a better understanding of the unit of speech perception in general by showing that cross-language perception occurs at the level of context-sensitive perceptual units (allophones), not at the abstract level of the phoneme, and that listeners have language-specific expectations about how coarticulation affects segment identity. Suggestions for future research include the role of visual information in L2-to-L1 mapping, and the question of whether this mapping is most accurately revealed in experiments examining ecphoric or perceptual similarity.

Research on the perceptual and learning difficulty/ease issue—how L1-to-L2 mapping may or may not cause perceptual and learning difficulty—has mainly been guided by two models of cross-language and L2 perception, the SLM and PAM(-L2). Although the predictions of these models have been supported by a large number of studies, other influences on perceptual and learning difficulty/ease than the mapping of segments or contrasts have been identified. Some of these influences are due to higher order generalizations about the sound systems of the L1 or the L2, and others are due to universal biases which listeners bring to the task of non-native speech perception irrespective of their L1. It is likely that future research will identify more of these two types of influences, and that this research will identify the conditions that allow higher order generalizations to affect perceptual and learning difficulty/ease.

Research on the plasticity issue—whether and how experience with the non-native language affects the perceptual organization of speech sounds in the mind of L2 learners—strongly suggests that perceptual reorganization, including the establishment of new categories for the sounds of the L2, is possible at any age studied so far. An important question of great applied and theoretical interest for future research concerns the generalizability of findings on the malleability of speech perception to the large and growing portion of mature adults.

## NOTES

1 The abbreviation "L2" is sometimes used to refer exclusively to a second language, but it will be used here to refer to any language learned in addition to the L1.

2 "Ecphory is a process by which retrieval information provided by a cue is correlated with the information stored in an episodic memory trace" (Tulving, 1983, p. 361).

3 PAM assumes that this happens on the basis of articulatory similarity.

4 Studies testing PAM have used different criteria for the "Uncategorized" assimilation: In Best, Faber, & Levitt (1996), "Uncategorized" was used for < 50% assimilations as instance of an L1 category, whereas Bundgaard-Nielsen, Best, and Tyler (2011) used the more stringent criterion of < 70% assimilations as instance of an L1 category.

5 Note that Norwegian "outrounded" [y] is front relative to IPA [y].

6 Other variables that are specific to the perceiver/learner such as integrative and instrumental motivation to learn the L2, attitude toward the language or the cultures, which are associated with the non-native language, and biological sex or social gender, have not been studied in any detail, and form no part of any current model of cross-language or L2 perception. Still, the influence of musical training and musical ability has received some attention (for an overview, see Gottfried, 2007), as has the role of talent in non-native speech perception (Jilka, 2009; Jilka, Lewandowski, & Rota, 2010).

# REFERENCES


Adank, P., Smits, R, & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America*, 116, 3099–3107. doi: 10.1121/1.1795335

Andruski, J. E., & Nearey, T. M. (1992). On the sufficiency of compound target specification of isolated vowels and vowels in/bVb/syllables. *Journal of the Acoustical Society of America*, 91, 390–410. doi: 10.1121/1.402781

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 167–200). Timonium MD: York Press.

Best, C. T., Bradlow, A. R., Guion-Anderson, S., & Polka, L. (2011). Using the lens of phonetic experience to resolve phonological forms. *Journal of Phonetics* 39, 453–455. doi: 10.1016/j.wocn.2011.08.006

Best, C. T., Faber, A., & Levitt, A. (1996). Assimilation of non-native vowel contrasts to the American English vowel system. *Journal of the Acoustical Society of America*, 99, 2602. doi: 10.1121/1.415316

Best, C. T., Hallé, P. A., Bohn, O.-S., & Faber, A. (2003). Cross-language perception of non-native vowels: Phonological and phonetic effects of listeners' native languages. *Proceedings of the 15th International Congress of Phonetic Sciences*, 2889–2892.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109, 775–794. doi: 10.1121/1.1332378

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 345–360. doi: 10.1037/0096-1523.14.3.345

Best, C. T., & Strange, W. (1992). Effects of language-specific phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20, 305–330.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Boas, F. (1889). On alternating sounds. *The American Anthropologist* 2, 47–53. doi: 10.1525/aa.1889.2.1.02a00040

Bogacka, A. (2004). On the perception of English high vowels by Polish learners of English. In E. Daskalaki, *et al.* (Eds.), *CamLing 2004: Proceedings of the University of Cambridge Second Postgraduate Conference in Language Research* (pp. 43–50). Cambridge, UK: Cambridge University Press.

Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 279–304). Timonium, MD: York Press.

Bohn, O.-S. (2002). On phonetic similarity. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 191–216). Trier: Wissenschaftlicher Verlag.

Bohn, O.-S. (2004). How to organize a fairly large vowel inventory: The vowels of

Fering (North Frisian). *Journal of the International Phonetic Association*, 34, 161–173. doi: 10.1017/S002510030400180X

Bohn, O.-S. (2013). Acoustic characteristics of Danish infant directed speech. *Proceedings of Meetings on Acoustics* 19, 060055. doi: 10.1121/1.4798488

Bohn, O.-S., & Best, C. T. (2012). Native-language phonological and phonetic influences on perception of English approximant contrasts by Danish and German listeners. *Journal of Phonetics*, 40, 109–128. doi: 10.1016/j.wocn.2011.08.002

Bohn, O.-S., Best, C. T., Avesani, C., & Vayra, M. (2011). Perceiving through the lens of native phonetics: Italian and Danish listeners' perception of English consonant contrasts. *Proceedings of the 17th International Congress of Phonetic Science*, 336–339.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11, 303–328. doi: 10.1017/S0142716400008912

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131–158. doi: 10.1017/S0272263100010792

Bohn, O.-S., & Garibaldi, C. (2017). In *Romance-Germanic bilingual phonology*, edited by M. Yavas, M. M. Kehoe, and W. C. Cardoso, 212–236. Equinox Publishing.

Bohn, O.-S., & Munro, M. J, eds. (2007). *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam and Philadelphia, PA: John Benjamins.

Bohn, O.-S., & Polka, L. (2001). Target spectral, dynamic spectral, and duration cues in infant perception of German vowels. *Journal of the Acoustical Society of America*, 110, 504–515. doi: 10.1121/1.1380415

Bohn, O.-S., & Polka, L. (2009). Accounting for the accented perception of vowels: Universal preferences and language-specific biases. *Journal of the Acoustical Society of America*, 125, 2758 doi: 10.1121/1.4784646.

Bohn, O.-S., & Steinlen, A. K. (2003). Consonantal context affects cross-language perception of vowels. *Proceedings of the 15th International Congress of Phonetic Sciences*, 2289–2292.

Bond, Z. S., & Fokes, J. (1991). Perception of English voicing by native and non-native adults. *Studies in Second Language Acquisition*, 13, 471–492. doi: 10.1017/S0272263100010299

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English/r/and/l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299–2310. doi: 10.1121/1.418276

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. I. (1999). Training Japanese listeners to identify English/r/and/l/: Long-term retention of learning in perception and production. *Perception and Psychophysics*, 61, 977–985. doi: 10.3758/BF03206911

Briére, E. (1968). *A psycholinguistic study of phonological interference*. The Hague: Mouton.

Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011a). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition*, 33, 433–461. doi: 10.1017/S0272263111000040

Bundgaard-Nielsen, R. L., Best, C. T., Kroos, C., & Tyler, M. D. (2011b). Vocabulary size matters: The assimilation of L2 Australian English vowels to L1 Japanese vowel categories. *Applied Psycholinguistics*, 32, 51–67. doi: 10.1017/S0142716410000287

Burnham, D., & Mattock, K. (2007). The perception of tomes and phones. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language*

*Speech Learning: In Honor of James Emil Flege* (pp. 239–258). Amsterdam: John Benjamins.
Butcher, A. (1976). The influence of the native language on the perception of vowel quality. *Arbeitsberichte 6, Institut für Phonetik Kiel University*, 1–137.
Calderon, J., & Best, C. T. (1996). Effects of bilingualism on non-native phonetic contrasts. *Journal of the Acoustical Society of America*, 99, 2602. doi: 10.1121/1.415315
Catford, J. C. (1968). Contrastive Analysis and Language Teaching. *Publications Department, School of Languages and Linguistics, Georgetown University, Washington, D.C. 20007 Monograph Series No. 21*.
Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel perception. *Journal of Phonetics*, 34, 372–387. doi: 10.1016/j.wocn.2005.08.003
Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, 40, 249–268. doi: 10.1016/j.wocn.2011.10.007
Collins, B., & Mees, I. M. (1984). *The Sounds of English and Dutch*. Leiden: Leiden University Press.
Darcy, I., & Krüger, F. (2012). Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics*, 40, 568–581. doi: 10.1016/j.wocn.2012.05.001
DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, 9, 133–150. doi: 10.1016/0163-6383(86)90025-1
Disner, S. F. (1978). *Vowels in Germanic languages*. Phonetics Laboratory, Department of Linguistics, University of California at Los Angeles.
Dooling, R. J. & Hulse, S. H, eds. (2014). *The comparative psychology of audition: perceiving complex sounds*. London, UK: Psychology Press.
Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A distressing 'deafness' in French? *Journal of Memory and Language*, 36, 406–421. doi: 10.1006/jmla.1996.2500
Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness". *Journal of the Acoustical Society of America*, 110, 1606–1618. doi: 10.1121/1.1380437
Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress "deafness": The case of French learners of English. *Cognition*, 106, 682–706. doi: 10.1016/j.cognition.2007.04.001
Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, 26, 551–585. doi: 10.1017/s0272263104040021
Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, 37, 452–465. doi: 10.1016/j.wocn.2009.07.006
Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36, 345–360. doi: 10.1016/j.wocn.2007.11.002
Flege, J. (1991a). The interlingual identification of Spanish and English vowels: Orthographic evidence. *Quarterly Journal of Experimental Psychology*, 43, 701–731. doi: 10.1080/14640749108400993
Flege, J. (1991b). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, 89, 395–411. doi: 10.1121/1.400473
Flege, J. (1992). The intelligibility of English vowels spoken by British and Dutch talkers. In R. D. Kent (Ed.), *Intelligibility on speech disorders* (pp. 157–232). Amsterdam: John Benjamins.
Flege, J. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues*

*in Cross-language Research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E., & Bohn, O.-S. (1989). The perception of English vowels by native speakers of Spanish. *Journal of the Acoustical Society of America*, 85, S85. doi: 10.1121/1.2027177

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25, 437–470. doi: 10.1006/jpho.1997.0052

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*, 106, 2973–2987. doi: 10.1121/1.428116

Flege, J. E., Munro, M. J., & Fox, R. A. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, 95, 3623–3641. doi: 10.1121/1.409931

Flege, J. E., Schmidt, A. M., & Wharton, G. (1996). Age of learning affects rate-dependent processing of stops in a second language. *Phonetica*, 53, 143–161. doi: 10.1159/000262195

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of/r/and/l/. *Journal of the Acoustical Society of America*, 99, 1161–1173. doi: 10.1121/1.414884

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41, 78–104. doi: 10.1006/jmla.1999.2638

Garibaldi, C., & Bohn, O.-S. (2015). Phonetic similarity predicts ultimate attainment quite well: The case of Danish/i, y, u/and/d, t/for native speakers of English and of Spanish. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, 10-14 August 2014.

Gottfried, T. L. (2007). Music and language learning: Effects of musical training on learning L2 speech contrasts. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 221–237). Amsterdam: John Benjamins. doi: 10.1075/lllt.17.21got

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107, 2711–2724. doi: 10.1121/1.428657

Hale, H. (1885). On some doubtful or intermediate articulations: An experiment in phonetics. *Journal of the Anthropological Institute of Great Britain and Ireland*, 14, 233–243. doi: 10.2307/2841981

Hallé, P. A., Best, C. T., & Levitt, A. (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics*, 27, 281–306. doi: 10.1006/jpho.1999.0097

Hallé, P. A., Chang, Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Chinese versus French listeners. *Journal of Phonetics*, 32, 395–421. doi: 10.1016/S0095-4470(03)00016-0

Hardison, D. M. (1999). Bimodal speech perception by native and non-native speakers of English: Factors influencing the McGurk effect. *Language Learning*, 49, 213–283.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, 24, 495–522.

Hardison, D. M. (2005). Second-language spoken word identification: Effects of perceptual training, visual cues, and phonetic environment. *Applied Psycholinguistics*, 26, 579–596.

Harnsberger, J. D. (2000). A cross-language study of the identification of non-native nasal consonants varying in place of articulation. *Journal of the Acoustical*

*Society of America*, 108, 764–783. doi: 10.1121/1.429610

Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *Journal of the Acoustical Society of America*, 110, 489–503. doi: 10.1121/1.1371758

Harrington, J., Palethorpe, S., & Watson, I. C. I. (2000). Does the queen speak the Queen's English? *Nature*, 408, 927–928. doi: 10.1038/35050160

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*, 28, 377–396. doi: 10.1006/jpho.2000.0121

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, 47, 360–378. doi: 10.1016/j.specom.2005.04.007

Højen, A. (2002). Short-term immersion can change the location of a phonetic category. *Journal of the Acoustical Society of America*, 112, 2388. doi: 10.1121/1.4779722

Horslund, C. S., Ellegaard, A. A., & Bohn, O.-S. (2015). Perceptual assimilation and identification of English consonants by native speakers of Danish. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, 10–14 August 2014.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *Journal of the Acoustical Society of America*, 118, 3267–3278. doi: 10.1121/1.2062307

Iverson, P., Kuhl, P. K., Akahane-Yamada, T., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87, B47–B57. doi: 10.1016/s0010-0277(02)00198-1

Jilka, M. (2009). Talent and proficiency in language. In G. Dogil & S. M. Reiterer (Eds.), *Language Talent and Brain Activity* (pp. 1–16). Berlin: Mouton De Gruyter.

Jilka, M., Lewandowski, N., & Rota, G. (2011). Investigating the concept of talent in phonetic performance. In M. Wrembel et al. (Eds.), *Achievements and Perspectives in SLA of Speech: New Sounds 2010* (pp. 171–180). Frankfurt am Main: Peter Lang.

Johannson, S. (1975a). Swedish and English phonemes: A perceptual study. *Lund Studies in English*, 50, 60–86.

Johannson, S. (1975b). Swedish and English stops: A perceptual study. *Lund Studies in English*, 50, 87–100.

Jongman, A. & Wade, T. (2007). Acoustic variability and perceptual learning. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 135–150). Amsterdam: John Benjamins. doi: 10.1075/lllt.17.14jon

Kewley-Port, D., Bohn, O.-S., & Nishi, K. (2005). The influence of different native language vowel systems on vowel discrimination and identification. *Journal of the Acoustical Society of America*, 117, 2399. doi: 10.1121/1.4785992

Ko, E.-S., Soderstrom, M., & Morgan, J. (2009). Development of perceptual sensitivity to extrinsic vowel duration in infants leaning American English. *Journal of the Acoustical Society of America*, 126, EL134–EL139. doi: 10.1121/1.3239465

Kondaurova, M. V. & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America*, 124, 3959–3971. doi: 10.1121/1.2999341

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical*

*Transactions of the Royal Society B*, 363, 979–1000. doi: 10.1098/rstb.2007.2154

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, F13–F21. doi: 10.1111/j.1467-7687.2006.00468.x

Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.

Levy, E. S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *Journal of the Acoustical Society of America*, 125, 1138–1152. doi: 10.1121/1.3050256

Levy, E. S. & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, 36, 141–157. doi: 10.1016/j.wocn.2007.03.001

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English/r/and/l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242–1252. doi: 10.1121/1.408177

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English/r/and/l/. III: Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, 96, 2076–2087. doi: 10.1121/1.410149

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l/: A first report. *Journal of the Acoustical Society of America*, 89, 974–886. doi: 10.1121/1.1894649

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1993). Training listeners to perceive novel phonetic categories: How do we know what is learned? *Journal of the Acoustical Society of America*, 94, 1148–1151. doi: 10.1121/1.406963

Logan, J. S., & Pruitt, J. S. (1995). Methodological issues in training listeners to perceive non-native phonemes. In W. Strange (Ed.), Speech Perception and Linguistic Experience: Issues in Cross-language Research (pp. 351–377). Timonium, MD: York Press.

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English/r/and/l/by Japanese bilinguals. *Applied Psycholinguistics*, 2, 369–390. doi: 10.1017/S0142716400009796

Major, R. C., & Kim, E. (1996). The Similarity Differential Rate hypothesis. *Language Learning*, 46, 465–596. doi: 10.1111/j.1467-1770.1996.tb01244.x

McClasky, C. L., Pisoni, D. B., & Carrell, T. D. (1983). Transfer of training of a new linguistic contrast in voicing. *Perception and Psychophysics*, 34, 323–330. doi: 10.3758/BF03203044

Mennen, I. & de Leeuw, E. (2014). Beyond segments. *Studies in Second Language Acquisition*, 36, 183–194. doi: 10.1017/S0272263114000138

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics*, 18, 331–340. doi: 10.3758/BF03211209

Moon, C., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: a two-country study. *Acta Paediatrica*, 102, 156–160. doi: 10.1111/apa.12098

Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: Visual articulatory information enables perception of second language sounds. *Psychological Research*, 71, 4–12. doi: 10.1007/s00426-005-0031-5

Nishi, K., Strange, W., Akahane-Yamada, R., Kubi, R., & Trent-Brown, S. A. (2008. Acoustic and perceptual similarity of Japanese and American English vowels.

*Journal of the Acoustical Society of America*, 124, 576–588. doi: 10.1121/1.2931949
Pajak, B. & Levy, R. (2014). The role of abstraction in non-native speech perception. *Journal of Phonetics*, 46, 147–160. doi: 10.1016/j.wocn.2014.07.001
Pallier, C., Bosch, L., & Sebastian-Galles, N. (1997). A limit on behavioral plasticity in speech perception acquisition. *Cognition*, 64, B9–B17. doi: 10.1016/S0010-0277(97)00030-9
Pisoni, D. B. & Lively, S. E. (1995). Variability and invariance in speech perception : A new look at old problems in perceptual learning. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 433–459). Timonium, MD: York Press.
Polivanov, E. D. (1931). La perception des sons d'une langue étrangère. *Travaux du cercle linguistique de Prague*, 4, 79–96.
Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America*, 89, 2961–2977. doi: 10.1121/1.400734
Polka, L. (1992). Characterizing the influence of native language experience on adult speech perception. *Perception and Psychophysics*, 52, 37–52. doi: 10.3758/BF03206758
Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America*, 95, 1286–1296. doi: 10.1121/1.412170
Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, 100, 577–592. doi: 10.1121/1.415884
Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41, 221–231. doi: 10.1016/S0167-6393(02)00105-X
Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, 39, 467–478. doi: 10.1016/j.wocn.2010.08.007
Rato, A. A. dos Santos. (2014). *Cross-language Perception and Production of English Vowels by Portuguese Learners: The Effects of Perceptual Training*. Unpublished PhD thesis, University of Minho, Braga.
Rauber, A. S., Escudero, P., Bion, R. A. H., & Baptista, B. O. (2005). The interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese. *Interspeech*, 2005, 2913–2916.
Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 379–410). Timonium, MD: York Press.
Ryan, E. B. (1983). Social psychological mechanisms underlying native speaker evaluations of non-native speech. *Studies in Second Language Acquisition*, 5, 148–159. doi: 10.1017/S0272263100004824
Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421–436. doi: 10.1006/jpho.1997.0051
Schmidt, A. M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *Journal of the Acoustic Society of America*, 99, 3201–3211. doi: 10.1121/1.414804
Schmitt, H. (2007). The case for the epsilon symbol (ɛ) in RP DRESS. *Journal of the International Phonetic Association*, 37, 321–328. doi: 10.1017/S0025100307003131
Sebastián-Gallés, N., & Soto-Franco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, 72, 111–123. doi: 10.1016/S0010-0277(99)00024-4
Sereno, J. A., & Wang, Y. (2007). Behavioral and cortical effects of learning a second language. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second*

*Language Speech Learning: In Honor of James Emil Flege* (pp. 239–258). Amsterdam: John Benjamins. doi: 10.1075/lllt.17.22ser

So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, 53, 273–293. doi: 10.1177/0023830909357156

So, C. K., & Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, 36, 195–221. doi: 10.1017/S0272263114000047

Strange, W., ed. (1995). *Speech perception and linguistic experience: Theoretical and methodological issues*. Timonium, MD: York Press.

Strange, W. (2007). Cross-language phonetic similarity of vowels. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 35–55). Amsterdam: John Benjamins. doi: 10.1075/lllt.17.08str

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39, 456–466. doi: 10.1016/j.wocn.2010.09.001

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, 26, 311–344.

Strange, W. & Bohn, O.-S. (1998). Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies. *Journal of the Acoustical Society of America*, 104, 488–504. doi: 10.1006/jpho.1998.0078

Strange, W., Bohn, O.-S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America*, 115, 1791–1807. doi: 10.1121/1.1687832

Strange, W., Bohn, O.-S., Nishi, K. & Trent, S. A. (2005). Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America*, 118, 1751–1762. doi: 10.1121/1.1992688

Strange, W. & Dittman, S. (1984). Effects of discrimination training on the perception of/r-l/by Japanese adults. *Perception and Psychophysics*, 36, 131–145. doi: 10.3758/BF03202673

Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *Journal of the Acoustical Society of America*, 122, 1111–1129. doi: 10.1121/1.2749716

Sundara, M., Polka, L., & Genesee, F. (2006). Language-experience facilitates discrimination of/d-ð/in monolingual and bilingual acquisition of English. *Cognition*, 100, 369–388.

Trapp, N. L, & Bohn, O.-S. (2002). Training Danish listeners to identify word-final/s/and/z/: Generalization of training and its effect on production accuracy. In A. James & J. Leather (Eds.), *New Sounds 2000: Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech* (pp. 343–350). Klagenfurt: University of Klagenfurt.

Trubetzkoy, N. S. (1939). Grundzüge der Phonologie. *Traveaux de Cercle Linguistique de Prague VII*.

Tulving, E. (1981). Similarity relations in recognition. *Journal of Verbal Learning and Verbal Behavior*, 20, 479–496. doi: 10.1016/S0022-5371(81)90129-8

Tulving, E. (1983). Ecphoric processes in episodic memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 302(1110), 361–371. doi: 10.1098/rstb.1983.0060

Wang, Y., Behne, D. M., & Jiang, H. (2008). Linguistic experience and audio-visual perception of non-native fricatives.

*Journal of the Acoustical Society of America*, 124, 1716–1726. doi: 10.1121/1.2956483

Werker, J. F., Frost, P. E., & McGurk, H. (1992). La langue et les lèvres: Cross-language influences on bimodal speech perception. *Canadian Journal of Psychology*, 46, 551–568. doi: 10.1037/h0084331

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, 75, 1866–1878. doi: 10.1121/1.390988

Wu, X., Munro, M., & Wang, Y. (2014). Tone assimilation by Mandarin and Thai listeners with and without L2 experience. *Journal of Phonetics*, 46, 86–100. doi: 10.1016/j.wocn.2014.06.005

Yamada, R. A. (1995). Age and acquisition of second language speeh sounds: Perception of American English/ɹ/ and/l/by native speakers of Japanese. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 305–320). Timonium, MD: York Press.

Yamada, R. A., & Tohkura, Y. (1991a). Age effect on acquisition of non-native phonemes: Perception of English/r/ and/l/for native speakers of Japanese. *Proceedings of the 12th International Congress of Phonetic Sciences*, Vol. 4, 450–453.

Yamada, R. A., & Tohkura, Y. (1991b). Perception of American English/r/ and/l/for native speakers of Japanese. In Y. Tohkura et al., (Eds.), *Speech Perception, Production and Linguistic Structure* (pp. 155–174). Tokyo: OHM.

# 11 Models of Lexical Access and Morphological Processing

PETAR MILIN[1], EVA SMOLKA[2], AND LAURIE BETH FELDMAN[3]

[1] The University of Sheffield
[2] University of Konstanz
[3] The University at Albany, State University of New York, and Haskins Laboratories

## Introduction

Humans possess the ability to discriminate among meanings for a practically unlimited number of possible utterances. This ability to make and understand new combinations defines productivity and has been extensively studied with respect to combining words into sentences. However, an analogous ability exists at the level of combining morphemes into words.[1] This is the essence of a *combinatorial* approach to language phenomena. Its point of departure is an assumption that language is a hierarchical system, where at each level in the hierarchy one can find distinct units, each of which makes an "imprint" in the mind, that is, brain.

In the combinatorial tradition, many studies in psycholinguistics use words composed of multiple morphemes, defined as morphologically complex, that have an intriguing property that their "meaning" can range from completely semantically compositional or "transparent," as in SUCCESS-FUL to relatively "opaque," as in SUCCESS-OR.[2] Given that the meaning of semantically opaque words cannot be derived from the meaning of their parts, linguistic, psycholinguistic, and neurolinguistic researchers, who adhere to the combinatorial tradition, are intrigued by questions such as: (i) how are the units at each level in the hierarchy represented and organized in the mental lexicon (e.g., are words stored as whole-word or as constituent units), and (ii) whether their combinatorics entail linguistic patterning (syntax) and/or a more general mechanism of association?

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

We can compare that approach to learning-based models where the emphasis is on how existing knowledge guides what functions as relevant units and how, with subsequent learning, those early units get better-differentiated one from another. In essence, learning-based models of word recognition emphasize progressive differentiation of larger chunks into smaller units and then use those to control the uncertainty in communication (see, e.g., Ramscar & Baayen, 2013; Ramscar & Port, 2015). Further, depending on what is already known, new relations can be easy or difficult to learn depending on how easy it is to differentiate them as distinct (c.f., Ellis, 2006; Milin, Feldman, Ramscar, Hendrix, & Baayen, 2017).

In the present chapter, we sample from the experimental literature using the lexical decision task with masked and unmasked priming techniques to investigate how we understand morphologically complex words and discuss how those findings inform both lexicon-based and learning-based models of word knowledge (the mental lexicon). Then we focus on behavioral and electrophysiological measures on derived words to discuss experimental findings pertaining to the processing of morphologically complex words with and without so-called semantic transparency. How best to capture knowledge about the patterns that form complex words and the extent to which the underlying process is specific to language or particular types of linguistic units is a theme that recurs throughout the domain of psycholinguistics. Discussion is particularly animated in the domain of morphology, therefore, word recognition for morphologically complex words is where we concentrate.

Languages can differ in the way in which morphemes combine and in the complexity of those patterns. In Indo-European languages with concatenative morphology like English, French, and German, morphemes are appended to one another in a more or less, linear fashion. In the combinatorial tradition, again, the creation of new words from existing stems entails attaching either prefixes or suffixes to a stem that captures the core meaning of the word, as in UN + SUCCESS + FUL + LY. In contrast, in Semitic languages like Hebrew and Arabic, words consist of two abstract morphemes, the root and the word pattern, that are intertwined one within the other. The root conveys semantic aspects of the word, while the word pattern carries phonological and morpho-syntactic information. For example, the root ZMR (relating to "sing") can be inserted into different word patterns to derive a large family of Hebrew words like ZiMeR ("sing"), hiZdaMeR ("sounded like a song"), ZaMaR ("singer"), ZeMeR ("song"), ZiMRa ("singing"), and tiZMoRet ("orchestra"). The way the root and word pattern combine is nonlinear in that the letters of the root are no longer contiguous. Some have claimed that the nonlinear (nonconcatenative) manner in which morphemes combine in a language such as Hebrew affects the way a word is processed and represented (e.g., Frost, Kugler, Deutsch, & Forster, 2005).

Inflection and derivation are two well-studied processes to form morphologically complex words. Inflectional affixes specify, among other small changes, the number or tense of an event, as in CATS or WALKED. That is, they introduce only minimal changes to the content (i.e., traditional "meaning") of the communication (e.g., from one to more in CAT-CATS, and from present to past tense in

WALK-WALKED). Derivational affixes form new words and occur in Indo-European languages like English, French, and German as prefixes or suffixes that append to base words. For example, the suffix -MENT often derives a noun that reflects the process or state of the underlying verb (e.g., DEVELOPMENT, RESENTMENT, ENLARGEMENT). In comparison to inflected forms, the semantic relation between derivations and their base is not as straightforward. For example, the suffix -MENT sometimes produces abrupt changes in the message, as from DEPART to DEPARTMENT.

Below we contrast lexicon-based (i.e., combinatorial) and learning-based approaches to the processing of prefixed and suffixed derived words. The lexicon-based approach to word processing follows in the combinatorial tradition where lexical entries like {WALK} carry the core meaning of a stem, from which further words and meanings are derived by attaching affixes like ER or ABLE. Within the learning-based approach we can distinguish parallel-distributed connectionist models (such as the Triangle model of Harm & Seidenberg, 2004), and discrimination learning models (like the Naive Discrimination Learning model—NDL of Baayen, Milin, Filipović Đurđević, Hendrix, & Marelli, 2011). The later make use of discrimination to understand and explain the dynamics of language, where, in essence, language is learned (acquired) through discrimination of contrastive properties rather than by the "content" of representations. In other words, for the NDL approach invariant meanings cannot be *encrypted* in utterances, and they cannot be mapped combinatorially (or compositionally) onto linguistic units at different levels of granularity. Instead, the essential form to meaning mapping is abstract and relational: elaborated meaning emerges only across communicative contexts (c.f., Ramscar & Baayen, 2013; Milin, Feldman, Ramscar, Hendrix, & Baayen, 2017). The NDL framework does not assume that lexical entries carry an invariant core of meaning, such as a core meaning for WALK plus a modification for past tense in ED. Rather, the listener's uncertainty about the speaker's intentions is reduced or even eliminated when she hears *John walked me to the train station*, and not *John walked me through his new book*. With respect to morphology, an appreciation of the variable form of morphologically related words such as RUN and RAN or RUNNER and RUNNABLE derives from the similarity of the contexts in which they appear relative to those in which, for example, WALK and WALKED or WALKER and WALKABLE are possible.

## Methodological issues

### *The lexical decision task*

Much of what we know about morphological processing comes from the outcomes of experiments where participants perform the lexical decision task. In this task, participants judge the lexical status of the letter string presented as the "target" (i.e., they make a "lexical decision" whereby they determine whether the letter string is a real word) and latencies to reach a decision along with judgment accuracy to each target constitute the typical dependent measures. In recent years, more and more often, electrophysiological brain responses (e.g., EEG) to "lexical events" supplement decision latencies and accuracies.

Many describe response latency as encompassing the time it takes to access a word in the lexicon and to ascertain that the lexical entry corresponds to the word that was presented (e.g., Rueckl & Galantucci, 2005). Interpretation of the data provides useful insights into a reader's lexical knowledge with an emphasis on words and how they are represented. Researchers introduce manipulations of context or presentation conditions and those contrasts help them to understand the underlying processes. For example, most typically targets appear after a single word that is called a *prime*.

The contrast of interest in any primed lexical decision task is the same target in the context of a related and an unrelated prime. For example, how does the determination of lexical status for the target SUCCESS differ depending on whether the prime is related, like SUCCESSFUL, or unrelated, like NEGLECTFUL? Any facilitatory or inhibitory effect (i.e., faster versus slower decision latencies relative to the unrelated prime) could occur because SUCCESSFUL and SUCCESS share form, meaning, or both. Hence, to differentiate between form and meaning, a further type of prime, like SUCCESSOR, would be introduced, that shares form but not meaning with the target SUCCESS. When there are multiple types of related primes, typically, "between-target" designs are adopted (e.g., Rastle, Davis, Marslen-Wilson, & Tyler, 2000) in which each target appears with one related and one unrelated prime and different targets appear with different types of related primes. For example, the target SUCCESS would appear with the semantically transparent prime SUCCESSFUL (and a matched unrelated prime), while the target DEPART would appear with the semantically opaque prime DEPARTMENT (and a matched unrelated prime). The typical way to control for differences targets in these experiments has been to eliminate significant differences between condition means for attributes such as word frequency and number of words that are similar in written (orthographic) or spoken (phonological) form in the various conditions and then to compare target decision latencies and accuracy judgments across conditions.

An alternative design applies a "within-target" manipulation and presents the same target in multiple related prime contexts. For example, the target SUCCESS would appear to different participants, with a prime that is transparent and thus semantically related (SUCCESSFUL), with a prime that is opaquely related (SUCCESSOR) and with a prime that is unrelated to the target word (NEGLECTFUL or MANAGER; for details, consult Feldman, 2000). The obvious advantage of comparing decision latencies to the *same* target across the different related prime contexts is that individual (distributional) differences among target words can be better controlled so as to avoid potential confounding of target properties with prime type. In principle, this confounding issue can also be controlled by statistical means. Rigorous implementations are infrequent, however (but see Feldman, O'Connor, & Moscoso del Prado Martín, 2009; Milin *et al.*, 2017).

## *Prime presentation*

Presentation duration for the prime and the presence or absence of a pattern mask further define the processing condition under which a participant makes a lexical decision to the target. Prime-target pairs can be presented visually or auditorily or in a cross-modal format. Particularly well-explored presentation conditions for the prime include manipulations of exposure duration of the prime and the presence or absence of pattern mask that precedes it.

## *The masked priming procedure*

The forward-masked lexical decision task is the most well accepted procedure to study early phases of word recognition. Here, a pattern mask appears for about 500 ms, then the prime appears in a lower-case equal-spaced font for about 48 ms, and finally a target printed in upper case is visible for 500 ms or more (Forster, Davis, Schoknecht, & Carter, 1987). For some, prime durations shorter than about 70 ms after a pattern mask meet the experimental conditions that qualify as "early" processing in this task. As is typical in other variants of the lexical decision task, latency, accuracy, and electrophysiological brain responses (EEG) to judge the lexicality of the target are measured.

Proponents of the forward-masked priming task argue that because the prime is not consciously perceived, participants cannot used it for strategic processing such as anticipating the upcoming target (Forster *et al.*, 1987). However, strategic effects can arise even in the forward-masked lexical decision task. For example, even at prime durations of 48 ms with a mask, the proportion of related prime-target pairs influences the magnitude of the difference in decision latencies between targets after related and unrelated primes. Facilitation increases as the proportion of related trials increases (Feldman & Basnight-Brown, 2008; Bodner & Masson, 2003). The implication is that processing of a masked prime can be informative about the orthographic and semantic properties of the target even though the prime is not available for conscious processing because of the mask. Despite its limitations, the forward-masked priming task provides the evidence for much of what we know about how morphologically complex words are understood during early stages of the word recognition process.

## *Overt priming*

In contrast to masked priming, unmasked or overt priming taps into lexical processing and representation in addition to early recognition. The prime is presented either auditorily or visually at a relatively long exposure duration (230 ms or longer) and is presented either immediately preceding the target (i.e., "immediate repetition priming") or with other items intervening between prime and target (i.e., "long lag priming"). In either case, the long exposure duration for the prime allows that it be consciously perceived. As a result, this procedure is vulnerable to conscious and strategic processing. One means to attenuate the anticipation of a particular target is to drastically reduce the proportion of related prime-target pairs (e.g., 25%) in the set of experimental materials (Napps & Fowler, 1987). Because processing is conscious, overt priming is used to examine not only lexical processing but also deeper, integrative processes across words. As in masked priming, most often response latency, accuracy and brain responses to targets (and sometimes to primes) are measured. In both masked and overt priming techniques, the critical measure refers to the difference between the latency/accuracy/brain response to targets following unrelated vs. related primes, this difference is called the priming effect.

## Lexicon-based and learning-based models of word recognition

The productivity of a language derives from the numerous ways in which a finite set of units can combine to form a new message. As noted above, productivity is characteristic not only at the level of combining words into sentences but also at the level of combining morphemes into words. For example, we can easily coin new words by combining morphemes like UNTAPABLE, or by combining words as in the compound SCHOOLHOUSE ROOF COLOR. Models of word recognition diverge on the question of how to represent morphologically complex words and the units from which they are composed and on the potential role that morphological rules play in describing the ways in which units combine. Likewise, models differ as to whether knowledge about morphemes is explicitly represented as lexical knowledge (Butterworth, 1983; Taft & Forster, 1975). Models further differ with respect to whether or not they require language-specific morphological rules to account for differences between experiments conducted with materials from English and from other languages, where the overall tendency to combine morphemes productively or the manner for forming those combinations (such as linear concatenation) differs from that of English (Frost, 2012; Feldman & Moscoso del Prado Martin, 2012; Smolka & Libben, 2017; Smolka, Preller, & Eulitz, 2014).

Below we review several basic types of models. These models differ with respect to the role granted to lexical representations and rules versus learning and the discovery of statistical patterning. Crucially, however, some of them are representative of a combinatorial and others of a discrimination tradition. One obvious manifestation of these differences is the contribution of whole-word units as contrasted with morpheme units and the concomitant role of rules to describe if and how morpheme units combine to form whole words.

Most well accepted are the *lexicon-based models* that assume the explicit representation of morphemes and of a word's morphological structure. Historically, dual-mechanism accounts of lexical representation have been particularly influential, especially in the domain of morphological processing. Their advocates often describe morphological knowledge in terms of a default option that provides for a rule-governed computation of complex linguistic forms from symbols, accompanied by storage of whole-word forms arrayed in lexical space when rules cannot succeed. The most distinct alternative are the *learning-based models.* These do not assume the explicit representation of morphemes, and describe "morphological effects" in terms of patterns of form and meaning that are processed conjointly. Nevertheless, even models within the general learning-based framework differ to a considerable extent in how they define and implement non-symbolic input/output representations. On the one hand, the major parallel-distributed processing models tend to covertly introduce item-specific lexical events, when they are trained on words presented in isolation and one at a time. This, for example, applies to the triangle model (Harm & Seidenberg, 2004). On the other hand, models such as the Naive Discrimination Reader (NDR; Baayen, Milin, Filipović Đurđević, Hendrix, & Marelli, 2011), build from sublexical form units, such as the letter bigrams and/or trigrams that occur in a particular language. As for the outcomes, the model introduces a hypothetical construct of *lexomes*—word-like units that represent neither word forms nor word meanings, but contribute to meaning in relation with other lexomes in the context (c.f., Milin *et al.*, 2017).

Broadly speaking, most of the learning-based models implement a single mechanism, characterized by graded effects whose underlying activation dynamics are based on mapping input to output similarities. In these frameworks, although both prime-target pairs are matched on form similarity, the pair FELL-FALL appears to be more similar in terms of the contexts in which the two appear than is BELL-BALL. The consequence is a greater difference in reaction times relative to a control word paired with the same target. In the same sense, decision latencies to FELL-FALL differ more from FULL-FALL than do BELL-BALL from BULL-BALL in this task.

Lexicon-based accounts focus on the role of rules and representations, and proponents reason that because formation of the inflected form FELL does not follow from a simple rule applied to FALL, it must be represented and greater facilitation for pairs such as FELL-FALL than for BELL-BALL must reflect something about how the lexical representations for the two words in a pair reference each other (Crepaldi, Rastle, Coltheart, and Nickels, 2010). Both Parallel-distributed processing and Naive Learning accounts, focus on what makes learning easy or hard and would thus quantify the processing cost derived from the systematicity of the mapping between form and meaning when trying to differentiate among words whose forms partially overlap with FALL as compared with BALL. Between the extremes, models vary along a continuum from purely rule-based, to fully probabilistic and inferential.

## *Dual-mechanism accounts*

The dual-mechanism class of lexical models for the recognition (or production) of words posits two independent mechanisms associated with different brain areas (Marslen-Wilson & Tyler, 1998; Pinker, 1999; Pinker & Ullman 2002, 2003; Silva & Clahsen 2008; Ullman, 2001, 2004) where the choice of mechanism depends on how adequately rules can describe word formation).[3] In the dual mechanism framework, differences between languages are captured in terms of different types of rules and/or the symbolic representations over which they apply. If morphological formations in a language follow a general pattern, so that word forms are compositional and can be described by rules, then word specific morphological structure need not be stored with each lexical entry. Past tense formation is the traditional domain of investigation where forms such as WALKED and HUMMED can be described by a rule that operates on a stem (WALK, HUM) and affix(es) such as ED, the inflection for past tense. However, for words that are irregular in that they cannot be formed or decomposed by rule, the incorrect application of the rule would produce the overgeneralized forms SPEAKED and RUNNED. This is the justification for a second non-combinatorial mechanism based on associations among full word representations; its function is to store the exceptions to the rules like SPEAK-SPOKE and RUN-RAN.

This type of dual-mechanism model further assumes different processing mechanism for morphological versus semantic processing. By dual-mechanism accounts, morphological facilitation from a regularly formed prime to its morphologically related target arises when the prime is decomposed into stem and affix and the stem of the prime preactivates the target. This results in faster recognition (facilitation) when it is time to recognize that target than when there is no preactivation. Thus ARTIST is decomposed into ART (stem) + IST (affix) and activation of the stem of the prime benefits the target ART. Similarly WALKED is decomposed into the stem WALK and the affix ED and its stem preactivates the target WALK. In this dual-mechanism framework, the mechanism that produces (regular) morphological facilitation entails decomposition and it differs from that for irregular inflection such as RUN-RAN or semantically related words such as CRAFT-ART where there is no shared stem to preactivate. In those cases, activation spreads from the whole prime word to the whole word target.

One challenge to the dual-mechanism account comes from the "word frequency effect", a comparison of words that vary on how frequently they occur: more frequent words are faster to recognize and faster to produce than less frequent words. A classical dual-mechanism interpretation of the whole word frequency effect in tasks such as lexical decision emphasizes access or activation of forms that are stored in the lexicon. According to the original dual-mechanism model whereby regular inflections are stored in terms of their stem while irregulars are stored in the mental lexicon as full words, the difference between high and low frequency words should be larger for irregularly than for regularly inflected forms in recognition tasks such as lexical decision (Alegre & Gordon, 1999) and production (Budd, Paulmann, Barry, & Clahsen, 2013). Frequency effects for regularly inflected forms thus pose a challenge to the dual-mechanism model (e.g., Baayen, Wurm, & Aycock, 2007).

A further test of the dual-mechanism account comes from nonword priming, where nonwords are formed from an illegal combination of existing stems and affixes like TAUGHTEN or SONGED. Dual-mechanism accounts assume that irregular verb forms are stored as whole word units; hence then these irregular stems (participle stems) should not be represented in the lexicon (Clahsen, Prüfert, Eisenbeiss, & Cholin, 2002). Facilitation from nonwords with irregular stems like

GE+SUNG+T or GE+WURF+T in German (analogs in English include SONG+ED or THREW+N) indicate that irregular and semi-regular stems can function similarly to regular verb stems and thus seriously challenge some assumptions of the dual-mechanism models (Smolka, Zwitserlood, & Rösler, 2007).

## *Single mechanism accounts*

Single mechanism learning models of morphological processing differ from dual-mechanism accounts in that they posit just one mechanism and that mechanism is sensitive to the frequency and sequential patterning of units in everyday language. Therefore, the foundation of Single Mechanism Accounts is the statistical structure that is present in language rather than distinct mechanism(s) that operate on the symbolic representations for words or the rules that operate on them. The basis of this framework is a serious treatment of the systematic mapping between form and meaning that characterizes the many words that share a base morpheme (Bybee, 1985, 1995; Bybee & McClelland, 2005). For example the form-meaning mapping is stronger in (SALT, SALTY, SALTINE, SALTSHAKER) than in (CALM, CALMNESS, CALMLY) because of the number of different words formed from the stem "morphological family size" (De Jong, Schreuder, & Baayen, 2000).

The most familiar option in the single mechanism research framework is parallel-distributed models (PDP: Gonnerman, Seidenberg, & Andersen, 2007; Kielar, Joanisse, & Hare, 2008; Joanisse & Seidenberg, 1999; Plaut & Gonnerman, 2000; Rueckl & Raveh, 1999; Seidenberg & Gonnerman, 2000, etc.). These models are also known as the connectionist models. The framework permits activation from the systematic mappings between form and meaning to vary in degree and to converge for noncompositional irregulars like RAN and SPOKE and FELL as well as for compositional regulars like WALKED and HUMMED. Stated alternatively, in a single mechanism framework, non-compositional irregulars and compositional regulars vary in degree not in type of morphological processing mechanism. Indeed, electrophysiological brain responses to regular (default), semi-irregular, and completely irregular participles in German have been observed to vary in degree and not in an all-or-none fashion (Smolka, Khader, Wiese, Zwitserlood, & Rösler, 2013).

At their core, PDP accounts are non-symbolic in nature and thus do not refer to stored representations or rules. Rather the organization emerges from distributed patterns of connectivity so that similarity patterns that encompass semantics, orthography and phonology contribute in a graded manner to the recognition and the production of all inflected forms. In the system, contributions of (constraints related to) task and more permanent differences between words are captured in a graded manner. Differences in facilitation in a priming study reflect the underlying dynamics of a system along with its earlier resting state or initial conditions (Rueckl, 2002). There is no shifting between mechanisms. For example, the point of departure for a PDP account of the apparent effect of inflectional regularity would focus on the consequences of greater orthographic overlap between prime and target for regularly than for irregularly inflected verbs (e.g., Bird *et al.*, 2003; Bybee

& McClelland, 2005; Patterson, Lambon Ralph, Hodges, & McClelland, 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996). Also relevant is greater semantic connectedness among irregularly than regularly inflected verbs (devoid of any prime) as this influences the starting point for recognition (Baayen & Moscoso del Prado Martin, 2005).

The finding that in native speakers irregularly inflected verb forms with high form overlap (e.g., DRAWN-DRAW) show facilitation that is comparable to regular verb forms (e.g., GUIDED-GUIDE) and not to change stem irregulars (e.g., FELL-FALL; in Basnight-Brown, Chen, Shu, Kostić, & Feldman, 2007) is consistent with single mechanism accounts based on convergent activation from form and meaning.

## *Information theoretic approach*

Information theory is the basis for another approach to morphological organization and processing. The essence is that information can be quantified, expressed in terms of an amount of information, that can serve as an alternative account of lexical knowledge, one that does not depend on a structure of representations arrayed by similarity in lexical space. A key measure is *entropy*: the average number of bits to communicate a message. The measure can be understood as quantifier of the uncertainty in predicting an outcome from a set of possible outcomes, which may or may not be equiprobable. In Shannon's (1948) own words, the number of discrete states of a system along with the way these states are organized determines the amount of information in that system. The cost of reducing the uncertainty is predictive of response latency in tasks such as lexical decision. Thus, in the information-theoretic approach, uncertainty and/or processing cost rather than a structure provide the organizing framework for lexical knowledge.

In the information theory framework, high information is characteristic of improbable events and low information is characteristic of probable events. Additionally, all events that may occur can be represented jointly by a probability distribution. On one extreme, as we said, this distribution can be equiprobable, representing events that have the same (equal) probability of occurrence. In that case, uncertainty is at its highest—with maximum entropy. On the other extreme, one event can have a probability of 1.0 and all others of 0.0. In that case, there is no uncertainty—entropy is zero—we know what will occur.

Stated crudely, the probability of a word appearing among the set of possible words can be described in terms of uncertainty, and uncertainty is correlated with the processing cost that is typically measured with reaction time. For example, a particular inflected form appears on an experimental trial (WORKING) and not one of its other related forms (WORKED), and this uncertainty can be quantified with information-theoretic measures, like the amount of information (c.f., Kostić, Marković, & Baucal, 2003) or the entropy (c.f., Moscoso del Prado Martín, Kostić, and Baayen, 2004), which are predictive of decision latencies (for an extensive overview see Milin, Kuperman, Kostić, & Baayen, 2009).

Processing rate per information unit increases as the amount of information becomes higher, which has been documented for words presented both in isolation (no context) and in contexts with various experimental manipulations (Kostić, 1991; 1995; Kostić, Marković, & Baucal, 2003). Similarly, as the probabilities of events get more equal, uncertainty gets higher and consequently processing time gets longer (Milin, Filipović Đurđević, Moscoso del Prado Martín, 2009; Milin *et al*., 2009; Moscoso del Prado Martín, Kostić, & Baayen, 2004).

The probability of an inflected variant of a word (like WORKING), in the information-theoretic framework is based on the frequency of that specific form when compared to the sum of the frequencies of all related forms (e.g., WORK, WORKS, WORKED, WORKING). This sum serves as a normalizing term. The probability distribution of the inflected variants of a particular word's inflected forms may differ from the probability distribution of its inflectional class in general. An analogy from English would be to compare the probability distribution of the inflected variants of a particular word (e.g., WORK: WORK; WORKS; WORKED; WORKING), with the probability distribution of all word endings—inflectional suffixes or exponents (-Ø; -S; -ED; -ING). The extent to which the general and word specific probability distributions differ is quantified by relative entropy, often called Kullback-Leibler divergence. This quantity is predictive of response latencies in lexical decision and other word recognition tasks: the greater the divergence, the longer the reaction time (Milin, Filipović Đurđević, & Moscoso del Prado Martín, 2009; Baayen, Milin, Filipović Đurđević, Hendrix, & Marelli, 2011).

In the information-theoretic framework, differences between regular and irregular verb forms get tied to properties of the words themselves, including their inflectional entropy—the frequency distribution of inflected forms including both regulars and irregulars. At the same time, it includes properties that pertain to semantics such as a word's imageability, number of senses in WordNet (Miller, 1995), contextual diversity compared to other words (Baayen & Moscoso del Prado Martín, 2005; but see also McDonald & Shillcock, 2001 and Adelman, Brown, & Quesada, 2006). In sum, within the information-theoretic framework, when regulars and irregulars incur differences in processing, the explanation is that their statistical properties differ, not that they are assigned different types of representations or different processing mechanisms from the outset. For example, the irregular past tense BUILT will be processed differently from the regular past tense HOUSED, because it differs in its uncertainty in context, that is, with respect to its frequency of occurrence and co-occurrence, distributional semantics and so on (see further discussions in Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014; Baayen, Milin, & Ramscar, 2016).

## *Naive discriminative learning*

Learning represents a counterpart to information processing, as it focuses on the *costs* of "inserting" new items into our memories. It increases the overall uncertainty, as it increases the elements in the system. It pays off, however, by increasing our capabilities to make ever-finer discriminations in our environment, which is the essence of adaptation (for general discussion related to learning and adaptation see Hinton & Nowlan, 1987). Similarly, language can be defined as a complex adaptive (sub)system (Beckner *et al*., 2009), whose principal function is to facilitate ways in which we interact with our social environment (see Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010). Thus, learning more language entails greater discrimination abilities in verbal communication.

The Naive Discriminative Learning (NDL; Baayen Milin, Đurđević, Hendrix, & Marelli, 2011; Baayen, Shaoul, Willits, & Ramscar, 2015; Milin, Ramscar, Baayen, & Feldman, 2015) framework provides an account of morphological processing that, inspired by Word and Paradigm Morphology (Matthews, 1991; Blevins, 2006), eschews the theoretical constructs of stems, morphemes, and affixes as units of form.

For language comprehension, this framework proposes a pair of two-layer networks, with connections between inputs (henceforth cues) and outputs (henceforth outcomes), which are obtained by applying the Rescorla-Wagner equations (Rescorla & Wagner, 1972) to time series of learning events—points in time at which weights between cues and outcomes are updated. The first network of the pair has letter n-grams or n-phones (typically, n is 2 or 3) as cues, and lexomes as outcomes.

The concept of the *lexome* is best explained by analogy to atoms in chemistry. Atoms have two important properties. First, from the perspective of physics, they are not indivisible, yet they suffice for understanding the chemical properties of molecules. Second, the chemical properties of molecules are specific to the molecules, and cannot be derived from the properties of the atoms. Like atoms, then, lexomes are theoretical constructs that have no meaning of their own, but instead their meaning is relational in that it emerges dynamically from the other lexomes with which they collocate. In the spirit of Landauer and Dumais (1997), semantic similarity between lexomes is approximated by the cosine of the angles between these lexomes' weight vectors. In the spirit of adaptive systems it is dynamical and evolves with learning.

In the NDL framework, the only form representations are those of letter bi/trigrams (or bi/triphones). There are no representations for stems, words, or affixes. Furthermore, no distinction is made between representations for derived, inflected or compounds words. The pivotal unit in the NDL approach is the *lexome*. Lexomes approximate experiences that are discriminated within a speech community, including not only experiences of different objects and actions, but also more "linguistic" experiences such as aspect, number, and tense. Morphological effects emerge in the first network (with letter or phone trigrams as cues and lexomes as outcomes), as a consequence of the co-occurrence statistics of these trigram or triphone cues and the lexomes (Baayen, Milin, Đurđević, Hendrix, & Marelli, 2011; Baayen, Shaoul, Willits, & Ramscar, 2015; Milin *et al.*, 2017). For example, differences in visual word recognition tasks between primes for PAST (like PASTOR and PASTA), reflect the extent to which sublexical letter trigrams are associated with the target lexome PAST. No decisions pertaining to whether trigrams such as STO, TOR, OR#, or STA and TA# (where # represents a terminal marker), are, or are not, affixes are invoked. No morphological decomposition is invoked whatsoever. Importantly, effects of form similarity of both embedded or embedding words (e.g., Bowers, Davis, & Hanley, 2005) as well as of orthographic neighbors (e.g., Baayen, Feldman, & Schreuder, 2006; Davis & Lupker, 2006; Forster & Taft, 1994; Kinoshita, Castles, & Davis, 2009) occur during learning, not from decomposition at decision time.

Learning is driven by both positive and negative evidence. As a cue occurs more often in contexts where it does not pertain to a given target lexome, the connection strength from this cue to this target lexome will be reduced. As a crucial consequence, measures based on discrimination learning go beyond frequency counts. Whereas frequency counts capture only the co-occurrence frequencies of a cue and a lexome, discrimination measures take into account how often a cue is "unfaithful' to this lexome (i.e., when it supports any other lexome; see more in Ramscar & Yarlett, 2007; Ramscar, Yarlett, Denny, & Thorpe, 2010).

The activation of a lexome is an index of how well it is discriminated from other lexomes, and typically correlates with its frequency of occurrence. A lexome's prior availability represents its degree of entrenchment in the network, and is also correlated with frequency of occurrence. But whereas the activation provides a frequency measure for bottom-up support, the prior availability is a measure of top-down expectation that is independent of the input.

Rescorla-Wagner networks specify the computational engine of the NDL approach, and are only part of a larger conceptual framework (for the specification of a discrimination-based model of auditory comprehension, see Baayen *et al.*, 2015; Hendrix, 2015, shows how NDL activation and NDL prior availability differentiate between the bottom-up and top-down processes guiding eye movements in reading compounds; Milin, Feldman, Ramscar, Hendrix, & Baayen, 2017 address "morpho-orthographic" segmentation in reading).

## Lexical access: Morphological and form effects in early visual word recognition[4]

Complex words like UNSUCCESSFULLY include several morphemic constituents, UN, SUCCESS, FUL, LY, that recur in the language in many other words. Since the seventies (Murrell & Morton, 1974; Taft & Forster, 1975) researchers have asserted that morphemic structure affects the recognition of complex words. The much-debated question—how complex words are accessed and represented in lexical memory—is closely related to the definition of a morpheme and to the possibility that morphological structure influences processing in the absence of semantics. Morphemes are units of form and semantics, as defined in traditional linguistic theory. Many interpret morphological structure devoid of semantics (comparable outcomes for semantically transparent SUCCESSFUL-SUCCESS and semantically opaque SUCCESSOR-SUCCESS) as evidence of an early prelexical (prior to access to knowledge stored in the lexicon) process. Experimental methods and measures define what is early and what is late, but whether the absence of semantically informed morphological structure necessarily implicates an earlier form stage that is independent from a later semantic process stage is more contested. We review some of the evidence for early morphological processing with and without a semantic contribution and evidence for form processing with and without a morphological contribution in the remainder of this section.

## *Semantic contributions to morphological processing: Early or late?*

Most variants of the lexicon-based account of morphological processing posit two-stages: Orthographically based but semantically blind morphological decomposition that occurs early during visual word recognition and a semantic interpretation of the decomposed constituents that occurs at a later, lexical stage: The early decomposition process is based on the form of the morpheme without regard to how that unit maps onto the "meaning" of the word in which it appears (hence "form-then-meaning"; e.g., Lavric, Rastle, & Clapp, 2011; Meunier & Longtin, 2007; Rastle, Davis, & New, 2004; Rastle & Davis, 2008). Because primes like UNCOVER and RECOVER are morphologically well structured and thus decomposable and because semantics plays no role until later in the sequence of processing stages, the rational is that they should facilitate the recognition of targets like COVER comparably regardless of whether they share both meaning and form with it, or form but little meaning.

Indeed, under masked priming conditions in many behavioral studies in English and French, different types of morphologically related primes and targets produce facilitation but have failed to show an effect of meaning (Longtin, Segui, & Hallé, 2003; Rastle *et al.*, 2000; Rastle, Davis, & New, 2004). Conversely, faster decision latencies to targets like SUCCESS after forward masked SUCCESSFUL than after SUCCESSOR provide evidence of an early role for meaning that might be characterized as semantically informed decomposition (hence "form-with-meaning"; see Feldman, Milin, Cho, Moscoso del Prado Martín, & O'Connor, 2014; Feldman, Smolka, Cho, & Milin, 2014). Similar effects arise with prefixed words like UNDRESS-DRESS and REDRESS-DRESS (Feldman, Smolka, Cho, & Milin, 2014). One difference between those studies that report an early effect of semantic transparency and those that fail to show the effect is that only the former tend to use a within target design so that the same target appears with a transparent, an opaque and an unrelated prime (for in-depth discussion about differences in the two experimental designs also consult Milin, Feldman, Ramscar, Hendrix, & Baayen, 2017).

Analogous to the behavioral measures, the above models have been tested with EEG measures using masked visual primes presented at short durations (below 50 ms SOA) and the lexical decision task (for a review of morphological EEG effects see Smolka, Gondan, & Rösler, 2015). Results with EEG provided strong N250 and/or N400 attenuations for letter sequences that are semantically related and exhaustively decomposable into morphemes (e.g., word pairs like FARMER-FARM), but results are inconsistent with respect to pairs that are semantically unrelated and differ with respect to exhaustive or partial decomposability into morphemic constituents (e.g., pairs like CORNER-CORN and BROTHEL-BROTH respectively where ER is an affix but EL is not.).

Most studies found no difference in facilitation induced by morphologically related primes and either exhaustively (e.g., + ER) or only partially (e.g., + EL) decomposable primes in either the N250 or the N400 latency range (Morris *et al.*, 2008, 2011, 2013), while one study found more priming by the former than by the latter two types (Morris *et al.*, 2007). Yet other studies revealed similar processing at an early (N250) processing stage combined with a differentiation at a later (N400) processing stage. The inconsistent patterning of words with exhaustively decomposable word primes continues to fuel the discussion about the model—"form-then-meaning" (e.g., Lavric *et al.*, 2007, 2011; Morris *et al.*, 2011), or "form-with meaning" where even early in processing true morphologically related pairs benefit from shared semantics in a way in which pairs that share only form cannot (e.g., Diependaele *et al.*, 2005; Feldman *et al.*, 2009; Morris *et al.*, 2008, 2011, 2013; Holcomb & Grainger, 2006).

Reports that semantic transparency of the prime reliably influences early morphological processing is incompatible with variants of the lexicon-based tradition where form is independently analyzed before meaning can influence recognition. The "form-with-meaning" view is based on conjoint effects of form and meaning, from the onset of the morphological processing (Baayen *et al.*, 2011; Feldman *et al.*, 2009, 2014; Feldman, Kostić, Gvozdenović, O'Connor, & Moscoso del Prado Martín, 2012). Finally, with respect to learning-based models more generally, only NDL anticipates that the effect of semantic transparency of the prime on morphological facilitation depends on the similarity of the target to the other words that constitute its form neighbors (Feldman *et al.*, 2014).

## *Morphological contributions to early orthographic processing*

Different patterns of facilitation for form similar pairs with a fully decomposable or exhaustive morphological structure like CORNER-CORN and for pairs with only a partially decomposable structure like BROTHEL-BROTH are crucial to the first stage of form-then-meaning lexical models. Neighborhood density or number of neighbors is a measure of form similarity. Analogs exist in the auditory as well as the visual domain (see Pisoni, this volume). Neighbors can be formed by letter substitution (Coltheart, Davelaar, Jonasson, & Besner, 1977), letter deletion or letter addition (for overviews, see Perea, Acha, & Fraga, 2008; Davis & Taft, 2005). A word's similarity to many other words as indexed by its orthographic neighborhood density increases single word recognition latencies when other measures are controlled (Baayen, Feldman & Schreuder, 2006; Yarkoni, Balota, & Yap, 2008). Similarly, when targets are preceded by masked primes that are neighbors, word recognition latencies get slower as target neighborhood density increases (Forster, Davis, Schoknecht, & Carter, 1987; Forster & Davis, 1991). Finally, targets from sparse orthographic neighborhoods tend to show stronger facilitation after orthographically similar primes than do targets from dense neighborhoods (Forster & Taft, 1994; Kinoshita, Castles, & Davis, 2009; Perry, Lupker & Davis, 2008).

We have recently documented that orthographic neighborhood density can systematically influence the magnitude of facilitation in morphological studies (Feldman, Smolka, Cho, & Milin, 2014; Milin *et al.*, 2017). For example, a word like

FORM with many neighbors (e.g., FORK, FOAM, FIRM, DORM) will tend to show smaller differences between masked neighbor primes (e.g., DEFORM, PERFORM) than a word with fewer neighbors like DRESS (e.g., PRESS, CRESS). Thus primes for FORM will differ less than primes for DRESS (e.g., UNDRESS, REDRESS).

These results are anticipated by NDL, using direct discriminative mappings from letter trigrams as input cues to lexome as outcomes. Contrariwise, the same results require theoretically unmotivated adjustments in lexicon-based accounts. Morphological facilitation after exhaustively decomposable forms (e.g., FARMER-FARM; CORNER-CORN; PASTOR-PAST) but not after partially decomposable forms (e.g., BROTHEL-BROTH; PASTA-PAST; LIMBO-LIMB) is central to the claim that early processing is based on semantically blind but, nonetheless morphological units, that is, not based only on form overlap (e.g., Diependaele, Sandra, & Grainger, 2005; Longtin, Segui, & Hallé, 2003; Marslen-Wilson, Bozic, & Randall, 2008; McCormick, Rastle & Davis, 2009; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Rastle, Davis, & New, 2004). However in those studies, different targets appeared with semantically unrelated exhaustively and partially decomposable primes and form similarity was only crudely matched. With more rigorous controls, that finding is not always replicable (see Milin *et al.*, 2017; also see Duñabeitia, Kinoshita, Carreiras, & Norris, 2011). The implication is that both capture form-based facilitation.

In summary, evidence is emerging that the effect of semantic transparency and of morphological structure of the prime on morphological facilitation can be linked to neighborhood properties of the target. As currently described, even lexicon-based accounts of morphological facilitation fail to anticipate the interaction of semantic transparency with target neighborhood density. This interaction not only is compatible but also is anticipated by an NDL account. However, as we explained above, NDL takes a different view on the questions of what consists a unit, how (and why) it emerges, and what its properties are. Crucially, it is not consistent with a compositional account and, thus, it avoids postulating constructs such as morphemes, stems and affixes as its units (following "Word and Paradigm Morphology": Matthews, 1991; Blevins, 2006). Accordingly, it also does not hypothesize "meaning" residing in those units. Instead meaning arises from the collocation dynamics between lexomes.

## *Explaining cross-language differences in morphological processing*

Effects of orthographic and morphological similarity that vary with visual neighborhood density of the target may help to explain discrepant priming outcomes across languages. For example, studies in Hebrew and Arabic replicated robust morphological priming by masked semantically transparent and opaque primes across long and short prime durations in the lexical decision task (Boudelaa & Marslen-Wilson, 2005; Deutsch, Frost, & Forster, 1998; Frost *et al.*, 1997). In addition, both form and morphological facilitation are easier to document in English (viz., for targets from sparse neighborhoods) than in Hebrew (Frost *et al.*, 2005). In fact the absence of form priming in Hebrew is sometimes interpreted as evidence for language-specific processing (Frost, Deutsch, & Forster, 2000). However, Hebrew roots are only three or four consonants in length and most vowels are not written in text for adult readers. The implication is that when neighborhood density is defined orthographically, Hebrew words will tend to be shorter than words in English or French and, generally, across languages including Hebrew, shorter words tend to have more neighbors.

## *Lexical representation: Morphological effects in late visual word recognition*

In contrast to masked priming, unmasked or overt priming taps into a later stage of lexical processing and representation. Under auditory priming or visual priming at long SOAs, in languages like English and French, stems were primed by suffixed derivations if they were semantically transparent, as in SUCCESSFUL-SUCCESS. By contrast, facilitation was absent for semantically opaque derivations like SUCCESSOR-SUCCESS (for cross-modal priming see Longtin, Segui, & Hallé, 2003; Marslen-Wilson, Tyler, Waksler, & Older, 1994; for visual priming at long SOAs, see Feldman & Soltano, 1999; Feldman, Soltano, Pastizzo, & Francis, 2004; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Feldman & Larabee, 2001; Meunier & Segui, 2002). Lexicon-based models posit different processing mechanisms for semantically transparent and opaque words at the lexical level. Semantic information can be integrated (in the two-stage model, e.g., Lavric *et al.*, 2011), shared representations can operate at the morpho-semantic level (in the dual-route model, e.g., Morris *et al.*, 2013), or form-to-meaning mappings can be realized (as in a connectionist network (e.g., Plaut & Gonnerman, 2000). EEG findings show a similar pattern. Attenuations to the N400 were largest when they were induced by true morphologically related words like FARMER-FARM, slightly reduced for pairs like CORNER-CORN and smallest for form pairs like BROTHEL-BROTH (Lavric *et al.*, 2011). Results are traditionally interpreted as consistent with the model of visual word recognition with an orthographically based morphological decomposition followed by a later semantically informed stage.

Similarly, prefixed words in English and Serbian facilitated lexical latencies, when they were semantically transparent as in DISOBEY-OBEY, or PRIVOLE-VOLIM. Characteristic of this task is the absence of facilitation when the derivations were semantically opaque as in RESTRAIN-STRAIN or ZAVOLE-VOLIM (Marslen-Wilson *et al.*, 1994; Feldman, Barac-Cikoja, & Kostic, 2002; Feldman & Larabee, 2001). Similarly with sentence primes, semantically transparent or ambiguous Dutch prefixed verbs (that possess a transparent and an opaque reading) showed facilitation, whereas truly opaque prefixed verbs did not (Zwitserlood, Bolwiender, & Drews, 2005).

The above findings are traditionally explained by assuming that the lexical representations of complex words that underlie performance when primes are overt depend on semantic compositionality: Only semantically transparent words can benefit recognition of their stem. By contrast, semantically opaque words like

SUCCESSOR are not represented as derived. Rather, they are represented as unanalyzed and independent lexical items that do not share a stem with words like SUCCESS or SUCCESSFUL (Marslen-Wilson *et al.*, 1994). In lexicon-based models, morphological decomposition as revealed by overt priming is constrained by semantic knowledge that influences the interrelation among lexical entries (Diependaele, Sandra, & Grainger, 2005; Marslen-Wilson, Bozic, & Randall, 2008; Meunier & Longtin, 2007; Rastle *et al.*, 2000; 2004; Taft & Kougious, 2004; Taft & Nguyen-Hoan, 2010).

In contrast, connectionist learning-based models posit a continuity between early and late tasks and emphasize graded effects of form and meaning, in behavioral (Gonnerman, Seidenberg, & Andersen, 2007) and in EEG (Kielar & Joanisse, 2011) studies. For example, stronger N400 priming effects for semantically transparent word pairs like GOVERNMENT-GOVERN than for less transparent word pairs like DRESSER-DRESS along with no facilitation for semantically opaque pairs like APARTMENT-APART or CORNER-CORN capture graded semantic similarity when form similarity is held constant.

## *The origin of cross-language differences in late morphological processing*

In German and in Hebrew effects of semantic transparency on morphological facilitation have not been detected regardless of processing time for the prime. Studies on prefixed verbs in German have found equivalent morphological facilitation after semantically opaque (VERSTEHEN-STEHEN, "understand-stand") and transparent (AUFSTEHEN-STEHEN, "stand up-stand") verbs with both auditory and visual presentations, even at long (300 ms and 1000 ms) SOAs (Smolka, Komlósi, & Rösler, 2009; Smolka, Preller, & Eulitz, 2014). Together with the absence of facilitation for form controls like VERKLEIDEN-LEIDEN ("disguise-suffer") these morphological effects cannot be attributed to form similarity. Similar to the behavioral findings, N400 attenuations are equivalent for semantically transparent (ANKOMMEN-KOMMEN, "arrive-come") and opaque (UMKOMMEN-KOMMEN, "perish-come") prefixed verbs in German (Smolka, Gondan, & Rösler, 2015). Moreover, these morphological effects were stronger than either pure semantic or pure form effects.

Approximating the findings in German, findings in Hebrew and Arabic also showed morphological effects that failed to vary with semantic transparency. A long term priming study in Hebrew (Bentin & Feldman, 1990) demonstrated that complex words in Hebrew are represented in the lexicon in terms of their root (e.g., GDL). The recognition of a target like miGDaL ("tower") was primed when it was preceded by morphologically related words (same GDL root) both when they were semantically related like GaDoL ("big") or semantically unrelated like GiDuL ("tumor"). Further studies in Hebrew and Arabic replicated robust and equivalent facilitation after both semantically transparent and opaque derivations (Boudelaa & Marslen-Wilson, 2004a, 2004b, 2005; Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000). In a traditional framework, the findings in

German, Hebrew and Arabic suggest that lexical representations capture morphological, specifically root structure regardless of meaning compositionality. To elaborate, whereas the lexical representation of a complex verb like UNDERSTAND refers to its base STAND (Smolka *et al.*, 2009; Smolka, Preller, & Eulitz, 2014; Smolka, Gondan, & Rösler, 2015), that of a Hebrew or Arabic word like GiDuL is with reference to its root GDL (Frost *et al.*, 2000; Boudelaa & Marslen-Wilson, 2005). Strong morphological effects without an effect of semantic transparency appear difficult to reconcile with connectionist or other accounts that depend on the general convergence of codes. Some offer these findings in support of rules and representations that *a priori* differ across languages. Underrepresented are attempts to track the role of general statistical properties that characterize words and their similarity to other words and how these tendencies might differ across languages.

In summary, different portraits of morphological facilitation across languages—restricted by semantic transparency in English, French, Dutch, and Serbian but independent of semantic compositionality in German, Hebrew, and Arabic—highlight the importance of cross-linguistic research. Hebrew and Arabic are morphologically rich languages (e.g., Ravid, 2012) and German represents the morphologically richest language within the Indo-Germanic language family, given that it has retained morphological markers to indicate grammatical functions (De Vogelaer, 2007; Roelcke, 1997). Therefore, one possibility is that the morphological richness of a language influences the representation of its morphological structure. However, within a connectionist framework, a greater effect of semantic transparency in morphologically impoverished than in a morphologically rich languages has been linked to the strength of form-meaning regularities (consult Plaut & Gonnerman, 2000; Raveh & Rueckl, 2000; Rueckl *et al.*, 1997; also, see Mirković, Seidenberg, & Joanisse, 2011 for connectionist model of Serbian noun paradigms). Hence, in a system with abundant form-meaning regularities, these patterns may guide visual word recognition. In a network simulation of a morphologically impoverished environment (a language like English), in which mappings between orthographic surface forms and their meanings are mostly idiosyncratic, morphological regularities played a minor role (Plaut & Gonnerman, 2000). In contrast, in the simulation of a morphologically rich environment (a language like Hebrew) with its dense overlapping mappings between many orthographic forms and meanings, form based morphological regularities dominated processing and simulated morphological priming effects were independent of semantic relatedness. Of course, even in morphologically rich environments, semantically transparent word forms should yield some advantage over semantically opaque ones, but this has not been demonstrated for prefixed words in German.

Within a language, evidence is accruing that morphological effects such as semantic transparency are likely to be more important for some types of words than for others and this may help us understand where and when differences emerge across languages. Semantic transparency can be operationalized in terms from distributional semantics such as the cosine similarity between a stem and a derived-form vector (Marelli & Baroni, 2015). We recently demonstrated in English that the influence of an effect of semantic similarity between morphologically-related prime and target is weaker when primes have many semantically similar words, with a dense semantic neighborhood, than when they have fewer semantically similar words (Feldman, Marelli, Amenta, & Milin, 2015). Whether differences across languages can be linked reliably to general properties like neighborhood density or mappings between form and semantics (Marelli, Amenta, & Crepaldi, 2015; Amenta, Marelli, & Sulpizio, 2016) or whether language-specific properties must be invoked to account for some of the variation in the patterns of morphological facilitation awaits further research.

## Summary: Lexical Access and Morphological Processing

Lexicon-based and learning-based models ask different questions about and offer different solutions to challenges in morphological processing. From a general combinatorial position, when words are composed of multiple morphemes, they range in semantic compositionality from semantically transparent, as in SUCCESS-FUL to "opaque," as in SUCCESS-OR and the meaning of semantically opaque words cannot be derived from linguistic rules that combine their meaningful components. Rules that operate on symbols, with storage of more opaque full word forms as a backup, is one popular way to characterize morphological knowledge. Insofar as rules are language specific, some describe differences between languages in terms of differences between rules.

Learning-based approaches look for universal, albeit more complex patterns that can vary with statistical properties of words. Here, we have used tools from distributional semantics and the lexical decision task in English to demonstrate that by differentiating among words that are semantically similar to many versus few other words we can better predict effects of semantic transparency in lexical processing. A better grasp of the variability among words within a language with respect to properties like orthographic and semantic neighborhood size, and how that variation differs across languages, may prove crucial to determine whether or not rules are necessary to capture knowledge about the patterns that form complex words, and the extent to which the underlying process is specific to a language.

Specifically within the NDL framework, meaning loses its encapsulation and becomes relational and contextualized—dependent on the communicative intentions between interlocutors. Form units are also reduced to naive (i.e., theoretically empty) sublexical n-grams of letters or phones, avoiding the burden of representing traditional morphemes and stems and affixes. Within this framework, in particular, there is no simple and direct mapping between form units and meaning units, but rather two interdependent networks, where morphological effects emerge from the network with letter or phone bi/trigrams as cues and lexomes as outcomes, and where word semantics emerge from a second network that takes neighboring lexomes as cues and target lexomes as outcomes, and the "meaning" of a lexome is best approximated by its relational behavior (vector of outgoing weights) with other lexomes.

## Author Note

This research was supported in part by the Alexander von Humboldt Foundation to Petar Milin through an Alexander von Humboldt research chair awarded to R. H. Baayen, by the National Institute of Child Health and Human Development HD 01994 to Haskins Laboratories, and by the VolkswagenStiftung, Grant FP 561/11 to Eva Smolka.

## NOTES

1 Traditionally, a morpheme is the minimal unit of meaning and morphology refers to the study of word structure for those units. As will become evident (see section 3.4), "meaning" is a term that is compatible with traditional lexical accounts but less so with current learning accounts.
2 This is, of course, conditioned on yet another implicit assumption that morphemes and words are independent "meaning carriers."
3 The term "mechanism" is used differently in this chapter than in the work of the proponents of dual mechanism accounts noted here. That framework posits two distinct brain mechanisms, rather than two different routes (e.g., parsing and retrieval) within a single mechanism.
4 In the discussion below, we revert to the traditional term "meaning" and attempt to remain agnostic with respect to how it should be represented.

## REFERENCES

Adelman, J. S., Brown, G. D., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science*, 17(9), 814–823. doi: 10.1111/j.1467-9280.2006.01787.x

Alegre, M. & Gordon, P. (1999). Rule-Based Versus Associative Processes in Derivational Morphology. *Brain and Language.* 68(1), 347–354. doi: 10.1006/brln.1999.2066

Amenta, S., Marelli, M., & Sulpizio, S. (2016). From sound to meaning: Phonology-to-Semantics mapping in visual word recognition. *Psychonomic Bulletin & Review*, 1–7.

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 55(2), 290–313. doi: 10.1016/j.jml.2006.03.008

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, 118(3), 438–481. doi: 10.1037/a0023851

Baayen, R. H., Milin, P., & Ramscar, M. (2016). Frequency in lexical processing. *Aphasiology*, 30(11), 1174–1220. doi: org/10.1080/02687038.2016.1147767

Baayen, R. H., & Moscoso del Prado Martín, F. (2005). Semantic density and past-tense formation in three Germanic languages. *Language*, 81(3), 666–698. doi: 10.1353/lan.2005.0112

Baayen, R. H., Shaoul, C., Willits, J. and Ramscar, M. (2015). Comprehension without segmentation: A proof of concept with naive discrimination learning. *Language, Cognition and Neuroscience*, 31(1), 106–128.doi: 10.1080/23273798.2015.1065336

Baayen R. H., Wurm, L. H., & Aycock, J. (2007). Lexical dynamics for low-frequency complex words: A regression study across task sand modalities. *The Mental Lexicon*, 2(3), 419–463. doi: 10.1075/ml.2.3.06baa

Basnight-Brown, D., Chen, H., Shu, H., Kostić, A. & Feldman, L. B. (2007). Monolingual and Bilingual Recognition of Regular and Irregular English Verbs: Does Sensitivity to Word Form Vary with Language Experience? *Journal of Memory and Language*, 57(1), 65–80. doi: 10.1016/j.jml.2007.03.001

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language learning*, 59(1), 1–26. doi: 10.1111/j.1467-9922.2009.00533.x

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *The Quarterly Journal of Experimental Psychology*, 42(4), 693–711. doi: 10.1080/14640749008401245

Bird, H., Lambon Ralph, M.A., Seidenberg, M.S., McClelland, J. L., & Patterson, K. (2003). Deficits in phonology and past-tense morphology: What's the connection?. *Journal of Memory and Language*, 48(3), 502–526. doi: 10.1016/s0749-596x(02)00538-7

Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, 42(3), 531–573. doi: 10.1017/S0022226706004191

Bodner, G.E., & Masson, M. E. J. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review*, 10(3), 645–652. doi: 10.3758/BF03196527

Boudelaa, S., & Marslen-Wilson, W. (2004a). Abstract morphemes and lexical representation: The CV-Skeleton in Arabic. *Cognition*, 92(3), 271–303. doi: 10.1016/j.cognition.2003.08.003

Boudelaa, S., & Marslen-Wilson, W. (2004b). Allomorphic variation in Arabic: Implications for lexical processing and representation. *Brain and Language*, 90(1), 106–116. doi: 10.1016/S0093-934X(03)00424-3

Boudelaa, S. & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, 20(1-2), 207–260. doi: 10.1080/01690960444000106

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Automatic semantic activation of embedded words: Is there a "hat" in "that"?. *Journal of Memory and Language*, 52(1), 131–143. doi: 10.1016/j.jml.2004.09.003

Budd, M.-J., Paulmann, S., Barry, C. & Clahsen, H. (2013). Brain potentials during language production in children and adults: An ERP study of the English past tense. *Brain and Language*, 127(3), 345–355. doi: 10.1016/j.bandl.2012.12.010

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language Production*, pp. 257–294. London: Academic Press.

Bybee, J. (1985). *Morphology*: A Study of the Relation between Meaning and Form (Vol. 9).John Benjamins Publishing. doi: 10.1075/tsl.9

Bybee, J. (1995). Diachronic and typological properties of morphology and their implications for representation. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (109–129). Hillsdale, NJ. Erlbaum.

Bybee, J. and McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, 22(2-4), 381–410. doi: 10.1515/tlir.2005.22.2-4.381

Clahsen, H., Prüfert, P., Eisenbeiss, S. & J. Cholin (2002). Strong stems in the German mental lexicon: Evidence from child language acquisition and adult processing. In: I. Kaufmann & B. Stiebels (Eds.), *More than Words: a Festschrift for Dieter Wunderlich*. Akademie-Verlag: Berlin, pp. 91–112.

Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention & Performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum

Davis, C. J., & Lupker, S. J. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 668. doi: 10.1037/0096-1523.32.3.668

Davis, C., & Taft, M. (2005). More Words in the Neighborhood: Interference in Lexical Decision Due to Deletion Neighbors. *Psychonomic Bulletin & Review*, 12(5), 904–910. doi: 10.3758/BF03196784

De Jong, N., Schreuder, R., and Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, 15(4-5), 329–365. doi: 10.1080/01690960050119625

Deutsch, A., Frost, R., & Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(5), 1238–1255. doi: 10.1037/0278-7393.24.5.1238

De Vogelaer, G. (2007). Extending Hawkins' comparative typology: Case, word order, and verb agreement in the Germanic languages. *Nordlyd*, 34(1), 167–182. doi: 10.7557/12.93

Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes*, 20(1-2), 75–114. doi: 10.1080/01690960444000197

Duñabeitia, J. A., Kinoshita, S., Carreiras, M., & Norris, D. (2011). Is morpho-orthographic decomposition purely orthographic? Evidence from masked priming in the same-different task. *Language and Cognitive Processes*, 26(4–6), 509–529.

Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied linguistics*, 27(1), 1–24.

Feldman, L. B. (2000). Are morphological effects distinguishable from the effects of shared meaning and shared form? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(6), 1431–1444. doi: 10.1037/0278-7393.26.6.1431

Feldman, L. B., Barac-Cikoja, D., & Kostić, A. (2002). Semantic aspects of morphological processing: Transparency effects in Serbian. *Memory and Cognition*, 30(4), 629–636. doi: 10.3758/BF03194964

Feldman L. B., & Basnight-Brown D. M. (2008). List context fosters semantic processing: parallels between semantic and morphological facilitation when primes are forward masked. *Journal of Experimental Psychology: Learning, Memory and Cognition*. May; 34(3), 680–7. doi: 10.1037/0278-7393.34.3.680.

Feldman, L. B., & Larabee, J. (2001). Morphological facilitation following prefixed but not suffixed primes: Lexical architecture or modality-specific processes?. *Journal of Experimental Psychology: Human Perception and Performance*, 27(3), 680–691. doi: 10.1037//0096-1523.27.3.680

Feldman, L. B., Marelli, M., Amenta, S. and Milin, P. (2015). Effects of target density on morphological processing at long and short SOAs. First Quantitative Morphology Meeting. Belgrade, Serbia. 11-12 July 2015.

Feldman, L. B., Kostić, A. Gvozdenović, V., O'Connor, P. A. and Moscoso del Prado

Martín, F. (2012). Semantic similarity influences early morphological priming in Serbian: A challenge to form-then-meaning accounts of word recognition. *Psychonomic Bulletin and Review*, 19(4), 668–676. doi: 10.3758/s13423-012-0250-x.

Feldman, L. B., Milin, P., Cho, K., Moscoso del Prado Martín, F. & O'Connor, P. A. (2014). Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience*, 9. doi: 10.3389/fnhum.2015.00111

Feldman, L. B. & Moscoso del Prado Martín, F. (2012). Does a focus on universals represent a new trend in word recognition?. *Behavioral and Brain Science*, 35(5). doi: 10.1017/S0140525X12000295.

Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, 16(4), 684–691. doi: 10.3758/PBR.16.4.684

Feldman, L. B., Smolka, E., Cho, K., & Milin, P. (2014). Effects of semantic transparency on morphological processing: Prefixed English Verbs. Paper presented at the Ninth International Meeting of the Mental Lexicon. Niagara on the Lake, ON. Oct. 2014.

Feldman, L. B., & Soltano, E. G. (1999). Morphological priming: The role of prime duration, semantic transparency, and affix position. *Brain and Language*, 68(1-2), 33–39. doi: 10.1006/brln.1999.2077

Feldman, L. B., Soltano, E. G., Pastizzo, M. J., & Francis, S. E. (2004). What do graded effects of semantic transparency reveal about morphological processing?. *Brain and Language*, 90(1), 17–30. Brln.1999.1077 doi: 10.1016/S0093-934X(03)00416-4

Forster, K. I., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language*, 30(1), 1–25. doi: 10.1016/0749-596X(91)90008-8

Forster, K. I., Davis, C., Schoknecht, C., and Carter, R. (1987). Masked priming with graphemically related forms: repetition or partial activation?. *Quarterly Journal of Experimental Psychology*, 39(2), 211–251. doi: 10.1080/14640748708401785

Forster, K. I., and Taft, M. (1994). Bodies, antibodies, and neighborhood density effects in masked form-priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 844–863. doi: 10.1037/0278-7393.20.4.844

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, 355, 263–279. doi: 10.1017/S0140525X11001841

Frost, R., Deutsch, A. & Forster, K. I. (2000). Decomposing Morphologically Complex Words in a Nonlinear Morphology. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 26(3), 751–765. doi: 10.1037//0278-7393.26.3.751

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. (2000). Morphological priming: Dissociation of phonological, semantic and morphological factors. *Memory and Cognition*, 28(8), 1277–1288. doi: 10.3758/BF03211828

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(4), 829–856. doi: 10.1037//0278-7393.23.4.829

Frost, R., Kugler, T., Deutsch, A., & Forster, K. I. (2005). Orthographic structure versus morphological structure: Principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6), 1293–1326. doi: 10.1037/0278-7393.31.6.1293

Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, 136(2), 323–345. doi: 10.1037/0096-3445.136.2.323

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading: Cooperative Division of Labor Between Visual and Phonological Processes. *Psychological Review*, 111(3), 662–720. doi: 10.1037/0033-295X.111.3.662

Hendrix, P. (2015). *Experimental explorations of a discrimination learning approach to language processing*. Doctoral dissertation. Eberhard Karls University Tübingen. doi: 10.15496/publikation-9333

Hinton, G. E., & Nowlan, S. J. (1987). How learning can guide evolution. *Complexsystems*, 1(3), 495–502.

Holcomb, P., and Grainger, J. (2006). On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, 18(10), 1631–1643. doi: 10.1162/jocn.2006.18.10.1631

Kielar, A., Joanisse, M. F., and Hare, M. L. (2008). Priming English past tense verbs: rules or statistics? *Journal of Memory and Language*, 58(2), 327–346. doi: 10.1016/j.jml.2007.10.002

Kielar, A., & Joanisse, M. F. (2011). The role of semantic and phonological factors in word recognition: An ERP cross-modal priming study of derivational morphology. *Neuropsychologia*, 49(2), 161–177. doi: 10.1016/j.neuropsychologia.2010.11.027

Kinoshita, S., Castles, A., & Davis, C. (2009). The role of neighbourhood density in transposed-letter priming. *Language and Cognitive Processes*, 24(4), 506–526. doi: 10.1080/01690960802289213

Kostić, A. (1991). Informational approach to processing of inflected morphology: standard data reconsidered. *Psychological Research*, 53(1), 62–71. doi: 10.1007/BF00867333

Kostić, A. (1995). Informational load constraints on processing inflected morphology. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing*, Hillsdale,NJ: Erlbaum. doi: 10.1007/BF00867333

Kostić, A., Marković, T., & Baucal, A. (2003). Inflectional morphology and word meaning: orthogonal or co-implicative cognitive domains? R. H. Baayen & R. Schreuder (Eds.): *Morphological Structure in Language Processing*. Mouton de Gruyter.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211. doi: 10.1037/0033-295X.104.2.211

Lavric, A., Clapp, A., and Rastle, K. (2007). ERP evidence of morphological analysis from orthography: a masked priming study. *Journal of Cognitive Neuroscience*, 19(5), 866–877. doi: 10.1162/jocn.2007.19.5.866

Lavric, A., Rastle, K., & Clapp, A. (2011). What do fully visible primes and brain potentials reveal about morphological decomposition?. *Psychophysiology*, 48(5), 676–686. doi: 10.1111/j.1469-8986.2010.01125.x

Longtin, C., Segui, J., & Hallé, P. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, 18(3), 313–334. doi: 10.1080/01690960244000036

McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech*, 44(3), 295–322. doi: 10.1177/00238309010440030101

Marelli, M., Amenta, S., & Crepaldi, D. (2015). Semantic transparency in free stems: The effect of Orthography-Semantics Consistency on word recognition. *The Quarterly Journal of Experimental Psychology*, 68(8), 1571–1583.

Marelli, M. & Baroni, M. (2015). Affixation in Semantic Space: Modeling Morpheme Meanings With Compositional Distributional Semantics. *Psychological Review*, 122(3), 485–515. doi: 10.1037/a0039267

Marslen-Wilson, W., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes*, 23(3), 394–421. doi: 10.1080/01690960701588004

Marslen-Wilson,W., & Tyler, L.K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Sciences*, 2(11), 428–435. doi: 10.1016/S1364-6613(98)01239-X

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 10(1), 3–33. doi: 10.1037/0033-295X.101.1.3

Matthews, P. H. (1991). Morphology. Cambridge: Cambridge University Press.

McCormick, S. F., Rastle, K., & Davis, M., H. (2009). Adore-able not adorable? Orthographic underspecification studied with masked repetition priming. *European Journal of Cognitive Psychology*, 21(6), 813–836. doi: 10.1080/09541440802366919

McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech*, 44(3), 295–322.

Meunier, F., & Longtin, C.-M. (2007). Morphological decomposition and semantic integration in word processing. *Journal of Memory and Language*, 56(4), 457–471. doi: 10.1016/j.jml.2006.11.005

Meunier, F., & Segui, J. (2002). Cross-modal priming in French. *Brain and Language*, 81(1), 89–102. doi: 10.1006/brln.2001.2509

Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R. H. (2017). Discrimination in lexical decision. *PLoS One*, 12(2), e0171935.

Milin, P., Filipović Đurđević, D. & Moscoso del Prado Martín, F. (2009). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language*, 60(1), 50–64. doi: 10.1016/j.jml.2008.08.007

Milin, P., Kuperman, V., Kostić, A., & Baayen, H. R. (2009). Words and paradigms bit by bit: An information-theoretic approach to the processing of inflection and derivation.In: J. P. Blevins and J. Blevins (Eds.), *Analogy in Grammar: Form and Acquisition* (pp. 214–252). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199547548.003.0010

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, 38(11), 39–41. doi: 10.1145/219717.219748

Mirković, J., Seidenberg, M. S., & Joanisse, M. F. (2011). Rules versus statistics: Insights from a highly inflected language. *Cognitive Science*, 35(4), 638–681. doi: 10.1111/j.1551-6709.2011.01174.x

Morris, J., Franck, T., Grainger, J., P. J. Holcomb, P. J. (2007.) Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology*, 44(4), 506–521. doi: 10.1111/j.1469-8986.2007.00538.x

Morris, J., Grainger, J., & Holcomb, P., J. (2008). An electrophysiological investigation of early effects of masked morphological priming. *Language and Cognitive Processes*, 23(7-8), 1021–1056. doi: 10.1080/01690960802299386

Morris, J., Grainger, J., & Holcomb, P., J. (2013). Tracking the consequences of morpho-orthographic decomposition using ERPs. *Brain Research*, 1529, 92–104. doi: 10.1016/j.brainres.2013.07.016

Morris, J., Porter, J., H., Grainger, J., & Holcomb, P., J. (2011). Effects of lexical status and morphological complexity in masked priming: an ERP study. *Language and Cognitive Processes*, 26(4-6), 558–599. doi: 10.1080/01690965.2010.495482

Moscoso del Prado Martín, F., Kostic, A., & Baayen, R. H. (2004). Putting the bits together: an information theoretical perspective on morphological processing. *Cognition*, 94(1), 1–18. doi: 10.1016/j.cognition.2003.10.015

Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology*, 102(6), 963–968. doi: 10.1037/h0036551

Napps, S. E., & Fowler, C. A. (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research*, 16(3), 257–272. doi: 10.1007/BF01067546

Patterson K., Lambon Ralph M.A., Hodges, J.R., and McClelland. J.L. (2001). Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia*, 39(7), 709–24. doi: 10.1016/S0028-3932(01)00008-2

Perea, M., Acha, J., & Fraga, I. (2008). Lexical competition is enhanced in the left hemisphere: Evidence from different types of orthographic neighbors. *Brain and Language*, 105(3), 199–210. doi: 10.1016/j.bandl.2007.08.005

Perry, J. R., Lupker, S. J. & Davis, C. J. (2008). An evaluation of the interactive-activation model using masked partial-word priming. *Language and Cognitive Processes*, 23(1), 36–68. doi: 10.1080/01690960701578112

Pinker, S. (1999). *Words and Rules*.New York, NY: Harper Perennial.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6, 456–463.

Pinker, S., and Ullman, M. (2003). Beyond one model per phenomenon. *Trends in Cognitive Sciences*, 7(3), 108–109. doi: 10.1016/S1364-6613(03)00021-4

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing?. *Language and Cognitive Processes*, 15(4-5), 445–485. doi: 10.1080/01690960050119661

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103(1), 56–115. doi: 10.1037//0033-295X.103.1.56

Ramscar, M., & Baayen, H. (2013). Production, comprehension, and synthesis: a communicative perspective on language. *Frontiers in Psychology*, 4, 233. doi: 10.3389/fpsyg.2013.00233

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, 6(1), 5–42. doi: 10.1111/tops.12078

Ramscar, M. & Port, R. (2015). Categorization (without categories). In E. Dawbroska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics*. Berlin, DE: De Gruyter Mouton.

Ramscar, M., & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science*, 31(6), 927–960. doi: 10.1080/03640210701703576

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*, 34(6), 909–95. doi: 10.1111/j.1551-6709.2009.01092.x

Rastle, K., Davis, M., H., Marslen-Wilson, W., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, 15(4-5), 507–537. doi: 10.1080/01690960050119689

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, 23(7-8), 942–971. doi: 10.1080/01690960802069730.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review*, 11(6), 1090–1098. doi: 10.3758/BF03196742

Raveh, M., & Rueckl, J. (2000). Equivalent effects of inflected and derived primes: Long-term morphological priming in fragment completion and lexical decision. *Journal of Memory and Language*, 42(3), 103–119. doi: 10.1006/jmla.1999.2673

Ravid, D. (2012). *Spelling morphology: The psycholinguistics of Hebrew spelling*. New York: Springer. doi: 10.1007/978-1-4419-0588-8_13

Rescorla, R. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43(3), 151–160. doi: 10.1037//0003-066X.43.3.151

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). Appleton-Century-Crofts.

Roelcke, T. (1997). Sprachtypologie des Deutschen: historische, regionale und funktionale Variation [Language typology of German: historical, regional, and functional variation] (Vol. 2812) Berlin: Walter de Gruyter.

Rueckl, J. G. (2002). The dynamics of visual word recognition. *Ecological Psychology*, 14(1-2), 5–19. doi: 10.1080/10407413.2003.9652750

Rueckl, J. G., & Galantucci, B. (2005). The locus and time course of long-term morphological priming. *Language and Cognitive Processes*, 20(1-2), 115–138. doi: 10.1080/01690960444000188

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C. S., & Mars, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language*, 36(3), 382–405. doi: 10.1006/jmla.1996.2489

Rueckl, J., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language*, 68(1-2), 110–117. doi: 10.1006/brln.1999.2106

Seidenberg, M., & Gonnerman, L. M. (2000). Explaining derivational morphology asthe convergence of codes. *Trends in Cognitive Sciences*, 4(9), 353–361.

Shannon's (1948). doi: 10.1016/S1364-6613(00)01515-1

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11(2), 245–260. doi: 10.1017/S1366728908003404

Smolka, E., Gondan, M., & Rösler, F. (2015). Take a stand on understanding: Electrophysiological evidence for stem access in German complex verbs. *Frontiers in Human Neuroscience*, 9, 62. doi: 10.3389/fnhum.2015.00062.

Smolka, E., Khader, P., Wiese, R., Zwitserlood, P., & Rösler, F. (2013). Neurobiological evidence for the continuous processing of linguistic categories of regular and irregular verb inflection in German. *Journal of Cognitive Neuroscience*, 25(8), 1284–1304. doi: 10.1162/jocn_a_00384

Smolka, E., Komlósi, S., & Rösler, F. (2009). When semantics means less than morphology: Processing German prefixed verbs. *Language and Cognitive Processes*, 24(3), 337–375. doi: 10.1080/01690960802075497

Smolka, E., & Libben, G. (2017). "Can you wash off the hogwash?" – semantic transparency of first and second constituents in the processing of German compounds. *Language, Cognition and Neuroscience*, 32(4), 514–531. doi: 10.1080/23273798.2016.1256492

Smolka, E., Preller, K., & Eulitz, C. (2014). 'Verstehen' 'understand' primes 'stehen' 'stand': Morphological structure overrides semantic compositionality in the lexical representation of German complex verbs. *Journal of Memory and Language*, 72, 16–36. doi: 10.1016/j.jml.2013.12.002

Smolka, E., Zwitserlood, P., & Rösler, F. (2007). Stem access in regular and irregular inflection: Evidence from German participles. *Journal of Memory and Language*, 57(3), 325–347. doi: 10.1016/j.jml.2007.04.005

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 638–647. doi: 10.1016/S0022-5371(75)80051-X

Taft, M., & Kougious, P. (2004). The processing of morpheme-like units in monomorphemic words. *Brain and Language*, 90(1), 9–16. doi: 10.1016/S0093-934X(03)00415-2

Taft, M., & Nguyen-Hoan, M. (2010). A sticky stick? The locus of morphological representation in the lexicon. *Language and Cognitive Processes*, 25(2), 277–296. doi: 10.1080/01690960903043261

Ullman, M. T. (2004). Contributions of neural memory circuits to language: The declarative/procedural model. *Cognition*, 92*(*1-2*)*. 231–270. doi: 10.1016/j.cognition.2003.10.008

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30(1). 37–69. doi: 10.1023/A:1005204207369

Yarkoni, T., Balota, D. A., & Yap, M. J. (2008). Beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15, 971–979. doi: 10.3758/PBR.15.5.971

Zwitserlood, P., Bolwiender, A., & Drews, E. (2005). Priming morphologically complex verbs by sentence contexts: Effects of semantic transparency and ambiguity. *Language and Cognitive Processes*, 20(1-2), 395–415. doi: 10.1080/01690960444000160

# 12 Orthography, Word Recognition, and Reading

DAVID BRAZE AND TAO GONG

Haskins Laboratories

## Introduction

The nature of connection between oral and written language has been worried over by researchers for decades, perhaps as much as a century (Huey, 1908). Huey credits V. Egger with the notion that "to read is, in effect, to translate writing into speech" (Huey, 1908, p. 123). Some 60 years later, this basic idea was well on its way to becoming a lynch pin of our modern understanding of the relationship between speech perception, printed word recognition and language comprehension. Alvin Liberman, in Kavanagh (1968), framed the idea somewhat differently when he wrote that "reading is in some sense parasitic on speech" (p. 123). This is to say that speech perception and reading are not wholly independent systems, but that they stand in different relations to Language. The medium of speech is privileged over that of writing in ways that are now widely recognized. In the history of our species speech precedes print; in the development of the individual, the ability to speak comes before facility with the written word, if the latter comes at all. In fact, speech and language development would seem to be biological imperatives whereas achieving skill with the written word is most definitely not.

The "parasitic" nature of the relation between speech and print is such that the connection between writing and language is mediated by speech. For those writing systems that have served as objects of most literacy-oriented research to date, the Alphabetic Principle underlies that connection. This mapping principle asserts that there is an explicit correspondence between orthographic characters and speech, specifically at the phonemic level of representation (Bloomfield, 1942). Mattingly (1972; also see Liberman, 1988; Liberman *et al.*, 1989) posited that the acquisition of reading ability in an alphabetic writing system relies on explicit knowledge of the phonemic structure of a language (phoneme awareness), and that that knowledge does not follow automatically from being a competent speaker

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

of the language. Subsequent experimental work has largely confirmed Mattingly's surmise about the connection between meta-linguistic awareness of phonemic structure and the development of alphabetic literacy (e.g., Byrne *et al.*, 2000; Bowey, 1995; Oakhill & Cain, 2011). We will see that there is, moreover, reason to believe that language/orthography appropriate variations of the Mattingly's alphabetic mapping principle and metalinguistic awareness may be important to the mastery of any type of writing system.

Implemented cognitive models of word recognition that are most relevant to discussions of reading comprehension integrate perception of external language both by ear and by eye. So, it is critically important to have some common terminology for talking about relationships of writing to speech and language. Perfetti *et al.* (2007; also Coulmas, 1989) distinguish among *scripts*, *writing systems* and *orthographies*. In their terms, a *script* is simply a set of symbols, the visual aspect of writing, independent of any association with linguistic elements. *Writing system* refers to the nature of the basic mapping from symbols to linguistic units; an *orthography* is a specific language-to-script mapping. So, alphabetic writing is one kind of writing system, as in the Korean Hangul orthography; syllabic writing is another; Japanese Kana is a possible example, although it may be more properly considered a *moraic* system as syllable weight is also significant in this writing system. In the first case the fundamental design principle is to map symbols to phonemes, and in the second the mapping is from symbols to syllables. It is possible for an orthography to make use of more than one kind of mapping. There are various ways in which that can occur.

For example, some have proposed that writing systems used for many Sinitic languages (languages of the Sino-Tibetan language family, e.g., Mandarin, Cantonese) are fundamentally morpho-syllabic in nature (Coulmas, 1989). Which is to say that each character is associated with a unit of meaning and a monosyllabic pronunciation. The large majority of characters/graphemes in these morpho-syllabic writing systems are composed of two elements, a "phonetic" element associated with pronunciation of a syllable and a "radical," which serves a morpho-semantic purpose. Albeit, single element characters are not uncommon, and three element characters occur as well; each character, regardless of constituency, corresponds to a single syllable at the phonological level. There are conventionally 214 radicals in the "simplified" character set used in the People's Republic of China. While the number of symbols occurring as phonetic elements is somewhat less clear, there may be on the order of 800 to 1000 in common use (DeFrancis, 1989). In the case of grapheme-to-morpheme mappings, the relationship is generally one-to-one, while for grapheme-to-syllable mappings the relationship is typically many-to-one, there being many more graphemes than syllables. For example, consider some examples from the "simplified" character set used for writing Mandarin in the People's Republic. The character 水 is a single element grapheme that corresponds to the word meaning "water" (pinyin *shui3*). In modified form this character serves as the semantic radical in more than a dozen two-element characters and the meaning of many of those has a relatively clear association with "water": 江 ("river"; pinyin *jiang1*), 洒 ("to sprinkle"; pinyin *sa3*), 漏 ("to leak/leaky";

pinyin *lou4*). The combining form of 水 can be seen at the left-hand side of each character, 氵. Next consider a few characters which all contain as their radical the character for "person," 人 (combining form 亻): 傲 ("proud/arrogant"; pinyin *ao4*), 仅 ("only"; pinyin *jin3*), 仿 ("to copy/to imitate"; pinyin *fang3*), 值 ("to value"; pinyin *zhi2*). Here, any semantic relationship among the words denoted by characters containing the "person" radical is considerably less clear. Other radicals may have an even more obscure relationship to the meanings of words denoted by the attendant characters. According to DeFrancis (1989) many "often offer no real semantic information at all and merely serve to differentiate one character from another, as do our spelling distinctions in *hair* and *hare*."

The phonetic components of characters are by contrast considerably more informative, although far from transparent. For example, the phonetic element on the right (square with vertical stroke) in the character 钟 ("bell/timer/alarm"; pinyin *zhong1*) is identical in pronunciation to the phonetic itself when used standalone (中, "in/within/during"; pinyin *zhong1*). As a component of other characters it is associated with similar but not identical pronunciations, as in: 冲 ("to charge in battle/to flush (toilet)"; pinyin *chong1*). More extreme examples of contextually determined pronunciation are not uncommon. Consider 淑 ("virtuous"; pinyin *shu1*) and 椒 ("pepper", pinyin *jiao1*), whereas their phonetic component, 叔, is pronounced *shu1*. The phonetic component of a character often provides a helpful clue as to its pronunciation, as when 皮 ("leather/skin"; pinyin *pi2*) is used as a phonetic in 披 ("to drape over one's shoulder"; pinyin *pi1*), or 坡 ("slope"; pinyin *po1*), or 跛 ("lame"; pinyin *bo3*). While this sort of inconsistency is common, DeFrancis (1989) estimates that the phonetic elements in about two-thirds of characters yield useful cues as to their pronunciation. Insofar as orthographic symbols map to phonology at the level of the syllable and to meaning at the level of the morpheme, then metalinguistic awareness important to acquisition of literacy in Sinitic languages might, arguably, target those two levels of analysis, rather than the phoneme as in alphabetic writing systems (Wang *et al.*, 2015; Zhang *et al.*, 2012; but also see Newman *et al.*, 2011).

Two additional points about Sinitic writing systems that are not obvious from the foregoing should be noted here. First, contrary to the examples just given, most words in Mandarin are written with two or more characters. A second feature, setting the Chinese orthographies apart from the more familiar European ones, is that in writing connected text there is no explicit indication of word boundaries (more on this later).

A different type of duality in mapping can be found in the Korean Hangul system. This is fundamentally an alphabetic system in that graphemes map to phonemes, but when written, graphemes are arranged in groups or blocks that correspond explicitly to syllables, with each block containing two to six graphemes (Coulmas, 1989). For example the name of the writing system, rendered in Korean, is 한 글. The leftmost block consists of three graphemes: ㅎ/h/(top-left), ㅏ/a/(top-right), and ㄴ/n/(bottom). The right block also consists of three graphemes, arranged vertically: ㄱ/g/, ㅡ/u/, ㄹ/l/. Here the dual nature of the mapping is entirely to do with different aspects of phonological structure. One might contrast this with the Vietnamese orthography, which is a roman-based alphabet, with the peculiar feature of using spaces to separate syllables and moreover lacking any specific indication of word boundaries. Modern Korean uses whitespace to mark word boundaries.

A third example of complex mapping can be found in orthographies used for writing most of the languages, of both Sanskritic and Dravidian origin, on the Indian sub-continent. These orthographies are sometimes called alphasyllabaries and their individual graphemes are referred to as *akshara* (Padakannaya & Ramachandra, 2011). While there is considerable diversity in this large group of orthographies they share a set of common characteristics (Coulmas, 1989): (a) each basic grapheme represents a consonant and an inherent vowel, and the inherent vowel is the same for all basic graphemes; (b) vowels other than the inherent vowel are represented by diacritics applied to basic graphemes; application of a diacritic replaces the inherent vowel with that represented by the diacritic; there is typically a special diacritic which suppresses the inherent vowel altogether; (c) consonant clusters are represented by ligatures of basic characters, and all but the final grapheme in a ligature loses its inherent vowel; (d) there is a set of graphemes used for initial vowels in syllables lacking a consonant onset. For example two basic akshara from the Gujarati alphasyllabary are ત and ક, representing the CV syllables/ta/and/ka/; the graphemes cannot be decomposed into separate elements corresponding to the consonant and vowel. Diacritics can be applied to these basic akshara to replace the inherent vowel with another, for example: તે/te/, તિ/ti/, તો /to/, તુ/tu/, and કે/ke/, કિ/ki/, કો/ko/, કુ/ku/. The individual akshara (basic grapheme plus diacritic) in these examples represent open syllables which can be broken down into phonemic components, although the diacritic vowels cannot stand alone (Padakannaya & Ramachandra, 2011). The diacritics are shown here, with the position of the dotted circle indicating the approximate relative location of the basic akshara: ◌ે/e/, ◌િ/i/, ◌ો /o/, ◌ુ/u/. Further, the inherent vowel can be suppressed with a special purpose diacritic called a *virama*, yielding a monophonemic (consonantal) grapheme: ત્/t/, ક્/k/. This alphasyllabic system differs from both the fundamentally alphabetic Korean Hangul system, described previously, and Japanese Kana in which each basic grapheme represents a syllable or mora in holistic fashion; Kana graphemes cannot be further analyzed into phonemic elements. Although, it has been argued that grapheme-to-phonology mappings in aksharic orthographies allow for (or even require) isolation of phonemes, and that this is evidence that such systems should be considered fundamentally alphabetic in nature (Rimzhim *et al.*, 2014).

Finally, we previously observed that the Chinese morpho-syllabic writing systems do not explicitly mark word boundaries, in contrast to the convention of inter-word spaces in European orthographies and elsewhere. Alphasyllabaries can be found in either camp. Gujarati makes use of spaces (probably the most common case for this type of writing system) while the Thai alphasyllabary does not indicate word boundaries.

As noted, the term *orthography* encapsulates details of the mapping from symbols to language elements. These details will necessarily include things like the specific symbol set, the specific linguistic units to be mapped (whether they be syllables, phonemes, morphemes or some admixture), and the specific mappings between linguistic units and symbols. Change any detail of the script, the linguistic units, or the mapping between them, and you have a different orthography. As a point of comparison, consider the cases of English and Dutch. The standard written forms of both languages use the same script, the same 26 grapheme set derived from the Roman script (leaving aside the issue of accents). However, because the phonemic structure of the two languages differs (as a result of language change over time), so too do many of the specific mappings between phonemes and graphemes, although there is some overlap due to historical connections between the languages and the origins of the script.

Regardless of the specific details of script and language, an ideal orthography would be one in which the mapping from symbols to linguistic units is one-to-one and perfectly consistent. But orthographies vary considerably in how closely they approximate the ideal (Lukatela *et al.*, 1980). The relative consistency of mapping from symbol to linguistic unit is referred to as *orthographic depth* (Lukatela *et al.*, 1980). Deep orthographies have complex mappings (e.g., standard writing conventions for English, Mandarin Chinese and Hebrew), while shallow orthographies more closely approximate the ideal one-to-one mapping (Korean Hangul, Japanese Kana, Spanish alphabet). The Orthographic Depth Hypothesis holds that shallower orthographies will be easier to learn due to their more consistent mappings from grapheme to linguistic unit, and that in such writing systems phonology will play a more prominent role in lexical access than will be the case for deeper orthographies (Frost *et al.*, 1987; Frost & Katz, 1989; Katz & Frost, 1992; Rao *et al.*, 2011; Schmalz *et al.*, 2015).

Orthographies (specific language-to-script mappings) with considerably greater overlap than that of the Dutch/English case mentioned above are not unusual. These may arise due to systematic differences in language usage across speakers of a single language; differences in pronunciation, word choice and even grammatical construction, are commonplace. Such differences are often linked to a speaker's identification with a particular social group that may be defined in part by culture or social class or geography (Wolfram, 2006). Details of a non-mainstream language variety that differ from the mainstream variety may complicate the acquisition of literacy by individuals or groups whose usage is not well-aligned with the mainstream. For example, it has been hypothesized that this kind of mismatch, with regard to contrasts in pronunciation between African American English (AAE) and Mainstream American English (MAE), may form an additional obstacle to the acquisition of literacy for children whose home dialect is AAE (Cunningham, 1976; Labov, 1995; LeMoine, 2001). While the evidence available for this particular case of orthographic mismatch finds little support for the hypothesis, researchers point out that children whose home dialect is predominantly AAE typically have enough knowledge of MAE by the time they reach school to mitigate any putative mismatch disadvantage (Patton Terry *et al.*, 2010; Patton Terry & Scarborough, 2011; Patton Terry, 2012). The question remains as to whether children coming from more insular non-mainstream language backgrounds might be hampered in their acquisition of orthographies tailored to unfamiliar language standards.

The terminological distinctions and conceptual foundation developed above give us the wherewithal to avoid certain confusions and ambiguities that are not unusual in even erudite discussions of "writing." For example, the Mandarin language is rendered in print using at least three different orthographies grounded in two different design principles: morpho-syllabic traditional characters in Taiwan, morpho-syllabic simplified characters in the People's Republic of China and alphabetic pinyin in the early grades in both (Cheung & Ng, 2003). In fact, it may not be unusual for the same language, or minimally different language varieties, to be written with very different scripts, as in the cases of Serbian/Croatian (Feldman *et al.*, 1985; Lukatela & Turvey, 1980) and Hindi/Urdu (Rao *et al.*, 2011). Moreover, two orthographies may be superficially similar in that they make use of the same script, yet map to languages that differ subtly (AAE/MAE) or markedly (English/Dutch) from one another.

Details of an orthography may have significant consequences for ease of lexical access by eye, the process by which the mental representation of a word's meaning or phonology becomes available for use. Evidence supports the hypothesis that the depth of an orthography modulates the ease with which beginning readers acquire it (Ellis *et al.*, 2004; Seymour *et al.*, 2003), or the ease with which fluent readers access the words that they know (Katz & Frost, 1992; Paulesu, 2006). Within an orthography, adult readers are faster to identify words with regular orthographic patterns versus irregular ones and this difference is greater for words that are less familiar (e.g., Katz *et al.*, 2005; Van Orden, 1990); other research shows that target letter identification is easier/faster when the letter string that contains the target is a word, versus nonword (e.g., Reicher, 1969; Wheeler, 1970). The ability to account for such effects is the minimum bar that must be cleared by any model of visual word recognition.

## Models of written word recognition

What follows is an admittedly incomplete and superficial survey of written word recognition models. The aim of such implemented models is to capture details of connections between single word recognition in print and speech modalities. A related issue, that of how to understand the relationship between lexical access and fluent reading of connected text, is discussed in the following section. A recurring theme in most models of visual word recognition is the presence of two modes of access to a lexical representation, given a particular orthographic stimulus. This dual path characteristic seems critical to accounting for the influence of certain lexical characteristics on the time course of word identification. The typical goal of a model is to predict human response times or accuracies for experimental tasks like word naming or lexical decision. In a typical naming task a participant is presented with the written representation of a single word and their charge is to simply pronounce it quickly and accurately. In a lexical decision task the participant is given an orthographic string, which may or may not represent a real word in the language at issue. Their task is to make a speeded judgment for each item as to whether or not it represents a real word and to press a button indicating that choice. A model's ability to emulate typical human performance is the standard against which it is judged. A significant limitation of essentially *all* current models of visual word recognition is that their organization typically incorporates specific assumptions about the target orthography and that this greatly limits the possibility of a completely general explanation of how word reading proceeds across orthographies (Rueckl, 2016). A recent collection of papers revealing the complexity and depth of the issues surrounding visual word recognition can be found in Grigorenko and Naples (2008).

In the early days of mechanistic word recognition models, Morton proposed that the mental lexicon could be simulated as a set of "logogens," essentially word detectors (Morton, 1969). Each logogen accumulates evidence in the form of sensory input or contextual information for the presence of a particular word. Only when evidence for a word exceeds a threshold does information associated with the word become available for subsequent processing (e.g., naming the word, or integrating it into a phrase). The earliest Logogen model held that lexical representations and mechanisms of access were amodal in all but the most superficial of sensory aspects. As such, a central prediction of the model is that there should be similar levels of lexical priming across modalities. Priming is an increase in speed or accuracy of word naming or lexical decision that occurs when the word under consideration is similar in some way (form or meaning) to a word that has been seen previously. The early Logogen model predicted that facilitation should be about equal whether a printed item was used to prime recognition of a spoken one, or vice versa. When data from experimental work with humans proved inconsistent with those predictions, later Logogen variants incorporated modality specific routes to the lexicon and constrained the potential for interaction between them (Morton, 1979, 1980). Other architecturally similar models emerged about this time in attempts to account for aspects of word recognition where logogen-based models failed. The Cohort model of speech recognition, for instance, dispensed with explicit thresholds for word detection, but retained the concept of individual word detectors, as well as the commitment that word-recognition is all-or-nothing; graded accumulation of evidence for words has no effect on subsequent processing until such a time as a word has been uniquely identified (Marslen-Wilson & Welsh, 1978).

The Dual Route Cascaded Model (DRC) of word identification is, like later versions of the Logogen model, explicitly designed to account for findings that supported a dissociation in modality specific paths to the lexicon, whence *dual-route* (Coltheart *et al.*, 1993; Coltheart & Rastle, 1994). One path to the lexicon is direct, relying on learned associations between orthographic forms, and phonology and semantics (the lexical route). A phonological path affords access to the lexicon by way of intermediate phonological representations, which are computed, or assembled, on the basis of learned orthography—phonology mappings. The correspondence rules that make this possible are built into the DRC, so it is capable of simulating skilled reading only; it is not a model of reading skill acquisition. However, mechanisms have been proposed for learning of grapheme-phoneme correspondence rules (Coltheart *et al.*, 1993; Pritchard *et al.*, 2016). The path to the lexicon by way of assembled phonological representations, while indirect, has the advantage of allowing access to words that are part of a reader's speech vocabulary, but whose orthographic representation is unfamiliar. It may also be critical to the ability to learn entirely new words from print (De Jong & Share, 2007; Share, 1995, 2011). Assembled phonology has the disadvantage, more so for some orthographies than others, that if a word's orthography—phonology mapping is inconsistent with the regular patterns for the orthography, then it may be difficult or impossible to derive a phonological form that is sufficient to support lexical access.

The DRC approach differs from the Logogen approach in that DRC models are cascaded, organized such that the output of each subprocess in a model is a set of continuous values that are always available for processing at the next level (McClelland, 1979). So DRC models provide for graded lexical activation across levels of processing. Word detection is neither thresholded nor all-or-nothing with regard to making lexical information available for subsequent processing. Further, consistent with the Interactive Activation Model (IAC) of printed word detection (McClelland & Rumelhart, 1981), information flow between subprocesses in the DRC can be bidirectional. The IAC and the DRC are explicit in representing words as *nodes*. As such, they can be considered symbolic systems. Nodes in the DRC are vaguely similar to logogens in the sense that, for each word in the lexicon, there is an individual node uniquely responsible for reflecting the current state of evidence for that word. Unlike logogens, this evidence is available to other subprocesses on a continuous basis (McClelland & Rumelhart, 1981; Coltheart *et al.*, 2001).

A fundamental premise of the DRC is that expert readers use both lexical and assembled phonological pathways to the lexicon during word recognition. Assembly of accurate phonological representations on the basis of orthographic input requires that the word under consideration conforms to the regular spelling patterns of the orthography. The assembled pathway will allow for the construction of a phonological representation even in the absence of a lexical entry, necessary when the word is unknown (perhaps a pseudoword); the assembled phonological representation can serve for naming the novel item. For known words, both pathways will be available. The lexical route provides a means to access a stored phonological representation without having to assemble it entirely on the fly. This is advantageous when a word's spelling does not conform to the regular spelling patterns for the orthography. However, it should be observed that an instance of phonological access in the DRC for the purpose of reading a known word aloud is not the result of a winner-take-all "race" between assembled and lexical routes, but rather a product of both (Coltheart *et al.*, 2001; Frost, 1998). In human performance, words whose spellings are in conflict with the regular spelling patterns of an orthography (irregular or exception words) result in longer reaction times for naming and identification, and trigger more errors than words with rule-governed spellings (Andrews, 1982; Treiman *et al.*, 1995), and the DRC captures these effects. This conjoint influence of the two paths, and their distinct mechanisms, allows for established lexicality effects in word and nonword naming, and also an explanation for the effect of spelling regularity being less evident in lexical decision than in word naming (e.g., Hino & Lupker, 2000; Andrews, 1982).

The DRC model was developed specifically to account for facts of printed word recognition in English, and it has been extended with some success to other alphabetic orthographies (Ziegler *et al.*, 2000; Ziegler *et al*., 2003). However, Coltheart *et al.* (2001) aver that "The Chinese, Japanese, and Korean writing systems are structurally so different from the English writing system that a model like the DRC model would simply not be applicable" (p. 236). The bases for parts of this claim are questionable: Korean Hangul is a fundamentally alphabetic orthography and so it's difficult to see why the DRC should not be expected to cover it. Regardless, Coltheart and colleagues make the strong claim that reading in alphabetic orthographies proceeds via altogether different mechanisms than reading in non-alphabetic orthographies. An overview of the development, structure and capabilities of the DRC can be found in Coltheart *et al.* (2001).

The Triangle Model (TM) of Seidenberg and colleagues (1989) is yet another influential framework for understanding word identification processes. It follows on earlier interactive models of word recognition (McClelland, 1979; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) and is strongly committed to non-symbolic lexical representations. Words are encoded as distributed patterns of activation in a connectionist network; there are no lexical nodes representing individual words, as exist in the DRC. The basic TM has seen several revisions since its introduction. In all versions, printed word identification proceeds through a combination of direct connections between orthography and semantics, and a phonologically mediated pathway from orthography to phonology to semantics. Orthographic input initiates patterns of activation that flow simultaneously through each pathway. Any given input will engender a distributed representation consisting of phonological, semantic, and orthographic information that corresponds to a lexical identity. Phonological access for both words and nonwords arises via the same pathways in the network. Flow of activation through the network from orthography to phonology can be viewed as roughly analogous to the lexical and assembled routes of the DRC (Harm & Seidenberg, 1999; Harm & Seidenberg, 2004; Plaut *et al.*, 1996). However, the distributed nature of lexical representations in the TM stands in distinct contrast to the DRC, with its local unitary representation of memorized word forms and production rules for decoding novel words. Further, the TM relies on the single mechanism of spreading activation through a connectionist network to derive its explanatory power.

A crucial difference between the two approaches is highlighted by considering the difference between orthographic regularity in the DRC, a categorical distinction among words, and orthographic consistency in the TM, a continuous graded property of words. The spelling of a word is either regular, its pronunciation is correctly determined by grapheme—phoneme correspondence rules, or not (MINT is regular; PINT is not), while orthographic consistency is graded in the sense that it can take values between 0 and 1. For example, the English spelling pattern -INK is consistent in that only one pronunciation is possible (e.g., *rink*, *fink*, *mink*), whereas the spelling pattern -AVE is not (*save*, *pave,* etc., versus *have)*. Consistency of a word is typically defined over its *orthographic neighborhood*, which is to say those words that are spelled similarly to it (Glushko, 1979). The –INK neighborhood is perfectly consistent, while the –AVE neighborhood is less so. Another important difference between the TM the DRC model is that the former directly incorporates a learning mechanism that allows it to induce regularities in orthographic mappings, rather than relying on a set of pre-specified correspondence rules (Seidenberg & McClelland, 1989).

It is interesting to note that orthographic consistency may have a relatively straightforward analog in the Mandarin morpho-syllabary. There, consistency can be defined as the congruence in pronunciation among characters sharing a phonetic component (Shu *et al.*, 2003). Several studies have reported effects of consistency, thus defined, on character naming that seem to cohere with more established results from word naming in alphabetic orthographies (e.g., Hsu *et al.*, 2009; Lee *et al.*, 2004; Li *et al.*, 2011). One recent effort to fit data from readers of Chinese using a modification of the Triangle Model has also met with some success. There, Yang and colleagues (2009) used a connectionist model to predict an interaction between frequency and consistency in Chinese word (character) naming, similar to that found in English word naming, in addition to an effect of regularity peculiar to Chinese. Data from a behavioral study of character naming by Chinese speakers subsequently confirmed their predictions (Yang *et al.*, 2009).

Variations on the Dual Route Cascaded Model and the Triangle Model are perhaps the strongest rivals among implemented mechanistic models of visual word recognition, with the edge going to the TM due to its explicit incorporation of a learning mechanism and to its stronger potential for applicability across writing systems. Yet, it must be acknowledged that most work toward understanding the reading process has focused on word recognition in alphabetic writing systems and more often than not rather explicitly on reading English (Share, 2008). That is certainly true of most of the research involving the models described in the preceding paragraphs. However, as Share observes, the English orthographies are rather unusual among alphabetic writing systems, being especially deep orthographies. Inferences derived from research on English speakers may not generalize well even to other alphabetic systems, and generalization to other types of writing systems is even more questionable. Share (2014) raises a somewhat related issue that he calls "alphabetism" in reading research. Alphabetism is the often unspoken assumption that alphabetic writing systems are inherently superior to those built around other design principles (e.g., syllabaries, morpho-syllabaries, alphasyllabaries). As Share observes, this assumption has been largely unexamined and seems to stem from a European bias in the reading research literature. As we have seen, orthographies show considerable diversity, and some researchers have questioned whether a unified account of reading across languages and orthographies is even possible (cf. Coltheart *et al.*, 2001). Both issues call for a better understanding of the diversity of writing systems and their implications for the process of visual word identification and reading in general.

The Lexical Constituency Model (LCM) reflects the recent move toward extending the coverage of mechanistic models of word recognition beyond alphabetic writing systems (Perfetti *et al.*, 2005). The model is motivated by emerging evidence that Chinese readers are sensitive to the phonological information provided by their writing system, that activation of phonological information by character graphemes can be as fast and automatic as activation of semantic information (Zhang & Perfetti, 1993; Perfetti & Tan, 1998; Xu *et al.*, 1999). In the implemented version of the LCM, inputs are encoded representations of characters. These connect to an orthographic level of representation, which itself has direct connections to a semantic level of representation and to a phonological level, the latter with separate encoding of syllabic onset (all phonemes in a syllable up to, but not including the vowel) and rhyme (those phonemes from the vowel onward) and tone. Further pathways also connect semantics to the syllable components represented in the model's phonological level. It bears repeating that a crucial difference between activation of phonological information by way of Chinese orthographies and similar initiation of phonological processing in alphabetic writing systems is that in Chinese the phonological units mapped to graphemes are syllables, not phonemes. So, any phonological route to the lexicon cannot proceed by way of phonological assembly of phonemic constituents, as is presumably the case with alphabetic systems. Therefore, Perfetti *et al.* (2005) make the assumption that written word identification necessarily includes "the recovery of a phonological object and its associated nonphonological components," (p. 46) even for writing systems like those used in Mandarin and other Sinitic languages. Perfetti (2007) reiterates the widely held notion that the mental representation of a word consists of an intersection of phonological, semantic and orthographic information. This general view is consistent with both distributed (e.g., the Triangle Model) and symbolic models (e.g., the DRC Model) of written word identification, including the Lexical Constituency Model. Under this view, the process of word recognition, or lexical access, becomes that of using available evidence, in the form of a written representation, to recover additional task-relevant components of a target word's lexical representation (Perfetti *et al.*, 2005).

These models and others focus on visual recognition of individual words. At the same time, it must be acknowledged that comprehension of connected text necessarily involves additional work beyond word recognition including, at least, derivation of syntactic (e.g., Rayner *et al.*, 1983; MacDonald *et al.*, 1994; Frazier & Clifton, 1997; Hale, 2003) and ultimately discourse representations (e.g., Kintsch & Kintsch, 2005 ; Marslen-Wilson *et al.*, 1993; McKoon & Ratcliff, 1988). The Dual Route Cascaded, Triangle and Lexical Constituency frameworks focus on word recognition, treating post-lexical integration as a black box with the assumption that whatever mechanisms are responsible for syntactic and discourse processing of spoken language are used also for print language input. Any potential differences in print and speech processing above the level of the word would seem to derive from the fundamental contrast in the temporal persistence of print and speech signals, a distinction that is independent of any specific orthography.

## Models of gaze behavior during reading

In order to read connected text, a necessary sequel to visual recognition of individual words is the scanning behavior that moves the eye from one word to the next as the linguistic content of the text is processed. Scanning serves to move words across the *fovea*, the highest resolution portion of the visual field, for optimally efficient word recognition (Inhoff, 1989; Rayner, 1975; Rayner *et al.*, 1982). The fovea extends about $2^0$ through the most central portion of the visual field, while an area of lesser acuity, the parafovea, extends another $3^0$ beyond the foveal limit. Words are identified more quickly and accurately in the foveal region, although information useful to the reader can be obtained for text in the parafoveal region as well (Rayner & Morrison, 1981). For example, spaces indicating word boundaries are readily detectable in the parafovea (Sheridan *et al.*, 2013; Slattery & Rayner, 2013). Gaze patterns over text are characterized by two distinct phases: *fixation*, in which the point of regard is relatively unchanging, and *saccade*, which is a rapid shift of gaze from one position to the next (Rayner, 1998). Several models have emerged in which details of the coordination of gaze behavior with word recognition processes are proposed to explain facts about the nominal durations of gaze on each word as well as the distribution of fixation locations within a text.

There is some controversy as to whether eye movements are directly regulated by cognitive dictates of linguistic processes (Engbert *et al.*, 2002; Reichle *et al.*, 1998), or whether surface perceptual features of the text (visual features at the *script* level) are the primary governors of gaze patterns (McConkie *et al.*, 1994; Reilly & O'Regan, 1998). In the first instance, linguistic processes like word recognition or contextual integration have a direct role in driving gaze behavior, while in the latter case linguistic processes serve only to modulate gaze behavior that is primarily driven by low-level visual information, word length inferred from spaces between words, for example. Nonetheless, it is well-established that gaze patterns over text are influenced by text characteristics, from lexical to syntactic to pragmatic, as well as reader characteristics like decoding skill (Summarized in: Rayner *et al.*, 2006; Staub & Rayner, 2007; Rayner *et al.*, 2013). The cognitively oriented model E-Z Reader model, which gives a direct role to linguistic processing (lexical access and contextual integration) in determining the timing and location of fixations over text (Reichle *et al.*, 1998), will be the focus of this section.

Reichle *et al.* (1998) introduced the E-Z Reader framework, or family of models, with the goal of accounting for interactions of visual processing, attention, and lexical processing in guiding the eye movements of readers (Reichle *et al.*, 2006). It is important to understand that E-Z Reader models do not aspire to provide deep accounts of word identification, sentence parsing, or eye-movement control, but instead should be seen as an account of how the relationships among these processes drive the location and timing of eye movements over print. A central feature of the E-Z Reader framework, is that attention for the purpose of lexical access is allocated serially, to one word at a time. Lexical access is modeled as a two-stage process: a "familiarity check" presumed to focus on recognition of formal properties of the word (phonology) begins when attention is allocated to its visual features; a subsequent "completion" stage involves retrieval of syntactic and sematic properties of the word from memory (Reichle *et al.*, 2009). More specifically, the second stage can be seen as representing some minimal amount of processing that must be carried out for the current word before attention can be shifted to the next. Two factors influence lexical access times in the various E-Z Reader models. Word frequency, estimated through corpus counts, stands proxy for general lexical properties (e.g., familiarity, consistency). Predictability in context, typically operationalized as cloze probability, can be seen as an estimate of difficulty associated with integrating a word into its syntactic and semantic context. In earlier versions of E-Z Reader, post-lexical integration of a word into preceding context was not handled separately from the completion phase of lexical processing, whereas contextual integration receives more explicit treatment in more recent models (Reichle *et al.*, 2009).

A second key feature of the E-Z Reader framework is its decoupling of saccadic programming and execution, and lexical processing. Once the familiarity check is complete, saccade programming proceeds in parallel with the completion phase of lexical access. This sets up a race between substantial completion of lexical access and completion of the motor program which shifts gaze to the next point of regard. Design of the model is such that lexical access sufficient to release attention always finishes before saccade programming. Upon completion of lexical access, attention shifts covertly to the next word, while gaze lags somewhat. The first stage of lexical processing begins as soon as a word becomes the focus of attention. Variation in the magnitude of the lag between attention and gaze as they step through a text is a function of the difficulty of processing the current word, $n$. When $n$ is difficult, low frequency or low predictability, then the lag is shorter and so the duration of pre-fixation attention devoted to the subsequent word is shorter. This variable lag is the mechanism by which the model accounts for "spillover" effects such that fixation time on word $n+1$ is influenced by properties of word $n$ (Just & Carpenter, 1978; Rayner *et al.*, 1989; Warren *et al.*, 2011).

In E-Z Reader, the time needed for saccade programming is a function of two random variables with fixed parameters, the first corresponding to a labile phase of programing, wherein the program can be influenced by external factors, and the second to a non-labile phase. Parafoveal processing of a word can also play a role in the timing of eye movements (Sheridan & Reichle, 2015). In general, the choice of where to move the eyes is a function of the optimal viewing location of the next word (McConkie *et al.*, 1988; McConkie *et al.*, 1989; Vitu *et al.*, 1990), plus a stochastic component to emulate error in the oculomotor program. Word skipping behavior can also be simulated by the model; E-Z reader will cause word $n+1$ to be skipped if its familiarity check finishes before the labile portion of the saccade program to shift gaze from word $n$ to $n+1$ is complete. In this case the program to fixate word $n+1$ is canceled and replaced by a new program to shift gaze to word $n+2$ (Pollatsek *et al.*, 2006; Reichle *et al.*, 2012). Simulations with E-Z Reader models by Reichle and colleagues discussed above have reproduced a number of benchmark phenomena from studies of gaze behavior in reading alphabetic orthographies, including effects of word frequency (e.g., Gong *et al.*, 2016; Just & Carpenter, 1980; Raney & Rayner, 1995; Valle *et al.*, 2013) and predictability (e.g., Braze *et al.*, 2002; Husain *et al.*, 2015; Kliegl *et al.*, 2006) on reading times, and word length on fixation positions (e.g., Vitu *et al.*, 1990; Joseph *et al.*, 2009).

In reading Chinese, just as for the alphabetic orthographies for which E-Z Reader models were first developed, word predictability in context and word frequency influence eye movements in expected ways: predictability is inversely related to fixation time (Rayner *et al.*, 2005), as is frequency (Yan *et al.*, 2006). However, the lack of any explicit indication of word boundaries in Chinese orthographies sets it apart from the European alphabets that have served as the forge for E-Z reader (a feature shared with some other orthographies, e.g., Thai, Japanese). This lacuna might increase the need for top-down information in extracting words from text in such writing systems. Evidence suggests that word boundaries in Thai, for example, are identified based on the distributional properties of graphemes (Kasisopa *et al.*, 2013; Reilly *et al.*, 2011), using mechanisms perhaps not dissimilar from those engaged by listeners in identifying words within continuous speech (e.g., Frank *et al.*, 2013; Saffran *et al.*, 1996). Some studies have asked whether inserting word-delimiting spaces into Chinese text would have a facilitative effect on reading, and the general finding is that the presence of such spaces either has no effect on reading times (Bai *et al.*, 2008) or that reading times are indeed reduced (Hsu & Huang, 2000b, 2000a). The same studies demonstrate that insertion of word-disrupting spaces has the effect of slowing reading times (also see Li *et al.*, 2009). These rather surprising results would seem to indicate that reducing the need for top-down information in guiding eye movements has a facilitative effect even in orthographies where unspaced text is the norm. One effort to adapt E-Z Reader to reading Chinese assumes that readers have deterministic knowledge of word boundaries when reading conventionally unspaced text (Rayner *et al.*, 2007). Given this somewhat unlikely assumption, the model is able to describe characteristic patterns of gaze over Chinese text with fair accuracy. That said, there is work to be done in understanding the mechanisms of word segmentation in orthographies that lack explicit cues. Some additional discussion of the problems for E-Z Reader vis a vis word segmentation and saccade targeting can be found in Liu *et al.* (2015).

## Language comprehension and reading

Language comprehension is the product of external sources of information, both linguistic and contextual, interacting with internal knowledge and processes. Regardless of whether the input modality is auditory or visual, the complexity of information and processes that generate a percept of linguistic meaning contribute to a state of affairs where, of necessity, theorists typically focus on generating testable models of *component* systems (e.g., lexical structure, lexical access, syntactic processing, discourse representation), rather than on soup-to-nuts models of comprehension as such. Some components of a comprehension model may represent greater theoretical and methodological challenges than others, or simply be better developed for a variety of reasons.

It will also be worthwhile to consider Gough and Tunmer's Simple View of Reading (SVR; 1986; Tunmer & Chapman, 2012), which has been influential in framing work on development of reading comprehension and its connections to oral language comprehension and visual word recognition. The Simple View states that comprehension of written language is the product of two capacities: the ability to decode, or to access lexical representations by way of their print forms, and the capacity for general language comprehension (typically operationalized through measures of oral language comprehension). If an individual has good oral language skills, but no familiarity whatsoever with their language's written form, then they will have no ability to read; if decoding skill is less than perfect this will impose real limits on the ability to comprehend language in its printed form. Conversely, if an individual has good decoding skills (visual word recognition skills), then their ability to comprehend print will be limited by their ability to comprehend language in general, where the constraints may arise from limits on specific vocabulary or lack of familiarity with complex grammatical structures or discourse devices. Regardless, the SVR holds that the single crucial difference between efficient processing of written versus spoken language lay in the input modality of words. Once words are recognized, subsequent processing proceeds in the same manner regardless of whether the original modality was visual or acoustic; the mechanisms involved in parsing, mental model construction and inferencing are essentially amodal.

In this connection, Braze *et al.* (2011) present evidence from a neuro-imaging study of sentence processing in print and speech. They used functional magnetic resonance imaging to examine brain activity in experienced readers while they read or listened to matched sentences that were designed to challenge specific aspects of comprehension; input modality was a within-subject manipulation. The brain regions engaged by challenging materials, largely confined to the left inferior frontal gyrus and the left posterior superior temporal gyrus, correspond approximately to regions that previous studies had identified as being sensitive to differences in sentence complexity (e.g., Constable *et al.*, 2004; Michael *et al.*, 2001). Additional analyses confirmed that predominantly left-hemisphere frontal and temporal regions responded in a similar way to comprehension challenges posed by the experimental sentences regardless of whether they were presented in printed or spoken form. Their findings support the existence of an amodal language system, which integrates linguistic inputs arising from different modalities such that speech and print engage a common underlying code (Braze *et al.*, 2011; also see: Frost *et al.*, 2009; Shankweiler *et al.*, 2008).

There has been considerable work indicating that the two components of the Simple View, decoding and language comprehension, are correlated but nonetheless distinct capacities for a variety of populations and developmental stages (e.g., Braze *et al.*, 2007; Braze *et al.*, 2016; Catts *et al.*, 2006; Dreyer & Katz, 1992; Joshi & Aaron, 2000; Landi, 2010). The SVR also makes an interesting prediction with regard to developmental changes in the specific contributions of oral language comprehension capacity and decoding skill. In early grades, where children are still learning to read, reading comprehension is clearly limited by a child's decoding skill, but as printed word recognition skills become automatized the importance of general language skills as a constraint on reading comprehension will increase and the importance of decoding skill as a limiting factor will decrease. The body of work in this area seems to confirm that supposition (e.g., Gough *et al.*, 1996; García & Cain, 2014). So, the idea that mechanisms of language processing are largely independent of input modality should be qualified by the impact of differences in reading skill regardless of the details of language or orthography.

## Conclusion

Orthographies differ from one another in terms of their scripts, the visual characteristics of their graphemes, and in terms of the nature of the mapping from script to linguistic unit. Orthographic depth, the complexity of the mapping from script to linguistic unit, modulates the ease with which an orthography is learned, and within an orthography the consistency or regularity of a particular orthographic pattern will temper the difficulty with which written words containing that pattern can be recognized. There is some hope for a unified cross-orthography account of visual word recognition, although details of how non-alphabetic writing systems (e.g., syllabaries, morpho-syllabaries, alphasyllabaries) may be fit into theoretical frameworks built on a foundation of empirical work on alphabetic reading are not entirely clear. Indeed, there remains a substantial gap in the literature with regard to research on reading and literacy in languages that make use of non-alphabetic writing systems. This euro-centric alphabetism in reading research continues to limit advances in our understanding of the potential for literacy as a universal human capacity.

## Acknowledgments

The authors are grateful to Jenna Duclos for assistance with preparation of this chapter. The authors' research is supported by National Institutes of Health grant R01 HD071988 to Haskins Laboratories.

## REFERENCES

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition*, *10*(6), 565–575. doi: 10.3758/BF03202439

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1277–1288. doi: org/10.1037/0096-1523.34.5.1277

Bloomfield, L. (1942). Linguistics and reading. *The Elementary English Review*, *19*(4), 125–130.

Bowey, J. A. (1995). Socioeconomic-status differences in preschool phonological sensitivity and first-grade reading-achievement. *Journal of Educational Psychology*, *87*(3), 476–487. doi: 10.1037/0022-0663.87.3.476

Braze, D., Katz, L., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., Gong, T., Johns, C. L., & Shankweiler, D. P. (2016). Vocabulary does not complicate the simple view of Reading. *Reading and Writing*, *29*(3), 435–451. doi: 10.1007/s11145-015-9608-6

Braze, D., Mencl, W. E., Tabor, W., Pugh, K. R., Constable, R. T., Fulbright, R. K., Magnuson, J. S., Van Dyke, J. A., & Shankweiler, D. P. (2011). Unification of sentence processing via ear and eye: An fMRI study. *Cortex*, *47*(4), 416–431. doi: 10.1016/j.cortex.2009.11.005

Braze, D., Shankweiler, D. P., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*, *31*(1), 25–44.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities*, *40*(3), 226–243. doi: 10.1177/00222194070400030401

Byrne, B., Fielding-Barnsley, R., & Ashley, L. (2000). Effects of preschool phoneme identity training after six years: Outcome level distinguished from rate of response. *Journal of Educational Psychology*, *92*(4), 659–667. doi: 10.1037//0022-0663.92.4.659

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech Language and Hearing Research*, *49*(2), 278–293. doi: 10.1044/1092-4388(2006/023)

Cheung, H., & Ng, L. K. H. (2003). Chinese reading development in some major Chinese societies: An introduction. In C. McBride-Chang & H.-C. Chen (Eds.), *Reading Development in Chinese Children* (pp. 3–17). Westport, Connecticut: Praeger Publishers/Greenwood Publishing Group

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608. doi: 10.1037/0033-295x.100.4.589

Coltheart, M., & Rastle, K. (1994). Serial Processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1197–1211. doi: 10.1037/0096-1523.20.6.1197

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. doi: 10.1037/0033-295X.108.1.204

Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., & Shankweiler, D. P. (2004). Sentence complexity and input modality effects in sentence comprehension: An fMRI study. *Neuroimage*, *22*(1), 11–21. doi: 10.1016/j.neuroimage.2004.01.001

Coulmas, F. (1989). *The Writing Systems of the World*. Cambridge, Massachussetts: Blackwell

Cunningham, P. M. (1976). Teachers' correction responses to Black-dialect miscues which are non-meaning-changing. *Reading Research Quarterly*, *12*(4), 637–653. doi: 10.2307/747445

De Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading*, *11*(1), 55–71. doi: 10.1207/s1532799xssr1101_4

DeFrancis, J. (1989). *Visible Speech: the diverse oneness of writing systems*

Dreyer, L. G., & Katz, L. (1992). An examination of "The simple view of reading." In D. J. Leu & C. K. Kinzer (Eds.), *Yearbook of the National Reading Conference: Literacy research, theory and practice: Views from many perspectives* (Vol. 41, pp. 169–175). Chicago, IL: National Reading Conference

Ellis, N. C., Natsume, M., Stavropoulou, K., Van Daal, V. H. P., Polyzoe, N., Tsipa, M.-L., & Petalas, M. (2004). The effects of orthographic depth on learning to read alphabetic, syllabic and logographic scripts. *Reading Research Quarterly*, *39*(4), 438–468. doi: 10.1598/RRQ.39.4.5

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*(5), 621–636. doi: 10.1016/S0042-6989(01)00301-7

Feldman, L. B., Lukatela, G., & Turvey, M. T. (1985). Effects of phonological ambiguity on beginning readers of Serbo-Croatian. *Journal of Experimental Child Psychology*, *39*(3), 492–510. doi: 10.1016/0022-0965(85)90053-0

Frank, M. C., Tenenbaum, J. B., & Gibson, E. (2013). Learning and long-term retention of large-scale artificial languages. *PLoS ONE*, *8*(1), e52500. doi: 10.1371/journal.pone.0052500

Frazier, L., & Clifton, C., Jr. (1997). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, *26*(3), 277–295. doi: 10.1023/A:1025024524133

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*(1), 71–99. doi: 10.1037//0033-2909.123.1.71

Frost, R., & Katz, L. (1989). Orthographic depth and the interaction of visual and auditory processing in word recognition. *Memory & Cognition*, *17*(3), 302–310. doi: 10.3758/bf03198468

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 104–115. doi: 10.1037/0096-1523.13.1.104

Frost, S. J., Landi, N., Mencl, W. E., Sandak, R., Fulbright, R. K., Tejada, E. T., Jacobsen, L., Grigorenko, E. L., Constable, R. T., & Pugh, K. R. (2009). Phonological awareness predicts activation patterns for print and speech. *Annals of Dyslexia*, *59*(1), 78–97. doi: 10.1007/s11881-009-0024-y

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in english. *Review of Educational Research*, *84*(1), 74–111. doi: 10.3102/0034654313499616

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 674–691. doi: 10.1037/0096-1523.5.4.674

Gong, T., Braze, D., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., & Shankweiler, D. P. (2016). Interactions between reading skills and lexical properties on on-line sentence reading. Gainesville, FL: Poster Presented at the 29th Annual CUNY Conference on Human Sentence Processing.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 1–13). Mahwah, NJ: Lawrence Erlbaum Assoc.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *RASE: Remedial & Special Education*, *7*(1), 6–10. doi: 10.1177/074193258600700104

Grigorenko, E. L., & Naples, A. J. (Eds.). (2008). *Single-Word Reading: Behavioral and Biological Perspectives*. New York: Lawrence Erlbaum Associates.

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, *32*(2), 101–123. doi: 10.1023/A:1022492123056

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528. doi: 10.1037/0033-295X.106.3.491

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), 662–720. doi: 10.1037/0033-295x.111.3.662

Hino, Y., & Lupker, S. (2000). Effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 166–183. doi: 10.1037//0096-1523.26.1.166

Hsu, C.-H., Tsai, J.-L., Lee, C.-Y., & Tzeng, O. J. L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: An event-related potential study. *Brain and Language*, *108*(1), 56–66. doi: 10.1016/j.bandl.2008.09.002

Hsu, S.-H., & Huang, K.-C. (2000a). Effects of word spacing on reading Chinese text from a video display terminal. *Perceptual and Motor Skills*, *90*(1), 81–92. doi: 10.2466/pms.2000.90.1.81

Hsu, S.-H., & Huang, K.-C. (2000b). Interword spacing in Chinese text layout. *Perceptual and Motor Skills*, *91*(2), 355–365. doi: 10.2466/pms.2000.91.2.355

Huey, E. B. (1908). *The Psychology and Pedagogy of Reading*. New York: MacMillan. doi: 10.2307/2177563

Husain, S., Vasishth, S., & Srinivasan, N. (2015). Integration and prediction difficulty in Hindi sentence comprehension: Evidence from an eye-tracking corpus. *Journal of Eye Movement Research*, *8*(2), 3, 1–12.

Inhoff, A. W. (1989). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception & Performance*, *15*(3), 544–555. doi: 10.1037/0096-1523.15.3.544

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, *49*(16), 2078–2086. doi: 10.1016/j.visres.2009.05.015

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, *21*(2), 85–97. doi: 10.1080/02702710050084428

Just, M. A., & Carpenter, P. A. (1978). Inference processes during reading: Reflections from eye fixations. In J. W. Sanders, D. F. Fisher & R. A. Monty (Eds.), *Eye Movements and the Higher Psychological Functions*. Hillsdale, NJ: Lawrence Erlbaum Assoc.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. doi: 10.1037/0033-295X.87.4.329

Kasisopa, B., G. Reilly, R., Luksaneeyanawin, S., & Burnham, D. (2013). Eye movements while reading an unspaced writing system: The case of Thai. *Vision Research*, *86*, 71–80. doi: 10.1016/j.visres.2013.04.007

Katz, L., & Frost, R. (1992). The Reading Process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology*, *94*, 67–84. doi: 10.1016/S0166-4115(08)62789-2

Katz, L., Lee, C. H., Tabor, W., Frost, S. J., Mencl, W. E., Sandak, R., Rueckl, J., & Pugh, K. R. (2005). Behavioral and neurobiological effects of printed word repetition in lexical decision and naming. *Neuropsychologia*, *43*(14), 2068–2083. doi: 10.1016/j.neuropsychologia.2005.03.022

Kavanagh, J. F. (1968). *Proceedings of the Conference on Communicating by Language: the reading process; New Orleans, LA: February 11-13, 1968*. Bethesda, MD: U.S. Dept. of Health, Education and Welfare.

Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Current Issues in Reading*

*Comprehension Assessment* (pp. 71–92). Mahwah, NJ: Lawrence Erlbaum Associates.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*(1), 12–35. doi: 10.1037/0096-3445.135.1.12

Labov, W. (1995). Can reading failure be reversed: A linguistic approach to the question. In V. L. Gadsden & D. A. Wagner (Eds.), *Literacy among African-American youth : issues in learning, teaching, and schooling* (pp. 39–68). Cresskill, N.J.: Hampton Press.

Landi, N. (2010). An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Reading and Writing*, *23*(6), 701–717. doi: 10.1007/s11145-009-9180-z

Lee, C.-Y., Tsai, J.-L., Kuo, W.-J., Yeh, T.-C., Wu, Y.-T., Ho, L.-T., Hung, D. L., Tzeng, O. J. L., & Hsieh, J.-C. (2004). Neuronal correlates of consistency and frequency effects on Chinese character naming: An event-related fMRI study. *Neuroimage*, *23*(4), 1235–1245. doi: 10.1016/j.neuroimage.2004.07.064

LeMoine, N. (2001). Language variation and literacy acquisition in African American students. In J. L. Harris, A. G. Kamhi & K. E. Pollock (Eds.), *Literacy in African American communities* (pp. 169–194). Mahwah, NJ: Lawrence Earlbaum Associates.

Li, Q.-L., Bi, H.-Y., Wei, T.-Q., & Chen, B.-G. (2011). Orthographic neighborhood size effect in Chinese character naming: Orthographic and phonological activations. *Acta Psychologica*, *136*(1), 35–41. doi: 10.1016/j.actpsy.2010.09.012

Li, X., Rayner, K., & Cave, K. R. (2009). On the segmentation of Chinese words during reading. *Cognitive Psychology*, *58*(4), 525–552. doi: 10.1016/j.cogpsych.2009.02.003

Liberman, A. M. (1988). Reading is hard just because listening is easy. *Haskins Laboratories Status Report on Speech Research*, *95-96*(July-Dec), 145–150.

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1–33). Ann Arbor, MI: The University of Michigan Press.

Liu, Y., Reichle, E. D., & Li, X. (2015). Parafoveal processing affects outgoing saccade length during the reading of Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1229–1236. doi: 10.1037/xlm0000057

Lukatela, G., Popadić, D., Ognjenović, p., & Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition*, *8*(2), 124–132. doi: 10.3758/BF03213415

Lukatela, G., & Turvey, M. T. (1980). Some experiments on Roman and Cyrillic alphabets. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, Reading, and Dyslexia*. Baltimore, MD: University Park Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Syntactic ambiguity resolution as lexical ambiguity resolution. In C. Clifton, Jr., L. Frazier & K. Rayner (Eds.), *Perspectives on Sentence Processing*, (pp. 123–153). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Marslen-Wilson, W., Tyler, L. K., & Koster, C. (1993). Integrative processes in utterance resolution. *Journal of Memory and Language*, *32*(5), 647–666. doi: 10.1006/jmla.1993.1033

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63. doi: 10.1016/0010-0285(78)90018-X

Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by Ear and by Eye* (pp. 133–147). Cambridge, MA: MIT Press.

McClelland, J. L. (1979). On the time relations of mental processes: an examination of systems of processes in cascade. *Psychological Review*, *86*(4), 287–330. doi: 10.1037/0033-295X.86.4.287

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407. doi: 10.1037/0033-295X.88.5.375

McConkie, G. W., Kerr, P. W., & Dyre, B. P. (1994). What are "normal" eye movements during reading: Toward a mathematical description. In J. Ygge & G. Lennerstrand (Eds.), *Eye Movements in Reading*. Amsterdam: Pergamon Press.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: 1. The location of initial eye fixations on words. *Vision Research*, *28*(10), 1107–1118. doi: 10.1016/0042-6989(88)90137-X

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1989). Eye movement control during reading: II frequency of refixating a word. *Perception and Psychophysics*, *46*(3), 245–253. doi: 10.3758/bf03208086

McKoon, G., & Ratcliff, R. (1988). Contextually relevant aspects of meaning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*(2), 331–343. doi: 10.1037/0278-7393.14.2.331

Michael, E. B., Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). fMRI investigation of sentence comprehension by eye and by ear: Modality fingerprints on cognitive processes. *Human Brain Mapping*, *13*(4), 239–252. doi: 10.1002/hbm.1036

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165–178. doi: 10.1037/h0027366

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kolers, M. E. Wrolstad & H. Bouma (Eds.), *Processing of Visible Language* (pp. 259–268). New York: Plenum Press. doi: 10.1007/978-1-4684-0994-9_15

Morton, J. (1980). The logogen model and orthographic structure. In U. Frith (Ed.), *Cognitive Processes in Spelling* (pp. 117–132). London: Academic Press.

Newman, E. H., Tardif, T., Huang, J., & Shu, H. (2011). Phonemes matter: The role of phoneme-level awareness in emergent Chinese readers. *Journal of Experimental Child Psychology*, *108*(2), 242–259. doi: 10.1016/j.jecp.2010.09.001

Oakhill, J. V., & Cain, K. (2011). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*(in press). doi: 10.1080/10888438.2010.529219

Padakannaya, P., & Ramachandra, N. B. (2011). Reading Akshara: Indian alphasyllabary. In P. McCardle, B. Miller, J. R. Lee & O. J. L. Tzeng (Eds.), *Dyslexica Across Languages* (pp. 76–95). Baltimore, MD: Paul H. Brookes Publishing Co.

Patton Terry, N. (2012). Examining relationships among dialect variation and emergent literacy skills. *Communication Disorders Quarterly*, *33*(2), 67–77. doi: 10.1177/1525740110368846

Patton Terry, N., Connor, C. M., Thomas-Tate, S., & Love, M. (2010). Examining relationships among dialect variation, literacy skills, and school context in first grade. *Journal of Speech, Language & Hearing Research*, *53*(1), 126–145. doi: 10.1044/1092-4388(2009/08-0058)

Patton Terry, N., & Scarborough, H. S. (2011). The Phonological Hypothesis as a valuable framework for studying the relation of dialect variation to early reading skills. In S. Brady, D. Braze & C. A. Fowler (Eds.), *Explaining Individual Differences in Reading: Theory and Evidence*. New York: Psychology Press.

Paulesu, E. (2006). On the advantage of "shallow" orthographies: Number and grain size of the orthographic units or consistency per se? *Developmental Science*, *9*(5), 443–444. doi: 10.1111/j.1467-7687.2006.00525.x

Perfetti, C., Liu, Y., Fiez, J. A., Nelson, J., Bolger, D. J., & Tan, L.-H. (2007). Reading in two writing systems: Accommodation and assimilation of the brain's reading network. *Bilingualism: Language and Cognition*, *10*(2), 131–146. doi: 10.1017/S1366728907002891

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. doi: 10.1080/10888430701530730

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review*, *112*(1), 43–59. doi: 10.1037/0033-295X.112.1.43

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 101–118. doi: 10.1037/0278-7393.24.1.101

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115. doi: 10.1037/0033-295X.103.1.56

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*(1), 1–56. doi: 10.1016/j.cogpsych.2005.06.001

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2016). Modelling the implicit learning of phonological decoding from training on whole-word spellings and pronunciations. *Scientific Studies of Reading*, *20*(1), 49–63. doi: 10.1080/10888438.2015.1085384

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale*, *49*(2), 151–173. doi: 10.1037/1196-1961.49.2.151

Rao, C., Vaid, J., Srinivasan, N., & Chen, H.-C. (2011). Orthographic characteristics speed Hindi word naming but slow Urdu naming: Evidence from Hindi/Urdu biliterates. *Reading and Writing*, *24*(6), 679–695. doi: 10.1007/s11145-010-9256-9

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81. doi: 10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. doi: 10.1037/0033-2909.124.3.372

Rayner, K., Ardoin, S. P., & Binder, K. S. (2013). Children's eye movements in reading: A commentary. *School Psychology Review*, *42*(2), 223–233.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 358–374. doi: 10.1016/S0022-5371(83)90236-0

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241–255. doi: 10.1207/s1532799xssr1003_3

Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review*, *12*(6), 1089–1093. doi: 10.3758/BF03206448

Rayner, K., Li, X., & Pollatsek, A. (2007). Extending the E-Z Reader model of eye movement control to Chinese readers. *Cognitive Science*, *31*(6), 1021–1033. doi: 10.1080/03640210701703824

Rayner, K., & Morrison, R. E. (1981). Eye movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society*, *17*(3), 135–138. doi: 10.3758/bf03333690

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, *4*(3-4), 21–49. doi: 10.1080/01690968908406362

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics*, *31*(6), 537–550. doi: 10.3758/BF03204186

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275–280. doi: 10.1037/H0027768

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences*, *13*(3), 115–119. doi: 10.1016/j.tics.2008.12.002

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125–157. doi: 10.1037/0033-295X.105.1.125

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E–Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, *7*(1), 4–22. doi: 10.1016/j.cogsys.2005.07.002

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in nonreading tasks: a unified framework for understanding the eye-mind link. *Psychological Review*, *119*(1), 155–185. doi: 10.1037/a0026473

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 1–21. doi: 10.3758/PBR.16.1.1

Reilly, R. G., Aranyanak, I., Yu, L., Yan, G., & Tang, S. (2011). Eye movement control in reading Thai and Chinese. *Xinli Yu Xingwei Yanjiu*, *9*(1), 35–44.

Reilly, R. G., & O'Regan, J. K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research*, *38*(2), 303–317. doi: 10.1016/S0042-6989(97)87710-3

Rimzhim, A., Katz, L., & Fowler, C. A. (2014). Brāhmī-derived orthographies are typologically Āksharik but functionally predominantly alphabetic. *Writing Systems Research*, *6*(1), 41–53. doi: 10.1080/17586801.2013.855618

Rueckl, J. G. (2016). Toward a theory of variation in the organization of the word reading system. *Scientific Studies of Reading*, *20*(1), 86–97. doi: 10.1080/10888438.2015.1103741

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*(1), 60–94. doi: 10.1037/0033-295X.89.1.60

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. doi: 10.1006/jmla.1996.0032

Schmalz, X., Marinus, E., Coltheart, M., & Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomics Bulletin & Review*, *22*(6), 1614–1629. doi: 10.3758/s13423-015-0835-2

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568. doi: 10.1037/0033-295X.96.4.523

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143. doi: 10.1348/000712603321661859

Shankweiler, D. P., Mencl, W. E., Braze, D., Tabor, W., Pugh, K. R., & Fulbright, R. K. (2008). Reading differences and brain: Cortical integration of speech and print in sentence processing varies with reader skill. *Developmental Neuropsychology*, *33*(6), 745–776. doi: 10.1080/87565640802418688

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218. doi: 10.1016/0010-0277(94)00645-2

Share, D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin*, *134*(4), 584–615. doi: 10.1037/0033-2909.134.4.584

Share, D. L. (2011). On the role of phonology in reading acquisition: The self-teaching hypothesis. In S. A. Brady, D. Braze & C. A. Fowler (Eds.), *Explaining Individual Differences in Reading: Theory and Evidence* (pp. 45–68). New York, NY: Psychology Press.

Share, D. L. (2014). Alphabetism in reading science. *Frontiers in Psychology*, *5*(752). doi: 10.3389/fpsyg.2014.00752

Sheridan, H., Rayner, K., & Reingold, E. M. (2013). Unsegmented text delays word identification: Evidence from a survival analysis of fixation durations. *Visual Cognition*, *21*(1), 38–60. doi: 10.1080/13506285.2013.767296

Sheridan, H., & Reichle, E. D. (2015). An analysis of the time course of lexical processing during reading. *Cognitive Science*, *40*(3), 226–243. doi: 10.1111/cogs.12246

Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development*, *74*(1), 27–47. doi: 10.1111/1467-8624.00519

Slattery, T. J., & Rayner, K. (2013). Effects of intraword and interword spacing on eye movements during reading: Exploring the optimal use of space in a line of text. *Attention, Perception, & Psychophysics*, *75*(6), 1275–1292. doi: 10.3758/s13414-013-0463-8

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In M. G. Gaskell & G. T. M. Altmann (Eds.), *The Oxford handbook of psycholinguistics*. Oxford, UK: The Oxford University Press.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of english orthography. *Journal of Experimental Psychology-General*, *124*(2), 107–136. doi: 10.1037/0096-3445.124.2.107

Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, *45*(5), 453–466. doi: 10.1177/0022219411432685

Valle, A., Binder, K. S., Walsh, C. B., Nemier, C., & Bangs, K. E. (2013). Eye movements, prosody, and word frequency among average-and high-skilled second-grade readers. *School Psychology Review*, *42*(2), 171–190.

Van Orden, G. C. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*(4), 488–522. doi: 10.1037/0033-295X.97.4.488

Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics*, *47*(6), 583–600. doi: 10.3758/BF03203111

Wang, Y., Yin, L., & McBride, C. (2015). Unique predictors of early reading and writing: A one-year longitudinal study of Chinese kindergarteners. *Early Childhood Research Quarterly*, *32*, 51–59. doi: 10.1016/j.ecresq.2015.02.004

Warren, T., Reichle, E. D., & Patson, N. D. (2011). Lexical and post-lexical complexity effects on eye movements in reading. *Journal of Eye Movement Research*, *4*(1), 1–10.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*(1), 59–85. doi: 10.1016/0010-0285(70)90005-8

Wolfram, W. (2006). Variation and language: An overview. In E. K. Brown (Ed.), *Encyclopedia of Language and Linguistics*. Oxford: Elsevier.

Xu, Y., Pollatsek, A., & Potter, M. C. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 838–857. doi: 10.1037/0278-7393.25.4.838

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology*, *97*(Pt 2), 259–268. doi: 10.1348/000712605X70066

Yang, J. F., McCandliss, B. D., Shu, H., & Zevin, J. D. (2009). Simulating language-specific and language-general effects in a statistical learning model of Chinese reading. *Journal of Memory and Language*, *61*(2), 238–257. doi: 10.1016/j.jml.2009.05.001

Zhang, J., McBride-Chang, C., Tong, X., Wong, A. M. Y., Shu, H., & Fong, C. Y. C. (2012). Reading with meaning: The contributions of meaning-related variables at the word and subword levels to early Chinese reading comprehension. *Reading and Writing*, *25*(9), 2183–2203. doi: 10.1007/s11145-011-9353-4

Zhang, S., & Perfetti, C. A. (1993). The tongue twister effect in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1082–1093. doi: 10.1037//0278-7393.19.5.1082

Ziegler, J., Perry, C., & Coltheart, M. (2003). Speed of lexical and nonlexical processing in French: The case of the regularity effect. *Psychonomic Bulletin & Review*, *10*(4), 947–953. doi: 10.3758/BF03196556

Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology*, *12*(3), 413–430. doi: 10.1080/09541440050114570

# 13 The Bilingual Lexicon

JUDITH F. KROLL[1] AND FENGYANG MA[2]

[1] University of California, Riverside
[2] University of Cincinnati

## Introduction

In the past two decades there has been increasing interest in research on bilingualism and second language learning. One reason is that cognitive psychologists, linguists, and cognitive neuroscientists have come to see that bilingualism provides a lens for investigating issues of language and mind that are otherwise impenetrable in speakers of one language alone (e.g., Kroll, Dussias, Bogulski, & Valdes Kroff, 2012). Much of the new research on bilingualism has examined the way that words are processed in each language. Although a comprehensive account of language processing requires that we understand much more than words alone, studies of the bilingual lexicon have contributed profound insights into the nature of bilingualism itself, the trajectory of language learning, and the dynamic interactions across the bilingual's two languages.

In this chapter we review the recent evidence on the way that second language (L2) learners and bilinguals understand and speak words in each of the two languages and what that tells us about the dynamics of the bilingual lexicon. We also consider accounts of how new vocabulary is acquired for individuals acquiring new words in an L2 and for bilinguals acquiring new words in a third language (L3). A theme in the recent research on the bilingual lexicon, and indeed on bilingualism more generally, is that both languages are continually active when even one language alone is required. That activity is thought to give rise to cross-language interaction and to competition that needs to be resolved to enable selection of the intended language. Much of the evidence on these issues has been reported in laboratory experiments that use behavioral methods to track language processing. More recently, neuroscience methods, using Event Related Potentials (ERPs) and functional magnetic resonance imaging (fMRI) have been exploited to identify the very earliest stages in comprehension and speech planning and to localize the brain areas that are engaged by these processes and that additionally allow bilinguals to regulate the language not in use.

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

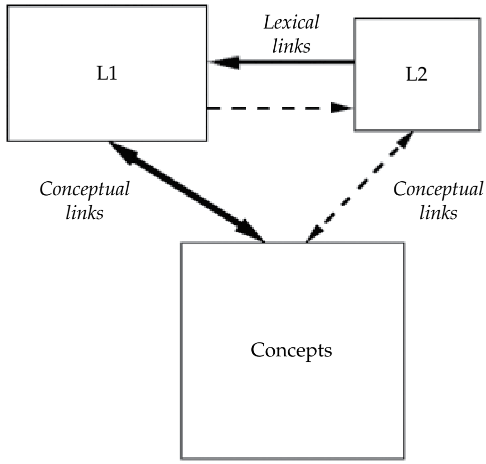
## Parallel activation of the bilingual's two languages

One of the most important discoveries in the past two decades of research on the bilingual lexicon is that information about words in both languages becomes active in parallel when bilinguals read, listen to spoken language, or plan speech in either of their two languages (see Kroll, Gullifer, & Rossi, 2013, for a recent review). The parallel activation of the two languages is counterintuitive, especially when we consider speech planning. One might assume that spoken production is initiated with great intelligence, exploiting the intention to speak a particular language to a specific audience in a designated context. Although the mechanisms of language selection have been debated (e.g., Finkbeiner, Gollan, & Caramazza, 2006; La Heij, 2005; and see Kroll & Gollan, 2014 for a recent review), there is little evidence that bilinguals are able to engage in this sort of smart decision making on the fly (but see Li *et al.*, 2014, and Zhang *et al.*, 2014, for recent findings that suggest otherwise). To the contrary, information in both languages becomes available and interacts. The source of cross-language interaction varies, depending on whether the bilingual is listening to speech, reading, or speaking, and on the structural overlap across the bilingual's two languages. In all cases, there is a requirement to regulate the activation of the language not in use. However, the hypothesized mechanisms of control to enable selection appear to differ across these different lexical tasks. But as we will see, the finding of parallel activation is compelling across a wide range of contexts and for virtually all bilingual language pairings, no matter how different the two languages may be (e.g., Emmorey *et al.*, 2008; Morford *et al.*, 2011; Thierry & Wu, 2007).

## Models of the bilingual lexicon

Models of the bilingual lexicon capture the consequences of different types of cross-language interaction (for a review of the historical development of lexical models, see Kroll & Tokowicz, 2005). Figure 13.1 contrasts the Revised Hierarchical Model (Kroll & Stewart, 1994), or RHM, and the Bilingual Interactive Action Model + (Dijkstra & Van Heuven, 2002), or BIA+. The RHM focuses on the question of how L2 word forms are mapped to meaning during early stages of learning and how that learning history then creates a set of asymmetries between form and meaning for even relatively proficient bilinguals. The BIA+ and the earlier Bilingual Interactive Activation (BIA) model on which it was based (Van Heuven, Dijkstra, & Grainger, 1998) use a localist connectionist architecture to characterize the data-driven activation of lexical codes in both languages when bilinguals recognize visually presented words in one language alone. We briefly consider the evidence for each type of cross-language activation and its implications for these two models.

(a)




(b)

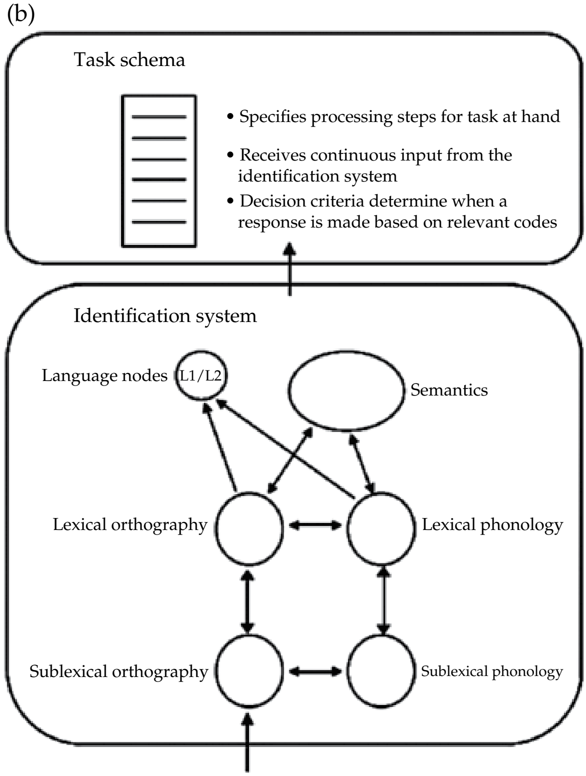



**Figure 13.1** Two models of the bilingual lexicon. The model in Figure 1a is the Revised Hierarchical Model, adapted from Kroll and Stewart (1994). The model in Figure 1b is the Bilingual Interactive Activation + or BIA+ model, adapted from Dijkstra and Van Heuven (2002).

## *The Revised Hierarchical Model*

The RHM focuses on the role of the L1 translation equivalent when bilinguals process words in their L2 and when they translate between their two languages. The model assumes that during initial learning of L2 vocabulary, it is necessary to mediate access to the semantics via the L1 translation equivalent. In this regard, the RHM is a model of transfer, not unlike models of transfer that have been proposed for the acquisition of the L2 grammar (e.g., MacWhinney, 2005; Pienemann, Di Base, Kawaguchi, & Håkansson, 2005). In acquiring any new skill, learners are assumed to transfer existing knowledge, in the case of L2 learning, the knowledge associated with the native language. As learners become more proficient in the L2, the RHM assumes that the reliance on transfer from the L1 diminishes, until L2 speakers are able to function independently in the L2. Whereas beginning learners need to rely on the L1 translation words to access the meaning of L2 words, advanced learners are assumed to be able to access the meaning of L2 words directly without the L1 mediation.

In the time since the RHM was proposed, there has been a great deal of research testing predictions derived from the model and a number of critiques that suggest that some of the assumptions within the model were incorrect (see Brysbaert & Duyck, 2010 and Kroll, Van Hell, Tokowicz, & Green, 2010, for a review of the critical evidence). In what follows, we discuss the major results of studies that have examined the mapping of word forms to meaning with increasing proficiency in the L2 and consider the specific contributions of neuroscience evidence in adjudicating debates on these issues.

Initial support for the RHM came from experiments on translation production. Kroll and Stewart (1994) compared the translation performance for Dutch-English bilingual speakers when they translated from L1 to the L2 (forward translation) and from L2 to the L1 (backward translation). The RHM predicts that only forward translation will necessarily engage semantics; backward translation is hypothesized to rely directly on the lexical level associations from the L2 word to the L1 translation equivalent. To test this prediction, Kroll and Stewart used lists of words to be translated that were either blocked by semantic category or randomly mixed. They found an effect of semantic blocking only in the forward direction of translation, such that translation was slower for blocked than mixed lists. In the backward direction of translation, there was no effect of semantic blocking. A comparison with a word naming control revealed longer latencies to name words in L2 than in L1 but no effect of semantic blocking, suggesting that the differential effect of semantic blocking in forward translation was not due to greater difficulty in speaking the L2.

Subsequent research addressed the question of whether lexical mediation via the L1 translation was observed in both comprehension and production of L2 words and attempted to track changes in the asymmetries hypothesized by the RHM with increasing proficiency in the L2. Because learners have difficulty producing words in the L2, spoken translation is not a good means to assess L2 processing. Instead, De Groot (1992) developed a translation recognition task in which two words are presented, one word in each language, with the task only to judge whether the second word is the correct translation of the first word. Talamas, Kroll, and Dufour (1999) used this task to investigate the role of the lexical form and meaning in native English speakers at two different levels of proficiency in Spanish as the L2. Half of the word pairs were correct translations and the remaining half were not. The condition in which the words were not translations were the critical focus of the study because these incorrectly paired words could be related in lexical form to resemble the translation (e.g., the word *hambre* [hunger] for *hombre* [man] in Spanish), to be related to the meaning of the correct translation (e.g., the word *mujer* [woman] for *hombre* [man]) or completely unrelated. Talamas *et al*. found that the interference imposed by these distractors differed as a function of L2 proficiency. Less-proficient L2 speakers were more sensitive to the lexical form than the meaning distractors and the reverse was true for the more proficient L2 speakers. The pattern of results supported the predictions of the RHM. The initial dependence on the L1 translation equivalent, revealed by the interference generated in the lexical for condition, appeared to give way to increasing semantic processing as learners became more proficient.

Later studies reported a more complex pattern of results (see Kroll *et al.*, 2010, for a review). In some cases, even non-proficient learners appear able to directly access the meaning of L2 words (e.g., Sunderman & Kroll, 2006) and even highly proficient bilingual speakers may sometimes reveal access to the L1 translation equivalent (e.g., Ferré *et al.*, 2006). It is of interest to note that the presence of lexical form interference in translation recognition does not appear to be modulated by the similarity of the bilingual's two languages. It has been reported in bilinguals whose two languages are typologically close, such as Spanish-Catalan (e.g., Guasch *et al.*, 2008) and for those whose two languages are more distant, such as Arabic-English (e.g., Qasem & Foote, 2010).

Thierry and Wu (2007) took a different approach to the question of whether the L1 translation is activated when bilinguals read words in the L2. They used ERPs to investigate performance by proficient Chinese-English bilinguals on a semantic judgment task presented in English. The bilinguals were living in the UK and therefore immersed and highly proficient in English as the primary language. The task was to decide whether two English words were semantically related or not. What the bilinguals did not know was that on some related and some unrelated trials, the Chinese translations of the English words shared characters. Thierry and Wu found that there was a modulation of the N400 component in the ERPs in the presence of shared Chinese characters, suggesting that even when these highly proficient bilinguals were functioning in their L2 alone, they implicitly activated the translation equivalents of the L2 words in their L1 (and see Morford, Wilkinson, Villwock, Piñar, & Kroll, 2011, for related behavioral evidence from deaf bimodal bilinguals reading words in English but activating translations in American Sign Language). Again, each of these studies demonstrates that even for bilinguals whose two languages take very different form, there is activation of the L1 when the L2 is processed. But critically, the bilinguals in these two studies were highly proficient in English as the L2, demonstrating that contrary to the claim of the

RHM, lexically mediated connections from the L2 to the L1 are salient not only during the earliest stages of L2 learning but for all L2 speakers.

The observation that the L1 translation is active for highly proficient bilinguals is problematic not only for the account given by the RHM, but also for its critics (e.g., Brysbaert & Duyck, 2010 for a review), who argue that all L2 speakers, regardless of proficiency, access meaning directly for L2 words without the need to mediate via the L1 translation. Guo, Misra, Tam, and Kroll (2012) used ERPs to test an alternative hypothesis about the role of L1 translation. They hypothesized that for learners, the L1 may indeed function as a mediator to enable access to meaning. In contrast, highly proficient bilinguals may access the L1 translation under some circumstances once the meaning of the L2 word has been understood. Guo *et al.* noted that all of the past studies using translation recognition and judgments of semantic relatedness used relatively long stimulus onset asynchronies (SOAs) between the two words, which may have allowed proficient bilinguals to activate the L1 translation after retrieving meaning of the L2 word. To test this hypothesis, they examined both behavioral data and ERPs in high-proficiency Chinese-English bilinguals performing translation recognition under two different SOA conditions. Highly sensitive to the time-course of language processing, ERPs reveal the early neural processes in word processing, which may not be evident in behavioral data alone (see Van Hell & Kroll, 2012, for a review). The Guo *et al.* behavioral data showed that high-proficiency bilinguals exhibited semantic and translation interference in translation recognition at both long and short SOAs. However, the form of these effects in the ERP data followed a different pattern. Furthermore, there were significant semantic and translation effects in the ERP data at the long SOA, but only a semantic effect at the short SOA. The findings confirmed Guo *et al.*'s (2012) hypothesis, suggesting that given enough time, high-proficiency bilinguals access the L1 translation word after they access the meaning of an L2 word directly.

Other findings with low proficiency learners also seem problematic to the RHM's assumption that beginning learners rely on the L1 to enable semantic access for L2 words (e.g., Brenders, 2012; Brenders, van Hell, & Dijkstra, 2011; Ma, Chen, Guo, & Kroll 2017; Sunderman & Kroll, 2006). In brief, there is overwhelming evidence that learners at relatively early stages of L2 acquisition are sensitive to the meaning of words in the L2 when the task is to comprehend that meaning (and see Dufour & Kroll, 1995, for an early report of learner sensitivity to L2 meaning in a semantic categorization task). Ma *et al.* replicated the Guo *et al.* (2012) translation recognition experiments but with native English speakers at early stages of learning Spanish as the L2. They found a pattern of results that was remarkably similar to the pattern reported by Guo *et al.* for high proficient Chinese-English bilinguals. In both cases, the sensitivity to the translation distractor was modulated by the time course, with larger effects at longer than at shorter SOAs. But low proficiency as well as high proficiency participants were sensitive to the semantic distractors at both SOAs, suggesting that the L1 itself does not mediate access to meaning for the low proficiency learners.

In sum, the current evidence seems to suggest that the L1 translation is active for all L2 speakers under a range of circumstances. At the same time, the recent studies suggest that the L1 has a limited role in accessing the meaning of an L2 word during comprehension, regardless of L2 proficiency and the linguistic distance between the two languages. It remains possible that the lower proficiency learners examined in the previous studies were already past a critical early stage of learning. As we will see in our later review of lexical production, the asymmetries predicted by the RHM are more likely to be seen when learners and bilinguals plan speech in the L2 than when they recognize L2 words.

## *The Bilingual Interactive Activation + Model*

Visual word recognition involves recognizing the orthographic and/or phonological features of words and mapping them to their corresponding semantic representations. The BIA+ Model (and the earlier BIA Model) proposed by Dijkstra and Van Heuven (2002) uses the connectionist architecture of the Interactive Activation Model that was originally proposed by McClelland and Rumelhart (1981) to account for word recognition within the native language. The bilingual version of the model resembles the monolingual model in that there is bottom-up processing of sub-lexical information that creates parallel activation of orthographic and/or phonological information in both of the bilingual's two languages. That process is hypothesized to create competition between alternative candidates. Only late in processing, at the level of a task schema, does the language-specific identification of words in each language become available. On this view, language control only occurs late in the sequence of word recognition, with early processes characterized by cross-language activation that is largely unaffected by context or expectations. The initial evidence for the BIA+ model came from experiments on visual word recognition that demonstrated that regardless of the intention of a bilingual reader to use one language only, cross-language interactions are evident under a wide range of circumstances (see Dijkstra, 2005, for a review of the behavioral evidence, and Van Heuven & Dijkstra, 2010, for a review of the neuroimaging evidence).

The basic evidence for the bottom-up activation of information in both languages comes from experiments that examine interactions when some shared information across languages is present in the written input. The logic of the approach is to ask whether bilinguals perform as monolinguals do when presented with visual input in one language alone. If the two languages were functionally separate, then bilinguals should not be affected by cross-language similarity or difference. In what is now an extensive literature, there is overwhelming evidence that bilinguals cannot ignore the influence of the language not in use, even when it might benefit performance to do so.

Many of the experiments investigating this issue have used words such as cognates or homographs that share some critical features across the two languages. Cognates are translations that have similar lexical form and the same meaning in two languages (e.g., the word *hotel* in Dutch is virtually identical to the word *hotel* in English). Like cognates, interlingual homographs share similar lexical form in two languages but conflict in meaning (e.g., the word *angel* in Dutch looks like

English word *angel* but means sting). That conflict is why they are often called "false friends" because learners, in particular, may be fooled by the lexical similarity. A number of studies (e.g., Brenders *et al.*, 2011; De Groot, Delmar, & Lupker, 2000; Dijkstra *et al.*, 1998; Dijkstra *et al.*, 2000) have shown that recognition of homographs is slower than that of matched controls existing exclusively in one language. Homograph interference has been attributed to the competition caused by the two readings with conflicting meaning across languages. It contrast, cognates generally produce facilitation, which has been understood as a reflection of convergence or resonance across lexical codes (e.g., Brenders *et al.*, 2011; Dijkstra *et al.*, 2010; Moon & Jiang, 2012; Van Hell & Dijkstra, 2002).

The competition that arises when similar lexical forms are activated appears to be modulated by the degree of orthographic and phonological similarity (e.g., Dijkstra *et al.*, 1998; Schwartz, Kroll, & Diaz, 2007), suggesting an early locus of cross-language interaction in reading words in either of the bilingual's two languages. Schwartz *et al.* asked relatively proficient English speakers of Spanish to name words in each language in separate blocks. The words included cognates and matched non-cognate controls. The cognate example given above (i.e., the word *hotel* in Dutch and English) has identical spelling in both languages but cognates can also be similar without being identical (e.g., the word *spinazie* in Dutch is *spinach* in English). Likewise, even for identical cognates, the phonology is rarely identical, pronunciations in each language that can be quite distinct (e.g., the cognate *base* in Spanish and English). Schwartz *et al.* demonstrated that the time to begin to name a cognate in either language was slower when the phonology of the cognate's mate in the other language was judged to be dissimilar. Because the other language was not present nor required (words were named in one language alone within a block of naming trials), and because these interactions occurred even when bilinguals named words in their L1, the results were taken to support the prediction of the BIA+ model that bottom-up activation of lexical and sub-lexical information creates cross-language interactions that regardless of information that might otherwise signal to a reader that he or she is reading in one language alone.

## Language nonselectivity in reading in context

The research that we have reviewed has focused on reading individual words in isolation in one of the bilingual's two languages. In real language use, bilinguals rarely read isolated words; instead, they are more likely to recognize words presented in sentence contexts. In theory, the language of a sentence context should provide a cue for bilinguals to restrict activation of the lexical alternatives to the language in which they are reading. If the context in which a person is reading appears reliably in only one language, it should be easy to identify words in their language-specific sense. Despite this obvious cue to the language in use, research on lexical activation in sentence context suggests that bilinguals activate both languages in parallel, as if the context were irrelevant. Many studies have now demonstrated that bilinguals continue to activate the language not in use when reading words in sentence context (e.g., Baten, Hofman, & Loeys, 2011; Duyck *et al.*, 2007; Gullifer *et al.*, 2013; Jouravlev & Jared, 2014; Libben & Titone, 2009; Schwartz & Kroll, 2006; Van Assche *et al.*, 2009; Van Hell & De Groot, 2008; and see Kroll & Dussias, 2013; Kroll, Gullifer, & Rossi, 2013; and Schwartz & Van Hell, 2012 for recent reviews). The basic finding in these studies is that the same facilitation or interference for language ambiguous words reported in isolated word recognition studies is also found when those words are embedded in sentence context. The only exception to this general observation comes from studies that have shown semantic constraints in sentence context may enable language-specific access (e.g., Schwartz & Kroll, 2006; Titone *et al.*, 2011; but see Van Assche *et al.*, 2011).

The studies showing parallel activation of words in sentence context are counterintuitive because it would seem that the language of the sentence itself should suffice to cue the bilingual reader that the upcoming words in the sentence are in that language. Studies of sentence processing in both the L1 and the L2 demonstrate that bilingual readers are sensitive to a range of syntactic and semantic constraints (e.g., Dussias & Cramer Scaltz, 2008; and see Kroll & Dussias, 2013, for a review). The apparent inability of language specific cues in sentence context to guide lexical access is therefore not an indication of insensitivity more generally. Furthermore, not only is there cross-language activation when bilinguals read in the L2, a situation in which we might expect that the relatively weaker of the two languages would be more likely to be influenced by the stronger of the two languages, but also in the L1. For relatively proficient bilinguals, there is evidence of lexical activation of the L2 that affects reading in the L1, even in sentence context in the native language (e.g., Gullifer *et al.*, 2013; Titone *et al.*, 2011; Van Assche *et al.*, 2009). There is also evidence that the parallel activation of the two languages in sentence context is not constrained by differences in the form of written script; similar effects of cross-language activation have been reported even when bilinguals are reading a language that is unambiguously one of their two languages, as in the case of Russian and English (e.g., Jouravlev & Jared, 2014).

Gullifer *et al.* (2013) took a different approach to examine the apparent lack of language-specific constraints in lexical access in sentence context. They compared the magnitude of cognate facilitation when words were embedded in sentences in either English or Spanish. A given sentence always appeared in one language alone but in an inter-sentential code switching condition, the language of the sentence alternated from Spanish to English. For a group of highly proficient Spanish-English bilinguals, the alternation of language from one sentence to the other had no effect on the presence of cognate facilitation for target words embedded within those sentences. There was a significant and similar cognate effect regardless of readers' ability to predict whether the sentence was in English or Spanish. Consistent with the predictions of the BIA+ model, these results suggest that high-level expectations about the language to be used does not play a critical role in guiding lexical access.

Only a few studies have investigated the issue of whether language-specific information provided by grammatical differences across the bilingual's two languages might effectively constrain lexical access to the target language. Baten *et al.* (2011) asked Dutch-English bilinguals to perform a lexical decision task to target words embedded in English (L2) sentences. Homographs with the same word class across two languages (e.g., *lever* is a noun in English and also in Dutch, but meaning liver) exhibited a facilitation effect compared to their matched controls. In contrast, homographs with different grammatical classes (e.g., *big* is an adjective in English, but a noun in Dutch, meaning piglet) showed no such effect. These findings suggest that grammatical class information plays an essential role in cross-language activation during bilingual L2 word comprehension. Notably, a facilitation effect was reported for the homographs with grammatical class overlap in this study, which differs from the interference effect observed in a number of other studies (e.g., Libben & Titone, 2009; Titone *et al.*, 2011). Baten and colleagues attributed the unanticipated homograph facilitation effect to the convergence of orthography and word class.

In sum, research on bilingual visual word recognition in sentence context provides compelling evidence for early cross-language activation that is not easily overcome. The pattern observed for both isolated presentations and for lexical access in context largely supports the predictions of the BIA+ model (Dijkstra & Van Heuven, 2002), which assumes a minimal role of top-down influences. A goal for ongoing research is to further identify the constraints that appear to characterize lexical access and to develop methods that may provide more sensitive indices of these processes. If bottom-up processes driven by the properties of words that are read or spoken create activation of lexical alternatives in each of the bilingual's two languages, then a mechanism of inhibitory control must be invoked to enable bilinguals to regulate the influence of the unintended lexical form. A recent report by Pivneva, Mercier, and Titone (2014) suggests one approach to this issue. Pivneva *et al.* identified individual differences within a large group of French-English bilinguals to differentiate language proficiency and inhibitory control ability. Using temporally sensitive measures of eye tracking, they found effects of both proficiency and inhibitory control but for different aspects of lexical processing. The magnitude of cognate facilitation in sentence context was reduced for bilinguals with higher L2 proficiency but independent of inhibitory control. In contrast, there was reduced interference for interlingual homographs for bilinguals with enhanced inhibitory control ability but the effect was independent of proficiency. The distinct pattern that Pivneva *et al.* described suggests that a more nuanced model of bilingual lexical processing is required to include an account of the time course of resolving activation and competition across different lexical codes and also the influence of higher level factors. The tendency in past research to view these effects categorically and to ignore individual differences may have failed to reveal the subtle dynamics of the bilingual lexicon. While the body of research overall supports the claims of the BIA+ model, results like those of Pivneva *et al.* demonstrate that it may be premature to conclude that there are no top-down influences on the earliest stages of visual word recognition.

## Neurocognitive evidence for language nonselectivity in lexical access

In the recent literature, there has been an upsurge of research that uses neuroscience methods, such as ERPs or fMRI, to investigate more sensitively the time course and localization of cross-language activation (e.g., see Van Heuven & Dijkstra, 2010). In the previous section on the RHM, we reviewed a number of studies that used ERPs to test predictions of the model. The temporal sensitivity of ERP provides a means to isolate the locus of particular cross-language interactions during lexical processing. The spatial resolution of fMRI provides a complementary source of information to determine how patterns of brain activation reflect the emergence and resolution of cross-language competition. Here we review some of the recent findings that illustrate the power of these methods to generate new information about the bilingual lexicon. As we will see, some of the results using these neuroscience methods converge closely with the behavioral evidence that has been reported. In other cases, the neuroscience evidence diverges from the behavioral findings and/or reveals aspects of lexical processing that were otherwise unavailable using behavioral methods alone.

Midgley *et al.* (2008) used ERPs to examine the influence of cross-language lexical neighbors in French-English bilinguals. Lexical neighbors are words that share all but a single letter, either within or between languages. One of the early studies that provided behavioral evidence for the BIA model (Van Heuven *et al.*, 1998) used lexical neighbors as a vehicle to examine cross-language interactions between Dutch and English. Van Heuven *et al.* found that lexical decision performance in each language was influenced not only by the presence of lexical neighbors within that language but also by the number of neighbors in the language not in use. Midgley *et al.* found that words with many cross-language neighbors elicited larger amplitudes in the N400 component (negative-going waveforms occurring around 300-500 ms and peaked at around 400 ms after the stimulus onset) relative to words with few cross-language neighbors. Since the N400 component is primarily related to semantic processing, the larger N400 amplitude was taken to reflect greater efforts to resolve competition from cross-language orthographic neighbors, and ultimately to access the meaning of the target word.

Other ERP studies (e.g., Comesaña *et al.*, 2012; Midgley *et al.*, 2011; Peeters *et al.*, 2013) have reported a cognate effect in ERP measures that mirrors the results reported in behavioral experiments. Critical to the current discussion, Midgley *et al.* (2011) and Peeters *et al.* (2013) found an N400 attenuation in cognates compared with matched non-cognate controls, suggesting greater ease of lexical-semantic processing for cognates than non-cognates. Cognates are special in that there is a high level of correspondence between form and meaning across the bilingual's two languages.

Kerkhofs *et al.* (2006) tested Dutch-English bilinguals in an L2 (English) lexical decision task. Both behavioral and ERP data showed that these bilinguals were sensitive to the frequency of interlingual homographs in the non-target L1

(Dutch). Specifically, homographs with a higher L1 frequency yielded longer reaction latencies and more negative-going waveforms in the N400 component, compared to homographs with a low L1 frequency. The interpretation for the cross-language effect is that an interlingual homograph with a higher L1 frequency resulted in higher availability of the L1 reading, which in turn created greater interference for the semantic integration of its L2 reading (and see Dijkstra, Van Jaarsveld, & Ten Brinke, 1998 on the effects of relative frequency of the L1 and L2 interpretation of interlingual homographs). Critically, what is notable across all of these ERP studies of bilingual word recognition, is that cross-language interactions are observed relatively early in processing, consistent with the hypothesis that they reflect the result of data-driven processes rather than later top-down processes. It will remain to be seen whether the sort of dissociation reported by Pivneva *et al.* (2014) for homographs and cognates in sentence context, can be detected in the ERP record.

Neuroimaging data also offers support for an integrated bilingual lexicon of the sort that is assumed by the BIA+ model (see Van Heuven & Dijkstra, 2010). Many fMRI studies suggest that the bilingual's two languages are supported by the same neural tissue but that when differences arise, they are more likely to be attributed to the need engage control mechanisms to regulate cross-language competition (e.g., Abutalebi, Cappa, & Perani, 2005; Abutalebi & Green, 2007). Several studies have adopted the functional magnetic resonance adaption (fMRA) technique (e.g., Grill-Spector, Henson, & Martin, 2006; Grill-Spector & Malach, 2001) to examine the neural overlap and dissociation of a bilingual's two languages. In this technique, adaption refers to a reduction of fMRI signals elicited by repeated stimuli compared to unrelated stimuli, and it has been attributed to the pre-activation of the representation of the repeated stimulus in voxels within a certain region (e.g., Grill-Spector *et al.*, 2006). In the fMRA's application to examine bilingual semantic representation, fMRI signals elicited by translations versus unrelated words from different languages have been compared. If translations or semantically related words in a bilingual's two languages share the same representational system, there should be reduced activation relative to dissimilar stimuli, producing cross-language adaptation effect. A few fMRA studies have reported common cortical substrates in L1 and L2 visual word recognition, including the left prefrontal and lateral temporal regions (Chee, Soon, & Lee, 2003), the bilateral left superior temporal gyrus and left inferior frontal region (Klein *et al.*, 2006), and the left anterior temporal cortex (Crinion *et al.*, 2006). For example, Crinion *et al.* (2006) tested late German-English bilinguals and late Japanese-English bilinguals in a primed semantic decision task. In this task, target words are preceded with either semantic related words or non-related words in either language. Results showed that both within-language and cross-language semantically related primes reduced activation of the targets in the left anterior temporal cortex. Based on this finding, they concluded that word meanings in the two languages converge on the same neuronal populations within this region. In summary, these fMRA studies indicate shared semantic representations of the bilingual's two languages, as proposed in both RHM and BIA+ models.

It is worth noting that L2 proficiency modulates the activation patterns of the bilingual's two languages. Previous research with balanced bilinguals (with comparable proficiency in their two languages) has reported similar recruitment or intensity of activation in overlapping regions during comprehension of words in two languages (e.g., Illes *et al.*, 1999). For unbalanced bilinguals, for whom one language is dominant, the past literature suggests largely overlapping activation patterns and greater activation generated by the weaker language. Again, the overlap of activated regions across language has been taken to suggest shared semantic representations for even less proficient bilinguals. At the same time, some previous investigations have shown greater activation in certain brain regions during word comprehension in the weaker language, including the inferior frontal gyrus (Chee, Hon, Lee, & Soon, 2001; Marian *et al.*, 2007), the left middle frontal gyrus and the left parietal region (Chee *et al.*, 2001), the left frontal cortex and anterior cingulate gyrus (Chee *et al.*, 2000; Sebastian, Kiran, & Sandberg, 2012), the left superior temporal gyrus (Meschyan & Hernandez, 2006; Klein *et al.*, 2006), putamen, insula, and SMA/cingulate in the right hemisphere (Meschyan & Hernandez, 2006), and the right middle frontal gyrus (Park *et al.*, 2012) and the right superior temporal gyrus (Klein *et al.*, 2006). For example, Meschyan and Hernandez (2006) used a silent word reading task to test Spanish-English bilinguals whose L2, English, had become the dominant language. Results showed that word recognition in the weaker language, Spanish, generated greater activity in left superior temporal gyrus (STG), supplementary motor area (SMA), the putamen, and the insular compared to word reading in the stronger language English. The greater activity in the weaker language indicates that more cortical resources are needed to meet the extra processing demands. It still remains to be examined whether the more extensive activation reflects more extensive representation patterns or a stronger cognitive control mechanism to inhibit the stronger language during word comprehension in the less dominant language.

Recent fMRI evidence also provides important information about the neural substrates underlying cross-language competition during bilingual word recognition. Van Heuven, Schriefers, Dijkstra, and Hagoort (2008) used fMRI to examine the homograph effect in bilingual word recognition. Dutch-English bilinguals performed a lexical decision task in English the L2, or a generalized lexical decision task in which they indicated whether the string of letters formed a word in either Dutch or English. Behavioral results showed the standard homograph interference in the English task, but not in the generalized task. More critically, the imaging data revealed differential activation patterns for the homographs and their matched controls at the left prefrontal cortex, a region associated with phonological and semantic processing. Based on these findings, Van Heuven *et al.* (2008) concluded that homographs' two readings were available during bilingual word reading, which caused stimulus-based cross-language conflict in the left prefrontal cortex. Again, the evidence based on neuroscience methods supports the conclusions of behavioral studies in that there appears to be non-selective access to an integrated lexicon, in line with the predictions of the BIA+ model.

## Lexical access in bilingual speech planning

The studies of lexical comprehension that we have reviewed suggest that there is persistent activation of the language not in use even under circumstances when the context is sufficiently rich and predictive to provide a basis on which the target language might be selected. Recent studies of lexical production (e.g., Costa, 2005; Hanulovà, Davidson, & Indefrey, 2011; Kroll, Bobb, & Wodniecka, 2006; and see Kroll & Gollan, 2014, for a recent review of the production research) converge with the conclusions of the comprehension studies. Finding language nonselectivity in spoken production is particularly counterintuitive because unlike word recognition, production is an inherently top-down process, with speech planning initiated by a thought that guides lexical selection (Levelt, 1989). In theory, bilingual speakers should be in control of the language they produce and should be able to take into account the context in which each language is spoken. Quite surprisingly, there is little evidence to suggest that bilinguals are able to exploit the contextual cues that might enable this sort of top-down selection process. At the same time, the parallel activation of the two languages does not appear to induce a high rate of errors in spoken production (e.g., Gollan, Sandoval, & Salmon, 2011). Together, these two observations suggest that bilinguals develop a mechanism of cognitive control that enables them to regulate the relative activation of the two languages to select the intended language to be spoken. Much of the research on bilingual lexical production has therefore focused on the nature of the selection mechanism and its cognitive consequences. Here we consider briefly the findings on each of these topics.

Like the research on bilingual word recognition, the approach in studies of lexical production has been to exploit cross-language ambiguities to demonstrate the activity of the language not in use. To illustrate, bilinguals are faster to speak the name of a pictured object when the name is a cognate in both languages (e.g., Costa, Caramazza, & Sebastián-Gallés, 2000) and to translate words from one language to the other when the translations are cognates (e.g., Kroll & Stewart, 1994). Recent ERP studies (e.g., Strijkers, Costa, & Thierry, 2010) have shown that the cognate effect in picture naming can be detected as early as 200 ms in speech planning. Because written words are not present in this task, the cognate effect must necessarily arise from the shared phonology that is activated as the picture's name is planned. The phonological basis of the cognate effect in production has been revealed in studies that compare bilinguals whose two languages use the same or a different script (e.g., Hoshino & Kroll, 2008). When the written form of the bilingual's two languages differs, it is only the phonology that can be the basis of the observed facilitation.

## Mechanisms of language selection in spoken production

If both languages become active during speech planning, then the activation of the language not to be spoken must be reduced to enable the intended language to be produced. Two families of models have been proposed to accomplish lexical selection (and see Kroll & Gollan, 2014, for a more detailed review). On one account, the activation of the non-target language is reduced by virtue of the intention to use one language and not the other (e.g., Costa, Miozzo, & Caramazza, 1999; La Heij, 2005). In a sense, the smartest solution of all would be to attend only to one language and to ignore the language not in use. The evidence mentioned above on cognate effects in speech production suggests that it is virtually impossible to switch off activation of the language not in use.[1] The model proposed by Costa *et al.* assumes what we have called a "mental firewall" so that both languages are activated but only alternatives in one of the two languages become candidates for lexical selection. This sort of selective account requires two strong assumptions. One is that activated alternatives in the language not in use do not compete for selection with those activated in the target language. The other is that bilinguals are fully able to exploit the presence of cues to the target language. That is, there must be a way to separate the languages from the start so attention can be properly allocated to the intended language. Although findings like cognate facilitation in picture naming would seem to suggest that the two languages cannot be easily separated, a number of different studies have attempted to demonstrate that there is not direct cross-language competition. To illustrate, Costa *et al.* used a picture-word interference paradigm in which bilinguals named a picture in the presence of a visually presented distractor word that was to be ignored. Critically, there were effects of distractors drawn from the non-target language, a result in and of itself that might be taken to suggest that the influence of the language not in use cannot be easily limited. In one of the distractor conditions, the word to be ignored was the name of the object in the language not to be spoken, that is, the translation equivalent. Contrary to the prediction that the translation in the language to be ignored should be the most troublesome distractor of all, Costa *et al.* reported facilitation and argued that the result was incompatible with an interpretation of active competition (and see Gollan & Acenas, 2004, for a similar finding about bilingual tip of the tongue states). Each of these findings, however, can be interpreted in a number of ways that may not require the assumed language selectivity (e.g., Hermans, 2004).

The second assumption according to language selective models is that bilinguals are able to exploit cues that signal the appropriateness of one language relative to the other. Without the ability to modulate attention to the target language, it would seem virtually impossible to achieve selectivity. Yet, the evidence to date provides only minimal support for the idea that even proficient bilinguals are able to seize the information in contextual cues to guide language selection. The few recent studies that claim to show sensitivity, for example, to the face of the person to whom the bilingual is speaking (see note 1), have not shown that bias from the context comes early in speech planning.

The alternative view is that selectivity is not possible early enough in planning to guide selection to only one of the two languages. Instead, lexical alternatives in both languages are hypothesized to become active and to compete for selection (e.g., Green, 1999; Kroll, Bobb, Misra, & Guo, 2008). A mechanism must then effectively reduce the activation of the language not in use to enable production in the target language. Recent studies of bilingual speech planning suggest that the language not in use is inhibited to achieve this goal. The approach to this issue has primarily involved studies of lexical switching, requiring bilinguals to either switch from one language to the other on cue or to switch across blocks of naming trials after speaking one language for a period of time.

In an early study on lexical switching, Meuter and Allport (1999) reported asymmetric switch costs for the L1 and L2, with greater costs following switching into the L1 than into the L2. The pattern of switch costs has been interpreted as support for the Inhibitory Control Model (Green, 1999). The assumption is that the L1 is active during speech planning in the L2 and must eventually be inhibited to enable fluent L2 speech. When the L1 must then be spoken just after L2, it is hypothesized to be in a state of reduced activation, requiring additional cognitive resources and producing greater switch costs. Although there has been debate about the interpretation of the switch cost asymmetry (e.g., Costa & Santesteban, 2004; and see Bobb & Wodniecka, 2013, for a recent review of the research on lexical switching), there is growing evidence that multiple mechanisms of inhibitory control are engaged when bilinguals plan to speak words in one language alone.

Misra, Guo, Bobb, and Kroll (2012) asked whether similar switch costs would be observed when bilinguals switch across blocks of naming trials, a context that resembles actual switching more than the artificial demands of trial to trial switching. They performed a simple experiment in which relatively proficient Chinese-English bilinguals were asked to name the same set of pictures, once in Chinese and once in English. The order of the two languages was counterbalanced across speakers and ERPs were recorded. Misra *et al.* predicted that the repetition of pictures from one language to the other should produce facilitation. For naming in English, the L2, that is exactly the pattern that was observed. There was reduced negativity in the ERP record when L2 naming following L1 naming of the identical set of pictures. The surprising result was that for naming in Chinese, the L1, the opposite pattern held. There was greater negativity in the ERPs when L1 followed L2. Because there had to be facilitation as well with the same pictures and concepts repeated, the result suggests that not only is there inhibition of L1 following spoken production in the L2, but that the observed inhibition was likely to underestimate its magnitude. Critically, the inhibitory pattern for L1, when it followed L2, persisted for the duration of the experiment, suggesting a temporally global mechanism of inhibition. Because the same pictures were repeated across languages, it was impossible to draw any conclusions about the scope of inhibition.

In an experiment similar to Misra *et al.* (2012) but using fMRI, Guo, Liu, Misra, and Kroll (2011) asked whether differential brain activation could be seen in blocked picture naming as a function of the order in which the two languages were produced relative to a condition in which the language of naming was mixed. They indeed found two different patterns of brain activation. The blocking effect was hypothesized to reflect global inhibition and activated the dorsal left frontal gyrus and the parietal cortex. In contrast, a comparison of blocked versus mixed picture naming, hypothesized to reflect local inhibition, revealed activation of the dorsal anterior cingular cortex (ACC) and the supplementary motor area (SMA).

Although additional research will have to investigate the distinctions between local and global inhibition, the critical result in the Guo *et al.* study is that at least two different patterns of inhibition were documented, suggesting that there is not a single overarching mechanism of control in bilingual speech planning. Other imaging studies (Abutalebi *et al.*, 2008; and see Abutalebi and Green, 2007 for a model of brain areas activated) converge with the conclusion that these brain areas are differentially engaged during bilingual speech planning. Abutalebi *et al.* demonstrated that selecting a word within a single language does not have the same consequences as selecting a word from one of the bilingual's two languages. What remains to be seen is how these inhibitory mechanisms may be activated under conditions that differ in their demands on bilinguals to regulate the use of each language.

## Learning new words

Monolinguals and bilinguals alike acquire new vocabulary over the course of their lives. For monolinguals, the process of learning new words in their L1, for example, when acquiring expertise in a new domain, or studying for an exam like the SATs, is a bit like acquiring new words in an L2. For bilinguals, that process occurs in both languages and of course learning yet another language becomes an L3. There are two different approaches that have been taken to examine how new lexical knowledge is acquired. One area of research uses what are essentially paired-associate training paradigms to introduce new vocabulary. In some studies the new words are designed to have particular properties to ask how different lexical features of the to-be-learned words affect the learning process. For example, De Groot and Keijzer (2000) demonstrated that new words that are cognates with native language translations or concrete rather than abstract, are easier to learn and better remembered. In other studies, it is the conditions of learning rather than the properties of the words themselves that are the focus. For example, Schneider, Healy, and Bourne (2002) taught French words to learners who knew no French. They manipulated the conditions of learning with respect to the order in which the learners were cued with the French word (L2) or the English word (L1). At test, learners were required to type the word in response to the cue. Supporting the findings of Kroll and Stewart (1994) on translation with actual bilinguals, producing the L2 in response to the L1 was more difficult than the reverse. Critically, the more difficult learning condition produced better later test performance, in line with studies of learning and memory that suggest that conditions of study that induce desirable difficulties by making encoding more effortful, that encourage elaboration, and that allow feedback from errors, improve later memory (e.g., Bjork, 1994; Potts & Shanks, 2014).

The other approach to vocabulary learning has focused on the question of whether learners who are already bilingual may be advantaged in acquiring new vocabulary in an L3 relative to monolingual learners acquiring the same new vocabulary in an L2. A number of past studies suggest that bilinguals are better word learners than monolinguals (e.g., Kaushanskaya & Marian, 2009a, 2009b; Van Hell & Mahn, 1997). A question that this observation raises is whether a bilingual advantage in word learning might reflect the same advantages that bilingualism has been hypothesized to confer to executive function (e.g., Bialystok, Craik, Green, & Gollan, 2009). It is possible that bilinguals are more efficient learners because of their general language learning experience or because they have greater expertise in those aspects of executive function that are engaged by multiple language use. Alternatively, the bilingual benefit in word learning may be due to specific aspects of bilingual learning experience. Bogulski, Bice, and Kroll (under review) asked whether the bilingual benefits to word learning resulted from the general cognitive consequences of bilingualism or from bilinguals' learning histories. They compared bilinguals learning a new L3 vocabulary via their L1 or their L2. All of the past studies on the bilingual advantage in word learning trained the new vocabulary via the L1. The research reviewed earlier on spoken production suggests that bilinguals inhibit the L1 to enable them to produce words in the L2. Bilinguals should therefore have a great deal of experience inhibiting L1 but not as much experience inhibiting L2. If that inhibitory history is critical to the word learning advantage, then only bilinguals learning new L3 words via the L1 would be expected to reveal the effect and that is precisely what Bogulski *et al.* reported. Only bilinguals learning the new vocabulary via their L1 revealed an advantage relative to monolingual learners. But Bogulski and Kroll also found that when comparing performance during initial training on the new words, it was only the bilinguals learning via the L1 who showed evidence at study for self-regulated learning that induced a cost but that translated into a benefit at later test (e.g., Bjork, Dunlosky, & Kornell, 2013). The pattern of results suggests that the word learning advantage for bilinguals is specific to the bilingual's language experience and the cognitive consequences of that experience and not a more general manifestation of executive function benefits.

## Conclusions

Although language is not by words alone, the research reviewed in this chapter on the bilingual lexicon shows that lexical processes can inform models of comprehension, production, and learning in ways that reveal the interface between language and cognition and the neural networks that support it. If there is a summary message, it is that words in the bilingual's two languages are continuously active and require regulation to enable control of each language. Those control mechanisms change with increasing proficiency in the two languages but even highly proficient bilinguals are unable to represent and access words in one language without influence from the other language. The goal of ongoing research on the lexicon is to better understand the dynamics and consequences of cross-language interaction and the implications for language use in contexts in which the grammar and phonology are also engaged.

## Author Notes

The writing of this chapter was supported in part by NIH Grant HD082796, NSF Grants BCS-1535124 and OISE-0968369, and a Guggenheim Fellowship to J.F. Kroll. Correspondence concerning this chapter should be addressed to Judith F. Kroll, Department of Psychology, University of California, Riverside, Riverside, CA 92521, USA. Electronic mail may be sent to judith.kroll@ucr.edu.

## NOTE

1 A number of recent studies have suggested bilinguals can exploit the cues present in the discourse context to selectively attend to the target language, such as the face of the person to whom the bilingual is speaking or culture-specific context (e.g., Jared, Pei Yun Poh, & Paivio, 2013; Zhang, Morris, Cheng, & Yap, 2013). Although these factors appear to have an effect, the locus in language processing at which they influence language selectivity is not yet clear.

## REFERENCES

Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., & Khateb, A. (2008). Language control and lexical competition in bilinguals: An event-related fMRI study. *Cerebral Cortex*, 18, 1496–1505. doi: 10.1093/cercor/bhm182

Abutalebi, J., Cappa, S. F., & Perani, D. (2005). What can functional neuroimaging tell us about the bilingual brain? In J. F. Kroll & A. M. B. De Groot (Eds.). Lexical access in bilingual production. *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 497–515). New York: Oxford University Press.

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20, 242–275. doi: 10.1016/j.jneuroling.2006.10.003

Baten, K., Hofman, F., & Loeys, T. (2010). Cross-linguistic activation in bilingual sentence processing: The role of word class meaning. *Bilingualism: Language and Cognition*, 14, 1–9. doi: 10.1017/S1366728910000246

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, 10, 89–129. doi: 10.1177/1529100610387084

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp.185–205). Cambridge, MA: MIT Press.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444. doi: 10.1146/annurev-psych-113011-143823

Bobb, S. C., & Wodniecka, Z. (2013). Language switching in picture naming: what asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology*, 25, 568–585. doi: 10.1080/20445911.2013.792822

Bogulski, C. A., Bice, K., & Kroll, J. F. (under review). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language.

Brenders, P. (2012). Cross-language interactions in beginning second language learners (Unpublished doctoral dissertation). Nijmegen, Netherlands: Behavioural Science Institute, Donders Institute for Brain, Cognition, Behaviour, Donders centre for cognition, Radboud University.

Brenders, P., Van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology*, 109, 383–396. doi: 10.1016/j.jecp.2011.03.012

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition*, 13, 359–371. doi: 10.1017/S1366728909990344

Chee, M. W., Hon, N., Lee, H. L., & Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. Blood oxygen level dependent. *Neuroimage*, 13(6), 1155–1163. doi: 10.1006/nimg.2001.0781

Chee, M. W., Soon, C. S., & Lee, H. L. (2003). Common and segregated neuronal networks for different languages revealed using functional magnetic resonance adaptation. *Journal of Cognitive Neuroscience*, 15, 85–97. doi: 10.1162/089892903321107846

Chee, M. W., Weekes, B., Lee, K. M., Soon, C. S., Schreiber, A., Hoon, J. J., & Chee, M. (2000). Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI. *Neuroimage*, 12(4), 392–403. doi: 10.1006/nimg.2000.0631

Comesaña, M., Sánchez-Casas, R., Soares, A. P., Pinheiro, A. P., Rauber, A., Frade, S., & Fraga, I. (2012). The interplay of phonology and orthography in visual cognate word recognition: An ERP study. *Neuroscience Letters*, 529, 75–79. doi: 10.1016/j.neulet.2012.09.010

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 308–325). New York, NY: Oxford University Press.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1283–1296. doi: 10.1037/0278-7393.26.5.1283

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365–397. doi: 10.1006/jmla.1999.2651

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50, 491–511. doi: 10.1016/j.jml.2004.02.002

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., … & Price, C. J. (2006). Language control in the bilingual brain. *Science*, 312, 1537–1540. doi: 10.1126/science.1127761

De Groot, A.M.B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1001–1018. doi: 10.1037/0278-7393.18.5.1001

De Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *Quarterly Journal of Experimental Psychology*, 53A, 397–428. doi: 10.1080/713755891

De Groot, A., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, 50, 1–56. doi: 10.1111/0023-8333.00110

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll, & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). New York, NY: Oxford University Press.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, 62, 284–301. doi: 10.1016/j.jml.2009.12.003

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On Being Blinded by Your Other Language: Effects of Task Demands on Interlingual Homograph Recognition. *Journal of Memory and Language*, 42, 445–464. doi: 10.1006/jmla.1999.2697

Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 189–225). Mahwah NJ: Lawrence Erlbaum Associates.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5, 175–197. doi: 10.1017/S1366728902003012

Dijkstra, A., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, 1, 51–66. doi: 10.1017/S1366728998000121

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, 23, 166–180. doi: 10.3758/BF03197219

Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta psychologica*, 128, 501–513. doi: 10.1016/j.actpsy.2007.09.004

Duyck, W., Assche, E. V., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 663–679. doi: 10.1037/0278-7393.33.4.663

Emmorey, K., Borinstein, H. B., Thompson, R. L., & Gollan, T. H. (2008). Bimodal Bilingualism. *Bilingualism: Language and Cognition*, 11, 43–61. doi: 10.1017/S1366728907003203

Ferré, P., Sánchez-Casas, R., & Guasch, M. (2006). Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and nonproficient Spanish-Catalan bilinguals. *Language Learning*, 56, 571–608. doi: 10.1111/j.1467-9922.2006.00389.x

Finkbeiner, M., Gollan, T., & Caramazza, A. (2006). Bilingual lexical access: What's the (hard) problem? *Bilingualism: Language and Cognition*, 9, 153–166. doi: 10.1017/S1366728906002501

Gollan, T. H., & Acenas, L.-A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 246–269. doi: 10.1037/0278-7393.30.1.246

Gollan, T.H., Sandoval, T., & Salmon, D.P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science*, 22, 1155–1164. doi: 10.1177/0956797611417002

Green, D. (1998). Mental control of the bilingual lexico-semantic system.

*Bilingualism: Language and Cognition* 1, 67–81. doi: 10.1017/S1366728998000133
Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in cognitive sciences*, 10, 14–23. doi: 10.1016/j.tics.2005.11.006
Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta psychologica*, 107, 293–321. doi: 10.1016/S0001-6918(01)00019-1
Gullifer, J. W., Kroll, J. F., & Dussias, P. (2013). When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology*, 4, 278. doi: 10.3389/fpsyg.2013.00278
Guasch, M., Sánchez-Casas, R., Ferré, P., & García-Albea, J. E. (2008). Translation performance of beginning, intermediate and proficient Spanish-Catalan bilinguals: Effects of form and semantic relations. *The Mental Lexicon*, 3, 289–308. doi: 10.3389/fpsyg.2013.00278
Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *NeuroImage*, 56, 2300–2309. doi: 10.1016/j.neuroimage.2011.03.049
Guo, T., Misra, M., Tam, J. W., & Kroll, J. F. (2012). On the time course of accessing meaning in a second language: An electrophysiological investigation of translation recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1165–1186. doi: 10.1037/a0028076
Hanulovà, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes*, 26, 902–934. doi: 10.1080/01690965.2010.509946
Hermans, D. (2004). Between-language identity effects in picture-word interference tasks: A challenge for language-nonspecific or language-specific models of lexical access? *International Journal of Bilingualism*, 8, 115–125. doi: 10.1177/13670069040080020101
Hoshino, N. & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, 106, 501–511. doi: 10.1016/j.cognition.2007.02.001
Illes, J., Francis, W. S., Desmond, J. E., Gabrieli, J. D., Glover, G. H., Poldrack, R., … & Wagner, A. D. (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain and language*, 70, 347–363. doi: 10.1006/brln.1999.2186
Jared, D., Pei Yun Poh, R., Paivio, A. (2013). L1 and L2 picture naming in Mandarin–English bilinguals: A test of Bilingual Dual Coding Theory. *Bilingualism: Language and Cognition*, 16, 383–396. doi: 10.1017/S1366728912000685
Jouravlev, O., & Jared, D. (2014). Reading Russian–English homographs in sentence contexts: Evidence from ERPs. *Bilingualism: Language and Cognition*, 17, 153–168. doi: 10.1017/S1366728913000205
Kaushanskaya, M., & Marian, V. (2009a). Bilingualism reduces native-language interference during novel-word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 829–835. doi: 10.1037/a0015275
Kaushanskaya, M., & Marian, V. (2009b). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, 16, 705–710. doi: 10.3758/PBR.16.4.705
Kerkhofs, R., Dijkstra, T., Chwilla, D. J., & De Bruijn, E. R. A. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain Research*, 1068, 170–183. doi: 10.1016/j.brainres.2005.10.087
Klein, D., Zatorre, R. J., Chen, J. K., Milner, B., Crane, J., Belin, P., & Bouffard, M. (2006).

Bilingual brain organization: A functional magnetic resonance adaptation study. *Neuroimage*, 31, 366–375. doi: 10.1016/j.neuroimage.2005.12.012

Kroll, J. F., Bobb, S. C., Misra, M. M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, 128, 416–430. doi: 10.1016/j.actpsy.2008.02.001

Kroll, J. F., Bobb, S., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9, 119–135. doi: 10.1017/S1366728906002483

Kroll, J. F., & Dussias, P.E. (2013). The comprehension of words and sentences in two languages. In T. Bhatia & W. Ritchie (Eds.). *The Handbook of Bilingualism and Multilingualism*, 2nd Edition (pp. 216–243). Malden, MA: Wiley-Blackwell Publishers.

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Valdes-Kroff, J. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. Ross (Ed.), *The Psychology of Learning and Motivation, Volume 56* (pp. 229–262). San Diego: Academic Press.

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.). *The Oxford Handbook of Language Production* (pp. 165–181). Oxford: Oxford University Press.

Kroll, J. F., Gullifer, J., & Rossi, E. (2013). The multilingual lexicon: The cognitive and neural basis of lexical comprehension and production in two languages. *Annual Review of Applied Linguistics*, 33, 102–127. doi: 10.1017/S0267190513000111

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilinguals memory representations. *Journal of Memory and Language*, 33, 149–174. doi: 10.1006/jmla.1994.1008

Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing. In J. F. Kroll & A. M. B. De Groot (Eds.). *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 531–553). New York: Oxford University Press.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, 13, 373–381. doi: 10.1017/S136672891000009X

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll & A. M. B. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 289–307). New York: Oxford University Press.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Li, Y, Yang, J., Scherf, S., & Li, P. (2013). Two faces, two languages: An fMRI study of bilingual picture naming. *Brain and Language*, 127, 452–462. doi: 10.1016/j.bandl.2013.09.005

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 381–390. doi: 10.1037/a0014875

Ma, F., Chen, P., Guo, T., & Kroll, J. F. (2017). When late second language learners access the meaning of L2 words: Using ERPs to investigate the role of the L1 translation equivalent. *Journal of Neurolinguistics*, 41, 50–69.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. De Groot (Eds.). *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.

Marian, V., Shildkrot, Y., Blumenfeld, H. K., Kaushanskaya, M., Faroqi-Shah, Y., & Hirsch, J. (2007). Cortical activation during word processing in late bilinguals: similarities and differences as revealed by functional magnetic resonance imaging. *Journal of Clinical and Experimental Neuropsychology*, 29, 247–265. doi: 10.1080/13803390600659376

McClelland, J. L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception, Part 1: An account of basic findings. *Psychological Review*, 88, 375–405. doi: 10.1037/0033-295X.88.5.375

Meschyan, G., & Hernandez, A. E. (2006). Impact of language proficiency and orthographic transparency on bilingual word reading: An fMRI investigation. *Neuroimage*, 29, 1135–1140. doi: 10.1016/j.neuroimage.2005.08.055

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25–40. doi: 10.1006/jmla.1998.2602

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). Effects of Cognate Status on Word Comprehension in Second Language Learners: An ERP Investigation. *Journal of Cognitive Neuroscience*, 23(7), 1634–1647. doi: 10.1162/jocn.2010.21463

Midgley, K. J., Holcomb, P. J., Van Heuven, W. J. B., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research*, 1246, 123–135. doi: 10.1016/j.brainres.2008.09.078

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, 67, 224–237. doi: 10.1016/j.jml.2012.05.001

Moon, J., & Jiang, N. (2012). Non-selective lexical access in different-script bilinguals. *Bilingualism: Language and Cognition*, 15, 173–180. doi: 10.1017/S1366728911000022

Morford, J. P., Wilkinson, E., Villwock, A., Piñar, P. & Kroll, J. F. (2011). When deaf signers read English: Do written words activate their sign translations? *Cognition*, 118, 286–292. doi: 10.1016/j.cognition.2010.11.006

Park, H. R., Badzakova-Trajkov, G., & Waldie, K. E. (2012). Language lateralisation in late proficient bilinguals: A lexical decision fMRI study. *Neuropsychologia*, 50, 688–695. doi: 10.1016/j.neuropsychologia.2012.01.005

Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language*, 68, 315–332. doi: 10.1016/j.jml.2012.12.003

Pienemann, M., Di Base, B. Kawaguchi, S., & Håkansson, G. (2005). Processing constraints on L1 transfer. In J. F. Kroll & A. M. B. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 128–153). New York: Oxford University Press.

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* doi: 10.1037/a0035583

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General*, 143, 644. doi: 10.1037/a0033194

Qasem, M., & Foote, R. (2010). Cross-language lexical activation. *Studies in Second Language Acquisition*, 32, 111–140. doi: 10.1017/S0272263109990271

Schneider, V. I., Healy, A. F., & Bourne Jr, L. E. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, 46, 419–440. doi: 10.1006/jmla.2001.2813

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, 55, 197–211. doi: 10.1016/j.jml.2006.03.004

Schwartz, A., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, 22, 106–129. doi: 10.1080/01690960500463920

Schwartz, A. I., & Van Hell, J. G. (2012). Bilingual visual word recognition in sentence context. In J. Adelman (Ed.), *Visual word recognition* (pp. 129–148). London, UK: Psychology Press.

Sebastian, R., Kiran, S., & Sandberg, C. (2012). Semantic processing in Spanish - English bilinguals with aphasia. *Journal of Neurolinguistics*, 25, 240–262.

Sunderman, G., & Kroll, J.F. (2006). First language activation during second language lexical processing: An investigation of lexical form meaning and grammatical class. *Studies in Second Language Acquisition*, 28, 387–422. doi: 10.1017/S0272263106060177

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlated of word frequency and cognate effects. *Cerebral Cortex*, 20, 912–928. doi: 10.1093/cercor/bhp153

Talamas, A., Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition*, 2, 45–58. doi: 10.1017/S1366728999000140

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, 104, 12530–12535. doi: 10.1073/pnas.0609927104

Thomas, M. S. C., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language*, 43, 44–66. doi: 10.1006/jmla.1999.2700

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1-L2 intermixing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1412–1431. doi: 10.1037/a0024492

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, 64, 88–107. doi: 10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, 20, 923–927. doi: 10.1111/j.1467-9280.2009.02389.x

Van Hell, J. G., & De Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica*, 128, 431–451. doi: 10.1016/j.actpsy.2008.03.010

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9, 780–789. doi: 10.3758/BF03196335

Van Hell, J. G., & Kroll, J. F. (2012). Using electrophysiological measures to track the mapping of words to concepts in the bilingual brain: A focus on translation. In J. Altarriba & L. Isurin (Eds.), *Memory, Language, and Bilingualism: Theoretical and Applied Approaches* (pp. 126–160). New York: Cambridge University Press.

Van Hell, J. G., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning*, 47, 507–546. doi: 10.1111/0023-8333.00018

Van Heuven, W. J. B., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain*

*Research Reviews*, 64, 104–122. doi: 10.1016/j.brainresrev.2010.03.002

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39, 458–483. doi: 10.1006/jmla.1998.2584

Van Heuven, W. J. B., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex*, 28, 2706–2716. doi: 10.1093/cercor/bhn030

Zhang, S., Morris, M. W., Cheng, C-Y., & Yap, A. (2013). Heritage-culture images disupt immigrants' second language processing, fostering first-language intrusion. Proceedings of the *National Academy of Sciences*, 110, 11272–11277. doi: 10.1073/pnas.130443511

# 14 Sentence Processing and Interpretation in Monolinguals and Bilinguals: Classical and Contemporary Approaches

MATTHEW J. TRAXLER,
LIV J. HOVERSTEN, AND TREVOR A. BROTHERS

University of California, Davis

## Syntactic parsing

Language comprehension involves an integrated suite of cognitive operations that run the gamut from very basic auditory processes to very high-level inferential and theory-of-mind processes (Jackendoff, 2002; Traxler, 2012; Townsend & Bever, 2000; Traxler & Gernsbacher, 2006; Van Gompel, 2013; many chapters in this volume). Somewhere in the middle there are some very important cognitive processes that, together, constitute syntactic parsing. These processes take lexical (word level) information as their inputs. They take into account concurrent prosodic and visual information, as well as prior referential and other discourse context information. They produce a description of the way words in sentences relate to one another, such that default semantic information associated with individual words can either be reinforced or overturned (e.g., the concept or referent to which a word refers may take on a characteristic or play a semantic role in a given sentence that it would not otherwise take on or play, due to the way syntax influences thematic role assignments). One of the most important functions of syntax is that it allows us to say surprising things (à la *The man bit the dog; Matt's ex-wife decided to be nice to him*, etc.). In fact, syntax can allow us to craft utterances that are so surprising that a sentence, although well-formed, cannot be assigned any real meaning at all (as in Chomsky's famous example, *Colorless green ideas sleep furiously*; Chomsky, 1957). Without syntax and syntactic

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

parsing, language and the messages it conveys would be unsurprising and therefore much less interesting.

This chapter provides an overview of classical and contemporary approaches to syntactic parsing, including recent developments in *Bayesian/Noisy Channel* and *dual-streams* (including *good-enough parsing*) approaches to sentence processing and interpretation. It mainly reviews sentence-processing research on monolingual English speakers, as this serves as the foundation of most popular approaches to syntax and syntactic parsing (see Hawkins, 2014). It will then review evidence from the field of cognitive neuroscience on how these sentence-processing mechanisms unfold in real time. The final section of the chapter will then contrast sentence processing in bilinguals and monolinguals. Second language processing provides an interesting window into the acquisition of new syntactic structures. In addition, bilinguals make up the majority of the world's population (Kroll & De Groot, 2005), and an investigation of the differences and similarities between monolinguals and bilinguals will build a more complete picture of human sentence processing.

## Classical approaches to syntax and sentence processing

The classic view of sentence processing views it as being driven by a sequence of mental events, such that conceptually lower-level processes are completed before higher-level processes are initiated. Phonological or orthographic processing precedes lexical access and word recognition, which precede the very important syntactic structure-building processes, which precede sentence-level semantic analysis, which precedes inferencing, contextual and discourse integration, and pragmatic processes. Because this type of account clearly differentiates between lexical level representations and processes and syntactic representations and processes, it requires strong theories of both lexical and syntactic representation. In the early days, phrase-structure grammars based on Chomsky's notions of syntax provided the representational substrate for cognitive theories of syntactic representation and parsing (e.g., *transformational grammar*; Chomsky, 1965). This version of phrase-structure grammar connected two levels of syntactic representation via sets of *transformations*. This account allowed an underlying *kernel* sentence to generate more than one *surface structure* representation, allowing for the same semantic content to be expressed in different overt forms (such as the active-passive voice alternation, among others). Early behavioral research in psychology tested out various predictions of this two-tiered syntactic representation framework, working off the assumption that each operation (or transformation) that was needed to translate between representational levels required the same amount of time to complete. This *derivational theory of complexity* was not very successful at predicting human behavior in a variety of experimental paradigms, but some of the early work did go a long way towards establishing the psychological reality of syntactic phrases as a perceptual unit (Garrett, Bever, & Fodor, 1966; Townsend & Bever, 1978).

Early cognitive work in parsing also adopted a modular view of the language processing system (Fodor, 1983; Frazier, 1979; Frazier & Fodor, 1978; although countervailing interactive views began to emerge during this period as well; Tyler & Marslen-Wilson, 1977). The fundamental perspective of modularity is that complex cognitive processes are broken down into simpler pieces, each of which is undertaken by a dedicated sub-system. To qualify as a module, a cognitive system must have a number of characteristics. The system must be *domain specific* and *informationally encapsulated*. In other words, the computations carried out by the module are based on specific inputs and produce specific outputs; after input is registered or encoded, processing events that are taking place simultaneously elsewhere do not affect the computations taking place within the module. Such systems are also claimed to be genetically determined and instantiated in distinct neural structures. As applied to parsing, the modularity hypothesis contends that auditory/orthographic and lexical processes are completed prior to syntactic parsing; which in turn is completed before other, higher level aspects of processing, such as discourse integration and inferencing. Considerable work in the 1970s through the 1990s addressed questions about whether or to what degree the mental processes involved in syntactic parsing really obeyed the principles of modularity. This work also questioned whether modularity resulted from the functional architecture of the processing system or whether it occurred because some kinds of information take so long to activate that other processes can be completed before that information becomes available ("de-jure" versus "de-facto" modularity; see, e.g., Lewis, 2000).

A number of experimental results from a variety of paradigms have made strict modularity an untenable position with regards to syntactic parsing. There is substantial evidence that parsing operations are affected by discourse context (Altmann, Garnham, & Henstra, 1994; Altmann & Steedman, 1988; Ni, Crain & Shankweiler, 1996), concurrent prosody (which is normally viewed as a separate language characteristic that has strong correlations with syntactic/phrase structure information; Schafer *et al.*, 2000; Speer & Blodgett, 2006; but see Ross & Monnot, 2008), concurrent visual context (Tanenhaus *et al.*, 1995), and aspects of lexico-semantic information (Clifton *et al.*, 2003; Pickering & Traxler, 1998; Traxler & Pickering, 1996; Traxler *et al.*, 2005; Trueswell, Tanenhaus, & Garnsey, 1994). Hence, rather than functioning as an independent intermediate stage between lexical and discourse-level processes, parsing may be more accurately described as an important interface system that mediates among lexical, sentence, and discourse semantics. This mediation involves reciprocal interactions with those connected systems. That interaction may be a consequence of lexically based syntactic representations (Boland & Blodgett, 1996; Boland & Boehm-Jernigan, 2002; Ford *et al.*, 1982; MacDonald *et al.*, 1994; Tooley *et al.*, 2009; Traxler *et al.*, 2014); or it may be driven by strong associations between lexico-semantic (and form) representations and aspects of syntactic structure information (Sag *et al.*, 2003; Vosse & Kempen, 2000, 2009).

## More contemporary approaches

Recent work in syntactic parsing has more fully developed and elaborated concepts relating to interaction between syntactic parsing and linked cognitive systems (like lexical and discourse processing; see van Gompel, 2013, for a recent comprehensive overview). These recent approaches also extend concepts relating to the extent to which parsing processes rely on stored information about the relative frequency of different syntactic structures and parsing operations (Hale, 2011; Jaeger & Snider, 2013; Levy, 2008; Jurafsky, 1996; Trueswell, Tanenhaus, & Kello, 1993). The fundamental notion here is that syntactic structures, like other aspects of language, can be placed along a continuous dimension from more frequent to less frequent. All other things being equal, the syntactic parser will assign higher activation to structural analyses that are more frequent. However, the precise means by which the parser computes frequency has been a topic of intense discussion for quite a while; see, for example, Mitchell's (1987) description of the *Tuning hypothesis* and his analysis of the *grain size* problem. The issue here is that different levels of analysis can produce different patterns of preferences. For example, a language-wide analysis of syntactic structures might reveal that, regardless of the specific words in a sentence, verbs might be most likely to appear with a direct object. At a finer grain, a particular verb might be most likely to appear without a direct object. At an even finer grain, verbs that most often appear without a direct object might appear with one for specific lexical items. The grain size problem has not yet been satisfactorily solved in that we have not established what exact statistical information is represented in memory at what grain sizes; nor how the language processing system weights information at different grains in the service of determining the likelihood of a given syntactic analysis in a given situation.

Adopting the frequency-preference hypothesis leads to a number of predictions for behavior. Some of these predictions have a fair amount of empirical support, and some do not. An example of the former is that, all things being equal, low-frequency syntactic structures produce longer reading times (in, e.g., eye-tracking experiments) than higher-frequency syntactic structures (see Traxler, 2012, Chapter 4, for an overview). An example of the latter is that there is no compelling evidence that parts of sentences where multiple structures are being considered are harder to process than parts of sentences where only a single syntactic analysis is compatible with the input (see, e.g., Traxler, Pickering, & Clifton, 1998).

Frequency-based, information-theoretic approaches to parsing nonetheless contend that processing load is a function of the degree of uncertainty that prevails at any given point in a sentence and the extent to which any part of the input is *surprising* (Hale, 2011; Levy, 2008). *Surprisal* can be computed in a variety of ways, but Bayesian approaches are the most highly favored at the moment. Such approaches view the posterior probability of a given syntactic structure as a function of the strength of "bottom-up" evidence for a given structural analysis and the prior probability (or frequency) of that analysis. Strong differences between posterior and prior probabilities are claimed to cause the highest cognitive processing load. Hence, processing times should increase with increasing *surprisal*. Similar claims have been embedded within non-Bayesian frequency-based accounts (e.g., *constraint based-lexicalist* accounts of parsing; see Jurafsky, 1996; MacDonald *et al.*, 1994; Spivey & Tanenhaus, 1998). There appears to be no fundamental incompatibility between a generic constraint based account and Bayesian approaches to parsing. Constraint-based accounts are particularly flexible with regards to the identification and combination of information that bears on structure-building decisions, especially when implemented within a neural network processing architecture.

A very recent view of syntactic parsing adopts a Bayesian approach to computing syntactic probabilities, and extends this view by emphasizing the concept of a noisy channel (Gibson, Bergen, & Piantadosi, 2013; Gibson *et al.*, 2013). The idea here is that linguistic information, like all perceptual input, is transmitted over channels (auditory and visual) that are noisy. That noise can be external to the individual, or internally caused by random activity within neural systems. As a result, the comprehender can never be 100% certain that the perceived input matches the actual or intended input. The comprehender's job therefore is to determine which of a number of possible messages was the intended one. This calculation is informed by assumptions about what kinds of signal distortion are more versus less likely and what kinds of messages are more or less likely. For example, it is more likely that a part of the signal will be degraded or destroyed by noise; it is less likely that noise will inadvertently add extra information to the signal (although it is possible that the speaker may misspeak in a way that adds extraneous information to the signal). Similarly, it is more likely that speakers will convey plausible messages than implausible ones.

These noisy channel assumptions lead to predictions for behavioral outcomes that do not fall in any obvious way out of regular constraint-based accounts, older style dual-stage accounts, or any other competing account of parsing and sentence interpretation. For example, the noisy channel hypothesis explains why comprehenders judge the quality of the sentence *The woman handed the candle the girl* as being more acceptable than the semantically equivalent *The woman handed the girl to the candle* (Gibson, Bergen, & Piantadosi, 2013; Gibson *et al.*, 2013). In the former case, the interpretation can be "repaired" by assuming that noise has wiped out a small part of the signal (the word *to* between *candle* and *girl*). In the latter case, no such repair is possible, so the comprehender must either assume that the speaker misspoke (as in a word-exchange error) or that the strange message really was intended (or that there is some obscure and inapt metaphoric interpretation). On the other hand, some researchers point out that comprehenders are extremely skilled at detecting even minor grammatical violations very soon after those violations appear in the input (Phillips & Lewis, 2013). It is not at all clear how comprehenders could rapidly and accurately detect syntactically ill-formed strings if they routinely under-specified the syntactic form of an expression, or routinely ignored the actual signal in favor of an "edited" version of the signal.

## Dual streams and good-enough parsing

Noisy channel accounts of syntactic processing stress that comprehenders can choose to accept or reject the cues that they perceive in the signal. Dual-streams approaches and the related good-enough parsing hypothesis make similar claims with regards to the balance of lexical and syntactic information in sentence interpretation (see Ferreira & Patson, 2007). The dual streams approach notes that individual words and syntactically driven interpretive processing are both potential sources of information about meaning. It notes further that meanings that are derived solely from the lexical level can conflict with syntactically driven interpretations. Both noisy channel and dual streams accounts propose that semantic interpretation can be derived solely from lexical information, without any syntax being computed. For instance, if one knew that a scenario involved a baseball player, a ball, an act of catching, and a glove, one could safely assume (almost all of the time) that the action involved an agent (baseball player) undertaking an action (catching) on a theme (baseball) using an instrument (glove), as in *The baseball player caught the ball in his glove*. It would be strange and off-putting to read *The baseball player caught the glove in his ball* or *The ball caught the glove in his baseball player* (these two sentences would require heavy contextual support to attain plausibility).

However, there are many instances where default relationships between role-players are insufficient to determine speaker meaning, either because the concepts themselves do not specify a unique scenario or because the speaker really is trying to convey something surprising or unusual. In those instances, syntactic relationships must be computed, so that the correct (intended) thematic roles can be assigned to the different constituents. Where dual-streams and good-enough parsing depart from prior accounts (whether modular, two-stage or interactive, constraint-based) is in making the strong claim that syntactic form routinely goes un-computed or under-specified (see Frazier & Clifton, 1996, for a related proposal).

This latter assumption helps explain why comprehenders across the age spectrum assign non-licensed or non-standard meanings to certain expressions. For example, comprehenders are quite happy to assign the plausible, but unlicensed, meaning, "The mouse ate the cheese" to the string *The mouse was eaten by the cheese* (Ferreira, Bailey, & Ferraro, 2002; Ferreira & Patson, 2007). Other studies indicate that syntactically licensed interpretations based on how words should be grouped into phrases are also ignored in favor of attractive, plausible, but unlicensed interpretations (Christianson *et al.*, 2001; Christianson, Luke, & Ferreira, 2010; see Traxler, 2014). For example, a comprehender who reads *While Mary was dressing the baby played in the crib*, will often interpret the sentence as meaning "Mary dressed the baby", even though the licensed interpretation is that "Mary dressed herself". On the dual streams and good-enough parsing accounts, the comprehender's derived meaning is based on default lexical associations alone, prior to any syntactic structure-building processes. Beyond that basic assumption, specific dual-stream and good-enough parsing accounts may differ. Some accounts claim that syntactic parsing processes are optional and are foregone by comprehenders when a sensible interpretation falls out of the lexical-level processes themselves. However, at least one recent study that addressed the issue of whether syntactic computations are optional or obligatory produced evidence that syntax is computed even in cases where it is not necessary and not helpful for the completion of a given task (Slattery *et al.*, 2013). The noisy channel account would make the same prediction, but would assume that comprehenders failed to register the syntactic cues that support the reflexive interpretation of …*Mary was dressing (herself)*…. Hence, it is likely that syntactic form is computed automatically, but that comprehenders ignore or suppress implausible, syntactically licensed interpretations in favor of (ungrammatical) but plausible lexically-based interpretations.

## Event-related potentials and sentence processing

Much of the research on sentence processing relies on observations of comprehenders' behavior as they read or listen to sentences of different kinds. Neurophysiological methods, such as event-related potentials (ERPs), offer another window into the processes underlying syntactic parsing and sentence interpretation. The electroencephalogram (EEG) is recorded non-invasively from the scalp and represents the summed electrical activity of large populations of cortical neurons. EEG signals from several trials in the same condition are then averaged together to create an ERP waveform. By providing an online measure of neural activity as comprehension unfolds, ERPs can reveal subtle processing differences before they appear in a measurable form in behavior. In this section we will review some recent developments in ERP research on sentence processing. The final section will then discuss how these techniques have been applied to investigate sentence processing in a second language.

One important ERP component relevant to sentence processing is the P600, which is a positive-going ERP wave typically observed between 500 and 900 ms post-stimulus onset. Differences in the amplitude of the P600 component were first observed by Osterhout and Holcomb (1992) when subjects read sentences containing dispreferred "garden path" continuations (e.g., *The broker persuaded <u>to</u> sell the stock was sent to jail.*). Similar late positivities were later described for syntactic violations involving number (e.g., *The spoiled child <u>throw</u> the toys on the floor.*) and gender agreement (e.g., *The woman congratulated <u>himself</u>…*; Hagoort Brown & Groothusen, 1993; Osterhout & Mobley, 1995). Initially the P600 was interpreted as the syntactic counterpart to an earlier ERP component, the N400, which is also sensitive to linguistic manipulations. While the N400 reflects the semantic fit between a word and its preceding context (Kutas & Hillyard, 1980), the P600 was interpreted as reflecting syntactic mechanisms of structural reanalysis or repair following a grammatical violation.

One aspect that appears critical for triggering the P600 is a participant's conscious awareness of the critical violation. For example, Hasting and Kotz (2008) presented participants with simple auditory phrases that were either grammatically licensed or resulted in an agreement violation (e.g., *they kicks*). When actively attending to these stimuli, participants produced a focal left anterior negativity (LAN) followed by a pronounced P600 effect. In contrast, when participants were presented with the same auditory stimulus while performing a visual distractor task, the anterior negativity was preserved but the P600 effect was completely abolished. These results suggest that early syntactic ERP effects like the LAN may reflect the automatic detection of a syntactic mismatch, while the P600 may reflect more deliberative and resource intensive mechanisms required to repair or resolve these conflicts (see also Batterink & Neville, 2013).

Additional evidence for this "resource-intensive" account of the P600 comes from a study by Kolk and colleagues (2003). In this study participants read a variety of semantic and syntactic violations during EEG recording. Critically, each violation was embedded in either a simple subject-relative clause (e.g., *The poachers that hunted the fox stalked though the woods*) or a more resource demanding object-relative clause (e.g., *The fox that the poachers hunted stalked though the woods*). In an offline error-detection task, violations presented within a syntactically complex, object-relative clause resulted in both longer reaction times and greater error rates. In the EEG task, syntactic complexity also dramatically reduced the amplitude of the P600. Based on these results, the authors concluded that the detection of syntactic anomalies requires the operation of a resource-limited working memory system. By taxing this system with external load, in this case syntactic complexity, detection of these anomalies was delayed or in some cases completely failed to occur.

In addition to factors such as working memory load, individual differences and parsing preferences also appear to play a role in generating the P600. In a study by Kemmerer and colleagues (2006), sentences that violated standard adjective order preferences in English (e.g. *Jennifer rode a huge gray elephant* versus *Jennifer rode a gray huge elephant*) also produced a P600 effect. Critically though, this effect was only present for the 50% of participants (10/20) who consistently classified these adjective combinations as unacceptable during an offline behavioral task. Participants with low rejection rates, who were classified as "insensitive" to these violations, failed to show a P600.

## Semantic P600 effects

While the P600 was initially interpreted as reflecting mechanisms specific to *syntactic* processing, this account was eventually challenged. In a series of ERP studies, a similar late positivity was observed in response to syntactically well-formed but semantically anomalous sentences such as *Every morning at breakfast the eggs would eat…*, (Kuperberg *et al.*, 2003; Kolk *et al.*, 2003; Kim & Osterhout, 2005). What was particularly surprising about these results was that, unlike other semantic anomalies (e.g., *He spread the warm bread with socks*), these thematic role violations did not produce an observable N400 effect. In fact, these violations displayed a *reduced* N400 relative to unassociated (but non-animacy violating) control sentences: *Every morning at breakfast the boys would plant…*

In the ensuing years, a variety of neuro-cognitive models were proposed to account for these results, including a variety of *dual stream* models(see Kuperberg, 2007; Brouwer, Fitz, & Hoeks, 2012 for reviews). As discussed earlier, these dual stream accounts posit a "syntactic parser," which combines parts of speech according to grammatical rules, as well as a "semantic analyzer," which combines words according to their semantic features to construct a rough, first-pass interpretation of incoming language input. According to the semantic analyzer, the content words *morning, breakfast, eggs,* and *eat* should be combined into a plausible scenario where eggs are being eaten. In contrast, the syntactic parser arrives at a conflicting interpretation (which accurately reflects the surface form of the sentence) in which the eggs are doing the eating. In many models, the conflict arising from these streams is ultimately what produces the P600 (Bornkessel-Schlesewsky & Schlesewsky, 2008; van Herten, Chwilla, & Kolk, 2006; van Herten, Kolk, & Chwilla, 2005; Kuperberg, 2007; Kos *et al.*, 2010).

Recently, an alternate interpretation of these semantic P600 effects has been proposed (Brouwer, Fitz, & Hoeks, 2012). Rather than appealing to multiple processing streams, the authors instead highlight the differential sensitivity of the N400 and P600 to different processing demands, namely semantic retrieval and discourse integration. Consistent with other recent accounts (Kutas & Federmeier, 2000; Federmeier and Laszlo, 2009, van Berkum, 2009) the authors propose that the N400 primarily indexes the difficulty of retrieving semantic information stored in long-term memory. In the case of violations such as *Every morning for breakfast the eggs would eat…* the meaning of the critical verb *eat* can be accessed without difficulty, by virtue of the surrounding sentence context. Because of the strong semantic association between *breakfast* and *eating*, the amplitude of the N400 to this target word will remain small. By contrast, upon reaching an integration stage, the reader will encounter substantial difficulty assigning this verb to the preceding noun *eggs*, thereby producing a P600.

This *Retrieval-Integration* account of the P600 can explain a wide variety of findings in the literature. For example, in Dutch, a semantically associated violation such as *The athlete that was by the spear thrown…* produces a P600 without a preceding N400, while a semantically unrelated violation *The athlete that was by the spear summarized…* produces a biphasic pattern, with larger amplitudes for both the N400 and the P600 component (Hoeks, Stowe, & Doedens, 2004).[1] Moreover, while a semantic violation like *The woman told the suitcase…* typically produces a biphasic N400/P600, when this same phrase is embedded in a highly congruent discourse context (where suitcases have been mentioned repeatedly), only a robust P600 effect is observed (Nieuwland & Van Berkum, 2005; for an additional discussion of referential late positivities, see Chapter 17, this volume).

Other evidence suggests that, like the syntactic P600, semantic P600s are also dependent on the explicit detection of an anomaly. For example, if someone is asked "how many animals of each type did Moses bring on the ark?" they will often respond with the answer "two," having failed to realize that it was Noah, not Moses, who gathered the animals. For illusions of this type, the strong semantic association between the violating material and the surrounding context can sometimes result in shallow processing strategies, with many people often missing the critical violation completely (see also Ferreira, Bailey, & Ferraro, 2002). In an ERP study with natural speech stimuli, Sanford and colleagues (2011) compared neural responses to detected and undetected semantic illusions, comparing them to neutral, non-violating control sentences. While all three conditions produced equally small N400 effects, only correctly detected semantic illusions resulted in a significant semantic P600.

Similarly, in an ERP study by Moreno and Kutas (2005), Spanish-English bilingual participants were presented with semantically congruent and incongruent sentences, in both their dominant and non-dominant languages. While participants showed a robust N400 effect of sentence congruity in both Spanish and English, surprisingly, the semantic P600 was significantly reduced for sentences in the participants' non-dominant language. This dissociation between N400 and P600 results suggests that, while semantic facilitation from a preceding context may be relatively automatic, the detection or resolution of semantic anomalies may require greater overall proficiency. Future work will be necessary to determine whether this effect was due to greater demands placed on working memory resources (as in Kolk *et al.*, 2003) or some other feature unique to second-language processing (see later sections for a more detailed discussion).

## The late frontal positivity

While sentence continuations that are semantically or syntactically anomalous produce a posterior positivity, other forms of unexpected linguistic input result in a qualitatively different pattern of ERP results. Critically, a distinct *frontal-distributed* positivity is observed when participants encounter a plausible but unexpected sentence continuation (see Van Petten & Luka, 2012 for a review). This positivity, which is maximal over frontal and left temporal electrode sites, is typically triggered by manipulations of cloze probability (Delong *et al.*, 2011; Federmeier *et al.*, 2007).[2] For example, sentences with highly predictable completions (e.g., *We could tell he was angry by the tone of his voice.*) will show large reductions in N400 amplitude at central-posterior scalp sites. If this same sentence context was instead completed with an unexpected but plausible continuation such as *…his message*, this would result in both a larger N400 and a sustained frontal positivity. The scalp topography of this effect suggests that it is functionally distinct from the posterior positivity observed following outright semantic anomalies (Delong, Quante, & Kutas, 2014). This suggests that readers engage fundamentally different neural mechanisms when encountering unexpected linguistic material that can (or cannot) be successfully integrated into a preceding discourse context.

While these frontally distributed positivities were observed early on in studies of ERP sentence processing (see Kutas, 1993), we still have limited understanding of what this frontal component represents. More recently, it has been suggested that this component may index the costs incurred when predictions about an upcoming lexical item are disconfirmed (e.g., encountering the word *message* instead of the predicted word *voice*; for discussions see Van Petten & Luka, 2012; Kutas, Delong & Smith, 2011). However, considering that this late frontal positivity is not observed in single-word priming studies, even in conditions when predictive strategies are highly encouraged (Lau, Holcomb, & Kuperberg, 2013) it is unlikely that this component is simply indexing the detection of an unpredicted stimulus. Considering the frontal positivity's late time course, as well as its link with sentence or discourse-level processing, this component may reflect updating of a prior discourse context in light of new, unanticipated information (for additional evidence see Brothers, Swaab, & Traxler, 2015).

In summary, it appears that there are two late positive waves following the N400, which likely reflect two distinct sentence processing mechanisms: 1) a posterior positivity (referred to as the P600), which reflects the detection or resolution of a variety of semantic and syntactic anomalies, and 2) a frontally distributed positivity triggered by plausible but unexpected continuations which may require a revision of the ongoing discourse representation. Future work determining the precise conditions that elicit these two ERP components, as well as their neuroanatomical underpinnings, will be critical in shaping our understanding of human sentence processing.

## Hypotheses relating to sentence processing in bilinguals

Bilinguals offer an enlightening contrast to monolinguals. In addition to acquiring two sets of labels for concepts, bilinguals need to acquire two different systems of syntactic or combinatory information. They also need to develop processes and mechanisms that allow them to activate their grammatical knowledge in real time to parse sentences in the second language (L2). One question about L2 sentence processing is whether it parallels processing in a native language (L1) or is qualitatively different from L1 sentence processing. If different, then further questions ensue as to how it differs and why and whether native-like processing is ever possible for a bilingual. Therefore, this review will focus on hypotheses about how and why syntax may be processed differently in a second language and whether bilinguals can achieve native-like proficiency in parsing second-language input. The ERP components reviewed in the previous section will be particularly useful for answering these questions since ERPs offer a direct and temporally sensitive measurement of neural signals during online sentence processing.

While simultaneous and early bilinguals acquiring their L2 in early childhood seem to process the L2 similarly to monolinguals (see Chapters 28 and 29, this volume), later bilinguals who acquire their L2 after puberty may have more difficulty processing the L2 in a native-like manner (Johnson & Newport, 1989; Clahsen & Musken, 1996; Meisel, 1998). Many studies have found differences in syntactic processing between native and late L2 speakers, especially when L2 syntactic features differ from the L1 and grammatical representations for the two languages may compete (Tokowicz & MacWhinney, 2005). These findings have engendered the critical period hypothesis of second language acquisition, which posits a period during infancy and childhood during which L2 exposure must occur for children to acquire native-like proficiency. The hypothesis has taken many forms, though all propose maturational constraints for syntactic processing in a second language.

The *fundamental differences hypothesis* (Bley-Vroman, 1989) was one of the first to propose that an L2 cannot be processed in a native-like manner if it is learned after a critical period. This hypothesis explains a critical period for sentence processing in terms of Universal Grammar, a hypothetical innate language mechanism used by infants to acquire syntactic knowledge of their native language (Chomsky, 1981, 1986; Pinker, 1994, 1999; see also Traxler *et al.*, 2013). Bley-Vroman favors the idea that Universal Grammar is only available for language acquisition during infancy and early childhood, after which parameters become fully tuned and entrenched. According to the fundamental differences hypothesis, if a second language is learned during this critical period, it can be processed in a native-like manner. However, Universal Grammar can no longer be used to tune new or different L2 parameters if it is learned after the critical period. Therefore, late L2 learners must rely on explicit learning and can never process their L2 like the L1, even if their proficiency in the L2 improves.

More recent critical period hypotheses for second language acquisition have explained maturational constraints on sentence processing in other ways. For example, Ullman (2001) proposed the *Declarative/Procedural model*, which posits that both L1 and L2 lexico-semantic processing rely on explicit, declarative memory to store meaning-form pairings of lexical items. The difference between L1 and L2 processing lies in the memory systems used for morphosyntactic processing. L1 syntactic processing is posited to be carried out by procedural memory, through which sequencing and automatic rule-based computation can take place. Ullman argues that only L1 and perhaps an early learned L2 can use procedural memory to process syntactic properties of language. A later learned L2 may instead depend more on declarative memory for these processes, since it may improve throughout adolescence (Di Giulio *et al.*, 1994). Behavioral and neuroimaging evidence confirms that L1 and L2 semantic processing are relatively indistinguishable in terms of cognitive mechanisms and underlying activated neural structures (e.g., Hahne, 2001; Wartenburger *et al.*, 2003). However, L1 and L2 morphosyntactic processing may be carried out by at least partially separate neural substrates (e.g. Dehaene *et al.*, 1997; Ullman, 2001, 2005). The Declarative/Procedural model allows the possibility that practice effects cause a shift in late L2 learners from reliance on declarative memory to reliance on procedural memory for sentence processing. This shift occurs for very experienced, highly proficient learners. Thus, it is possible for this hypothesis to accommodate findings of native-like sentence processing in highly proficient late L2 learners.

Clahsen and Felser (2006), on the other hand, have proposed the *shallow structure hypothesis* (*SSH*), which states that only certain aspects of L2 syntactic processing in late learners are doomed to remain less-specified, or "shallow," after a high degree of proficiency in the language is attained (reminiscent of the good-enough processing account; Ferreira, Bailey, & Farraro, 2002). While they agree that processing of morphological aspects of an L2 may become native-like in nature, they assert that other aspects of syntactic processing can never become fully native-like. These conclusions largely rest on data from experiments assessing L2 processing of nonlocal dependencies. One experiment (Marinis *et al.*, 2005) showed that late L2 learners did not make use of intermediate gaps in a long-distance *wh-* dependency (e.g., *The manager* who$_i$ *the consultant claimed* e$'_i$ *that the new proposal had pleased* e$_i$ *will hire five workers tomorrow*), whereas native speakers did. They use this data to make the claim that processing differences on complex syntactic structures such as these reflect qualitatively different mechanisms in native versus non-native speakers, rather than simply quantitative differences.

## ERP evidence in bilingual sentence processing

The major thread that connects these critical period hypotheses is that they all claim an underlying qualitative difference between L1 and L2 sentence processing. However, much of the initial evidence for critical period effects on L2 sentence processing has come from behavioral performance on tasks like grammaticality judgment (e.g., Johnson & Newport, 1989). While these tasks reflect the output of syntactic parsing, they do not directly measure online processing in real time. Since ERPs are a direct measure of online processing, they can help address whether L2 parsing difficulties indeed result from the use of qualitatively different neural mechanisms or from less efficient use of the same mechanisms. Two major ERP components that can address this question are the P600 and the LAN. As discussed in the previous section, the P600 has been associated with conscious, controlled detection or repair of grammatical violations, while the LAN may reflect automatic, implicit syntactic processing. Because the critical period hypotheses reviewed above propose that only native speakers can process syntax automatically and efficiently, they would predict that the LAN should be most affected by age of acquisition (AoA).

Weber-Fox and Neville (1996) conducted one of the first studies that used ERPs to examine the effect of AoA on online sentence processing. They tested English monolinguals and five groups of Chinese-English bilinguals that varied in their L2 AoA. A P600 was obtained for phrase structure violations in all but the latest L2 learners (AoA > 16 years), but was delayed and reduced in amplitude for bilinguals who acquired their L2 between 11-13 years. An early anterior negativity was also obtained for all groups, but its distribution varied according to AoA. Whereas the early bilinguals (AoA < 11 years) showed a similar left-lateralized distribution to the monolinguals (i.e., a LAN), the late bilinguals (AoA > 11 years) had a more broadly distributed anterior negativity. Weber-Fox and Neville (1996) took these results to mean that late L2 learners do not process syntactic properties with the same underlying neural mechanisms that native speakers or early L2 learners use. Many other studies failed to find a LAN or even a more broadly distributed anterior negativity in late L2 learners (Hahne & Friederici, 2001; Hahne, 2001; Wartenburger *et al.*, 2003; Mueller *et al.*, 2005), suggesting that these learners cannot acquire native-like processing mechanisms for morphosyntactic aspects of language.

However, several studies claiming an absence of an early anterior negativity in late L2 learners may have missed reliable effects due to methodological issues (see Steinhauer & Connolly, 2008 for a discussion). In particular, aspects of the stimuli may have affected the baseline for comparison. Additionally, many of these early studies did not take L2 proficiency into account. In fact, proficiency and AoA were negatively correlated in the Weber-Fox and Neville (1996) study. High L2 proficiency was related to early AoA and vice versa, so it is unclear whether the effects were due to a maturational constraint for L2 sentence processing or lower proficiency in the late learners. Other studies did not report proficiency levels (Hahne & Friederici, 2001) or reported worse grammaticality judgment performance in their L2 participants than in native speakers (Mueller *et al.*, 2005). Green (2003) suggested that increasing proficiency may decrease possible differences in the neural structures underlying L1 and L2 syntactic processing found in neuroimaging studies. If this is the case, then the differences observed between L2 and L1 sentence processing might reflect quantitative rather than qualitative differences. Highly proficient late bilinguals, then, may be able to process their L2 in a native-like manner.

Because AoA and proficiency are often highly correlated, some studies have trained participants on an artificial or miniature language to dissociate the two. Artificial languages allow researchers to tightly control the properties of the language being tested (see Ettlinger *et al.*, 2015 for evidence linking artificial language learning to natural language learning). A miniature language, on the other hand, can allow participants to learn some aspect of a full-blown natural language in a relatively short period of time (see Mueller *et al.*, 2006). The goal of these studies was to see whether the native-like biphasic LAN/P600 can be evoked when participants are highly proficient in the syntax of the artificial or miniature language. Friederici, Steinhauer, and Pfiefer (2002) first showed a native-like biphasic LAN/P600 in response to phrase structure violations in proficient late learners of an artificial miniature language. The same violations did not elicit either component in a control group that was trained on the semantics but not the grammatical aspects of the language. Mueller *et al.* (2006) performed a study that failed to show native-like LAN effects in low-proficiency late learners of a miniature Japanese grammar. However, after further training, the same participants showed both improved grammaticality judgment accuracy (95-99% versus 84% in the first study) and a biphasic LAN/P600 response to violations that was indistinguishable from that of native speakers of Japanese to the same syntactic violations.

Critically, Rossi and colleagues (2006) and Bowden and colleagues (2013) also showed a similar native-like biphasic LAN/P600 response in highly proficient late L2 learners of full-blown natural languages (English and Spanish L2, respectively). In both of these studies, participants with lower proficiency yielded only a P600 in response to violations, similar to earlier studies claiming to support the critical period hypothesis. Higher proficiency participants, on the other hand, yielded both the P600 and the LAN that is the hallmark of native-like processing. These studies confirm the results of studies that used miniature languages, showing that native-like online processing may be attainable for both morphological and syntactic aspects of language even for late L2 learners, provided that they learn the L2 to a high degree of proficiency.

## Neural markers of L2 sentence processing

Based on the evidence reviewed thus far, it may be expected that the ERP signature of morphosyntactic processing changes throughout the course of second language acquisition (SLA). A series of studies have directly investigated the online processing of lexico-semantic and morphosyntactic aspects of language during the early stages of late SLA. After just one month of French language instruction, a semantic violation elicited a small, delayed N400, which became larger and earlier with increasing L2 exposure (McLaughlin, Osterhout, & Kim, 2004). Interestingly, a syntactic violation also elicited an N400 after one month of instruction (Osterhout *et al.*, 2006). This syntactic N400 response could mean that the learners relied heavily on lexical rather than rule-based knowledge to process the violation, a result predicted by Ullman's Declarative/Procedural model. After four months of instruction, the same syntactic violation elicited a small, delayed P600 that was preceded by an N400 in some learners. This was interpreted as a transition stage toward controlled grammatical processing of the violation. At nine months of instruction, the violation elicited a larger, earlier P600, much like that elicited in native speakers. This P600 effect likely reflects more native-like structural reanalysis or repair processes at an intermediate proficiency level.

Steinhauer, White, and Drury (2009) referenced this longitudinal within-subjects data, in combination with cross-sectional data from other ERP studies of morphosyntactic processing across the L2 proficiency spectrum. They proposed a hypothetical time course of the shift in L2 grammatical processing over the course of SLA as indexed by ERP measures. This time course begins with the changes in ERP responses from low to intermediate proficiency reported by Osterhout and colleagues—an N400 at very low proficiency, a small, delayed P600 (perhaps preceded by an N400) at low to intermediate proficiency, and a larger and earlier P600 at intermediate proficiency. Their account proposes that an anterior negativity will begin to precede the P600 as proficiency increases to a higher intermediate level (Weber-Fox & Neville, 1996; Friederici, 2002; Bowden *et al.*, 2013). This negativity will become left-lateralized (i.e., a LAN) only at the highest proficiency level (Ojima *et al.*, 2005; Rossi *et al.*, 2006; Bowden *et al.*, 2013). At this level, even late L2 learners will exhibit native-like automatized syntactic processing.

Steinhauer and colleagues (2009) acknowledge that, while a shift from heavily lexically influenced interpretation to more rule-bound interpretation may be the prototypical shift in processing with increasing L2 proficiency, the exact timing of the shift and its endpoint may vary across individual learners (see also Tanner *et al.*, 2013). Additionally, the time course may vary according to each particular structure due to a learner's differential experience with different types of structures. For example, Osterhout and colleagues (2006) showed the early stages of these transitions for phrase structure violations, but no differential ERP responses were elicited by number agreement violations by the end of nine months of L2 instruction. Similarly, Hahne, Mueller, and Clahsen (2006) showed that Russian-German bilinguals exhibited a LAN only for past participle violations but not for noun pluralization violations. Behavioral results indicated high proficiency for past participle constructions but lower proficiency with noun plurals. Thus, proficiency seems to vary based on the structure examined, and LAN results have been shown to track proficiency with the particular structure tested.

The time course may also differ according to the method of training in the L2. Morgan-Short and colleagues (2012) tested this hypothesis by training two groups of participants on an artificial language with two different methods. The explicitly trained group received clear instructions on the grammar rules of the language. Low proficiency participants in this group produced no ERP effects of a syntactic violation, while high proficiency participants in this group produced a P600 to the violation. Conversely, the implicitly trained group was exposed to many examples of the correct use of the grammar without any explicit training in the rules of the language. A syntactic violation elicited an N400 in the low proficiency participants in this group but a native-like biphasic LAN/P600 response in the high proficiency participants in this group. This result suggests that implicit training may have allowed participants to better internalize the grammatical rules of the language. The result parallels the distinctions made in the Declarative/Procedural model: the explicitly trained group may have relied heavily on declarative knowledge of the grammar rules to perform the task, whereas the implicitly trained group may have proceduralized the rules to process violations in a more automatized, native-like manner. Similarly, Linck, Kroll, and Sunderman (2009) found proficiency differences between English-Spanish bilinguals based on whether they were trained in a classroom setting or through immersion in the L2. Together, these results suggest that the acquisition setting may play an important role in the time course of SLA and possibility of native-like achievement in syntactic processing.

Finally, the time course of L2 syntactic acquisition of particular structures may depend on the typological distance between L1 and L2. According to the *shared syntax account*, bilinguals build their L2 on the syntax of their L1 as much as possible. For shared structures or grammatical aspects of the two languages, this strategy may result in positive transfer effects where the bilingual maps L2 syntax onto their already efficient L1 processing mechanisms. However, if the L1 and L2 rules differ for a particular structure, the shared syntax account would predict negative transfer effects. In this case, the two structures compete and bilinguals may have a harder time overriding their automatic L1 processing in order to properly process the L2 structure.

Evidence of this effect comes from French-English and English-French bilinguals who read L2 sentences that conflicted in the grammatical rules between the two languages. French-English bilinguals seemed to transfer their French L1 knowledge to their English L2 because they were more likely to accept an ungrammatical English sentence as correct when the structure was licensed in French (White, 1989a, 1989b, 1991). English-French bilinguals generalized their English L1 knowledge to their French L2 such that their eye movement record showed a greater processing load for sentences that were licensed in French but unlicensed in English (Frenck-Mestre, 1998). Tokowicz and MacWhinney (2005) and Morett and MacWhinney (2012) similarly showed ERP evidence that structures that are shared between L1 and L2 are processed in a more native-like manner than those that are more dissimilar. Other studies have demonstrated shared syntactic structures using a cross-language syntactic priming paradigm (e.g., Hartsuiker & Pickering, 2008; Loebell & Bock, 2003; Salamoura & Williams, 2006). It appears that priming only occurs when the structure examined is completely shared (i.e., same word order rules) across languages (Weber & Indefrey, 2009).

Such evidence supports the shared syntax account because it shows that bilinguals may create integrated representations of syntactic structures across their languages, perhaps to maximize their efficiency and automaticity in L2 processing. It also suggests that the time course of acquisition may vary according to whether a particular grammatical aspect is shared or distinct across the two languages. If instead, some aspect is L2-specific (i.e., gender agreement for a native English speaker learning French), the time course may differ from either of the above options (shared vs. competitive). However, it is clear from several studies that even late L2 learners can learn L2-specific grammatical aspects to a proficiency that appears native-like (e.g., Foucart & Frenck-Mestre, 2012).

While late L2 learners may be slower and follow an acquisition trajectory that may not parallel that of L1 acquisition in children, the evidence reviewed here suggests that they are capable of native-like, automatized syntactic processing once they reach a high enough proficiency. This evidence would be difficult to reconcile with either the fundamental differences hypothesis or the SSH. However, the scarcity of evidence for late L2 processing of long-distance dependencies, on which the SSH depends heavily, precludes complete abandonment of this hypothesis. On the other hand, the Declarative/Procedural model is compatible with a substantial portion of the evidence we have reviewed here. Several studies supported a difference in late L2 learners' processing along declarative/procedural memory lines, either in terms of acquisition mode (Morgan-Short *et al.*, 2012) or levels of proficiency (e.g., Bowden *et al.*, 2013; Mueller *et al.*, 2005; Mueller *et al.*, 2006; Weber-Fox & Neville, 1996). In line with this model, current evidence reviewed here suggests that late L2 learners can incrementally proceduralize their syntactic knowledge of L2 such that it becomes automatized and indistinguishable from native speakers in terms of behavioral performance, cognitive processing, and neural underpinnings.

## Conclusions

We have made a lot of progress in the last 20 years in our understanding of how proficient readers parse and understand the meaning of sentences during comprehension. In concert with behavioral measures, ERP evidence suggests that both syntax and semantics are combined early on in sentence comprehension to shape the final meaning of an utterance. These studies reveal dissociable neural mechanisms for i) the retrieval of semantic concepts (N400), ii) the syntactic and thematic integration of new constituents (P600), and iii) discourse integration of new and unexpected information (frontal positivity). ERPs also provide evidence that late L2 learners develop sentence-processing abilities with increasing exposure and can attain native-like processing with high enough proficiency despite a late age of acquisition. While the learning trajectory seems to be primarily influenced by proficiency, it may also differ based on factors such as individual variability, type of training, the particular structure examined, or typological distance between the native and second languages.

## Author Notes

This project was supported in part by the National Science Foundation (1024003; 1355776) and the National Institutes of Health (1R01HD073948; MH099327; 1R21DC013617-01A1).

## NOTES

1 Some studies have shown different patterns of results for semantically unrelated anomalies, observing either just an N400 effect (Kim & Osterhout, 2005) or just a P600 effect (Kuperberg *et al.*, 2007). Overall, a review of the literature suggests that, for these types of violations, a biphasic N400/P600 effect is the norm (e.g., Stroud, 2009; Diaz & Swaab, 2007; Kuperberg *et al.*, 2010; Pijnaker *et al.*, 2010; Nieuwland & Van Berkum, 2006). Multiple factors including list composition, component overlap, and individual differences in working memory capacity (Nakano, Saron, & Swaab, 2010) are likely to be responsible for these differences across studies.

2 "Cloze probability" is defined as the likelihood that a participant will produce a particular continuation in an offline sentence completion task.

## REFERENCES

Altmann, G. T. M.., Garnham, A., & Henstra, J.A. (1994). Effects of syntax in human sentence parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 209–216. doi: 10.1037/0278-7393.20.1.209

Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human

sentence processing. *Cognition*, *30*, 191–238. doi: 10.1016/0010-0277(88)90020-0
Batterink, L., & Neville, H. J. (2013). The human brain processes syntax in the absence of conscious awareness. *The Journal of Neuroscience*, *33*(19), 8528–8533. doi: 10.1523/JNEUROSCI.0618-13.2013
Bley-Vroman, R. (1989). What is the logical problem of foreign language learning. *Linguistic perspectives on second language acquisition*, *4*, 1–68. doi: 10.1017/cbo9781139524544.005
Boland, J. E., & Blodgett, A. (2006). Argument status and pp-attachment. *Journal of Psycholinguistic Research*, *35*, 385–403. doi: 10.1007/s10936-006-9021-z
Boland, J. E., & Boehm-Jernigan, H. (1998). Lexical constraints and prepositional phrase attachment. *Journal of Memory and Language*, *39*, 684–719. doi: 10.1006/jmla.1998.2591
Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*, 55–73. doi: 10.1016/j.brainresrev.2008.05.003
Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, *51*, 2492–2511. doi: 10.1016/j.neuropsychologia.2013.09.004
Brothers, T., Swaab, T. Y., & Traxler, M. J. (2015). Effects of prediction and contextual support on lexical processing: Prediction takes precedence. *Cognition*, *136*, 135–149. doi: 10.1016/j.cognition.2014.10.017
Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143. doi: 10.1016/j.brainres.2012.01.055
Chomsky, N. (1957). *Syntactic structures*. The Hague/Paris: Mouton.
Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
Chomsky, N. (1986).*Knowledge of language: Its nature, origin, and use*. Greenwood Publishing Group.
Chomsky, N. Lectures on government and binding, 1981. *Foris, Dordrecht*.
Christianson, K., Luke, S., & Ferreira, F. (2010). Effects of plausibility on structural priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 538–544. doi: 10.1037/a0018027
Christianson, K. A., Hollingworth, A., Halliwell, J., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*, 368–407. doi: 10.1006/cogp.2001.0752
Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(01), 3–42. doi: 10.1017/S0142716406060024
Clahsen, H., & Muysken, P. (1996). How adult second language learning differs from child first language development. *Behavioral and Brain Sciences*, *19*(04), 721–723. doi: 10.1017/S0140525X00043600
Clifton, C., Jr., Traxler, M. J., Williams, R., Mohammed, M., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, *49*, 317–334. doi: 10.1016/S0749-596X(03)00070-6
Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., & Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport*, *8*, 3809–3815. doi: 10.1097/00001756-199712010-00030
DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia*, *61*, 150–162. doi: 10.1016/j.neuropsychologia.2014.06.016
Delong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology*,

*48*, 1203–1207. doi: 10.1111/j.1469-8986.2011.01199.x

DiGiulio, D. V., Seidenberg, M., O'leary, D. S., & Raz, N. (1994). Procedural and declarative memory: A developmental study. *Brain and Cognition*, 25, 79–91. doi: 10.1006/brcg.1994.1024

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P. (2015). The relationship between artificial and second language learning. *Cognitive Science*, *40*, 822–847. doi: 10.1111/cogs.12257

Federmeier, K. D., & Laszlo, S. (2009). Time for meaning: Electrophysiology provides insights into the dynamics of representation and processing in semantic memory. *Psychology of Learning and Motivation*, *51*, 1–44. doi: 10.1016/S0079-7421(09)51001-8

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*, 75–84.

Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, *1*, 71–83.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Fodor, J. A. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Ford, M., Bresnan, J. W., & Kaplan, R. M. (1982). A competence-based theory of syntactic closure. In J. W. Bresnan (Ed.), *The Mental Representation of Grammatical Relations* (pp. 727–796). Cambridge, MA: MIT Press.

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *66*, 226–248.

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Doctoral dissertation, University of Connecticut.

Frazier, L., & Clifton, C., Jr. (1996). *Construal*. Cambridge, MA: MIT Press.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*, 291–325.

Frenck-Mestre, C. (1998). Overcoming parameters: An on-line look at bilingual sentence processing. In *39th annual meeting of the Psychonomic Society, Dallas, TX*.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, *99*, 529-534.

Garrett, M., Bever, T. G., & Fodor, J. (1966). The active use of grammar in speech perception. *Perception and Psychophysics*, *1*, 30–32.

Gibson, E., Bergen, L., & Piantadosi, S. (2013). The rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1216438110

Gibson, E., Piantadosi, S., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of cross-linguistic word-order variation. *Psychological Science*, *4*, 1079–1088.

Green, D. W. (2003). The neural basis of the lexicon and the grammar in L2 acquisition. In R. van Hout, A. Hulk, F. Kuiken, & R. Towell (Eds.), *The interface between syntax and the lexicon in second language acquisition*. Amsterdam: John Benjamins.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*, 439–483.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*, 251–266.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*, 123–141.

Hahne, A., Mueller, J. L., & Clahsen, H. (2006). Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal of cognitive neuroscience*, *18*, 121–134.

Hale, J. T. (2011). What a rational parser would do. *Cognitive Science*, *35*, 399–443.

Hartsuiker, R. J., & Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychologica*, *128*, 479–489.

Hasting, A. S., & Kotz, S. A. (2008). Speeding up syntax: On the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of Cognitive Neuroscience*, *20*, 1207–1219.

Hawkins, J. A. (2014). *Cross-Linguistic Variation and Efficiency*. Oxford, England: Oxford University Press. doi: 10.1093/acprof:oso/9780199664993.001.0001

Hoeks, J. C., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: the interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, *19*, 59–73. doi: 10.1016/j.cogbrainres.2003.10.022

Jackendoff, R. (2002). *Foundations of Language*. Oxford, England: Oxford University Press. doi: 10.1093/acprof:oso/9780198270126.001.0001

Jaeger, F. T., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*, 57–83. doi: 10.1016/j.cognition.2012.10.013

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive psychology*, *21*, 60–99. doi: 10.1016/0010-0285(89)90003-0

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*, 137–194. doi: 10.1207/s15516709cog2002_1

Kemmerer, D., Weber-Fox, C., Price, K., Zdanczyk, C., & Way, H. (2007).Big brown dog or brown big dog? An electrophysiological study of semantic constraints on prenominal adjective order. *Brain and Language*, *100*, 238–256. doi: 10.1016/j.bandl.2005.12.002

Kim, A. & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*, 205–225.

Kolk, H. H., Chwilla, D. J., van Herten, M., & Oor, P. J. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*, 1–36. doi: 10.1016/S0093-934X(02)00548-5

Kos, M., Vosse, T., Van Den Brink, D., Hagoort, P., et al. (2010). About edible restaurants: conflicts between syntax and semantics as revealed by ERPs. *Frontiers in Psychology*, *1*, 1–11. doi: 10.3389/fpsyg.2010.00222

Kroll, J., & De Groot, A. M. B. (2005). Handbook of Bilingualism: Psycholinguistic Approaches. Oxford, England: Oxford University Press.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49. doi: 10.1016/j.brainres.2006.12.063

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*, 117–129. doi: 10.1016/S0926-6410(03)00086-7

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*, *8*, 533–572. doi: 10.1080/01690969308407587

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205. doi: 10.1126/science.7350657

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. *Predictions in the Brain: Using Our Past to Generate a Future*, 190–207. doi: 10.1093/acprof: oso/9780195395518.003.0065

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, *25*, 484–502. doi: 10.1162/ jocn_a_00328

Lewis, R. (2000). Specifying architectures for language processing: Process, control, and memory in parsing and interpretation. In M Crocker, MJ Pickering, & C Clifton, Jr. (Eds.), *Architectures and Mechanisms for Language Processing* (pp. 56–89). Cambridge, England: Cambridge University Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177. doi: 10.1016/j.cognition.2007.05.006

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, *20*, 1507–1515. doi: 10.1111/j.1467-9280.2009.02480.x

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, 41(5; ISSU 387), 791–824. doi: 10.1515/ ling.2003.026

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703. doi: 10.1037/0033-295X.101.4.676

Marinis, T., Roberts, L., Felser, C., &Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, *27*, 53–78. doi: 10.1017/s0272263105050035

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature neuroscience*, *7*, 703–704. doi: 10.1038/nn1264

Meisel, J. M. (1998). Parametric change in language development: Psycholinguistic and historical perspectives on second language acquisition. *Lynx: Panorámica de estudioslingüísticos*, *6*, 18–36.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and Performance XII* (pp. 601–618). Hillsdale, NJ: Erlbaum.

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: An electrophysiological exploration in both languages of Spanish-English bilinguals. *Cognitive Brain Research*, *22*, 205–220. doi: 10.1016/j. cogbrainres.2004.08.010

Morett, L. M., & Macwhinney, B. (2013). Syntactic transfer in English-speaking Spanish learners. *Bilingualism: Language and Cognition*, *16*, 132–151. doi: 10.1017/ S1366728912000107

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, *24*, 933–947. doi: 10.1162/jocn_a_00119

Mueller, J. L. (2006). L2 in a nutshell: The investigation of second language processing in the miniature language model. *Language Learning*, *56*, 235–270. doi: 10.1111/j.1467-9922.2006.00363.x

Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of*

*Cognitive Neuroscience*, *17*, 1229–1244. doi: 10.1162/0898929055002463

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden paths: Assessing the contributions of syntax, semantics, and plausibility in resolving ambiguities. *Language and Cognitive Processes*, *11*, 283–334. doi: 10.1080/016909696387196

Nieuwland, M. S., & Van Berkum, J. J. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, *24*, 691–701. doi: 10.1016/j.cogbrainres.2005.04.003

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience*, *17*, 1212–1228. doi: 10.1162/0898929055002436

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806. doi: 10.1016/0749-596X(92)90039-Z

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*, 739–773. doi: 10.1006/jmla.1995.1033

Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice Learners, Longitudinal Designs, and Event-Related Potentials: A Means for Exploring the Neurocognition of Second Language Processing. *Language Learning*, *56*, 199–230. doi: 10.1111/j.1467-9922.2006.00361.x

Phillips, C., & Lewis, S. (2013). Derivational order in syntax: Evidence and architectural consequences. *Studies in Linguistics*, *6*, 11–47.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 940–961. doi: 10.1037/0278-7393.24.4.940

Pinker, S. (1994). *The Language Instinct*. New York, NY: Harper. doi: 10.1037/e412952005-009

Pinker, S. (1999). *Words and Rules*. New York, NY: Basic Books.

Ross, E. D., & Monnot, M. (2008). Neurology of affective prosody and its functional anatomic organization in right hemisphere. *Brain and Language*, *104*, 51–74. doi: 10.1016/j.bandl.2007.04.007

Rossi, S., Gugler, M., Friederici, A., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *18*, 2030–2048. doi: 10.1162/jocn.2006.18.12.2030

Sag, I. A., Wasow, T., & Bender, E. M. (2003). *Syntactic Theory: A Formal Introduction*. Palo Alto, CA: Stanford University Press.

Salamoura, A., & Williams, J. N. (2006). Lexical activation of cross-language syntactic priming. *Bilingualism: Language and Cognition*, *9*, 299–307. doi: 10.1017/S1366728906002641

Sanford, A. J., Leuthold, H., Bohan, J., & Sanford, A. J. (2011). Anomalies at the borderline of awareness: An ERP study. *Journal of Cognitive Neuroscience*, *23*, 514–523. doi: 10.1162/jocn.2009.21370

Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000).Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, *29*, 169–182. doi: 10.1023/A:1005192911512

Slattery, T., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from flawed semantic processing. *Journal of Memory and Language*, *69*, 104–120. doi: 10.1016/j.jml.2013.04.001

Speer, S. R., & Blodgett, A. (2006). Prosody. In M. J. Traxler, & M. A. Gernsbacher (Eds.), *The Handbook of Psycholinguistics*, 2nd Edition (pp. 505–538). Amsterdam, The Netherlands: Elsevier. doi: 10.1016/B978-012369374-7/50014-6

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1521–1543. 10.1037/0278-7393.24.6.1521

Steinhauer, K., & Connolly, J. F. (2008). Event-related potentials in the study of language. *Handbook of the neuroscience of language*, 91–104. doi: 10.1016/B978-0-08-045352-1.00009-4

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, *25*, 13–41. doi: 10.1177/0267658308098995

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language. *Science*, *268*, 1632–1634. doi: 10.1126/science.7777863

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, *16*, 367–382. doi: 10.1017/S1366728912000302

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in second language acquisition*, *27*, 173–204. doi: 10.1017/s0272263105050102

Tooley, K. M., Traxler, M. J., & Swaab, T. Y. (2009). Electrophysiological and behavioral evidence of syntactic priming in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 19–45. doi: 10.1037/a0013984

Townsend, D. J., & Bever, T. G. (1978). Interclause relations and clausal processing. *Journal of Verbal Learning and Verbal Behavior*, *17*, 509–521. doi: 10.1016/S0022-5371(78)90304-3

Townsend, D., & Bever, T. G. (2001). *Sentence Comprehension*. New York, NY: Bradford.

Traxler, M. J. (2014). Trends in syntactic parsing: Anticipation, Bayesian estimation, and good-enough parsing. *Trends in Cognitive Sciences*, *18*, 605–611. doi: 10.1016/j.tics.2014.08.001

Traxler, M. J., Boudewyn, M., & Loudermilk, J. (2013). What's special about human language? The contents of the "Narrow Language Faculty" revisited. *Language and Linguistics Compass*, *6*, 611–621. doi: 10.1002/lnc3.355

Traxler, M. J., & Gernsbacher, M. A. (2006). *The Handbook of Psycholinguistics*. Amsterdam, The Netherlands: Elsevier. doi: 10.1016/B978-012369374-7/50000-6

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, *35*, 454–475. doi: 10.1006/jmla.1996.0025

Traxler, M. J., Pickering, M. J., & Clifton, C. (1996). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, *39*, 558–592. doi: 10.1006/jmla.1998.2600

Traxler, M. J., Williams, R. S., Blozis, S. A., & Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language*, *53*, 204–224. doi: 10.1016/j.jml.2005.02.010

Traxler, M. J., Tooley, K. M., & Pickering, M. J. (2014). Syntactic priming during sentence comprehension: Evidence for the lexical boost. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 905–918. doi: 10.1037/a0036377

Traxler, M. J. (2012). *Introduction to Psycholinguistics: Understanding Language Science*. Boston, MA: Wiley-Blackwell.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences

on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285–318. doi: 10.1006/jmla.1994.1014
Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 528–553. doi: 10.1037/0278-7393.19.3.528
Tyler, L. K., & Marslen-Wilson, W. D. (1977).The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 683–692. doi: 10.1016/S0022-5371(77)80027-3
Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*, 105–122. doi: 10.1017/s1366728901000220
Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 141-178). Georgetown University Press.
Van Berkum, J. J. (2009). The neuropragmatics of 'simple' utterance comprehension: An ERP review. *Semantics and Pragmatics: From Experiment to Theory*, 276–316.
Van Gompel, R. P. G. (2013). *Sentence Processing: Current Issues in the Psychology of Language*. New York, NY: Psychology Press.
Van Herten, M., Chwilla, D. J., & Kolk, H. H. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*, 1181–1197. doi: 10.1162/jocn.2006.18.7.1181
Van Herten, M., Kolk, H. H., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*, 241–255. doi: 10.1016/j.cogbrainres.2004.09.002
Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*, 176–190. doi: 10.1016/j.ijpsycho.2011.09.015
Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and lexicalist grammar. *Cognition*, *75*, 105–143. doi: 10.1016/S0010-0277(00)00063-9
Vosse, T., & Kempen, G. (2009). The unification space implemented as a localist neural net: Predictions and error-tolerance in a constraint-based parser. *Cognitive Neurodynamics*, *3*, 331–346. doi: 10.1007/s11571-009-9094-0
Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron*, *37*, 159–170. doi: 10.1016/S0896-6273(02)01150-9
Weber, K., & Indefrey, P. (2009). Syntactic priming in German-English bilinguals during sentence comprehension. *NeuroImage*, *46*, 1164–1172. doi: 10.1016/j.neuroimage.2009.03.040
Weber-Fox, C., & Neville, H. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*, 231–256. doi: 10.1162/jocn.1996.8.3.231
White, L. (1989a). The principle of adjacency in second language acquisition: Do L2 learners obey the subset principle? In S. Gass & J. Schacter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 134-159). Cambridge University Press. doi: 10.1017/CBO9781139524544.010
White, L. (1989b). *Universal grammar and second language acquisition*. Amsterdam: Benjamins. doi: 10.1075/lald.1
White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research*, *7*, 133–161. doi: 10.1177/026765839100700205

# 15 The Comprehension of Anaphora and Verb Agreement

JANET L. NICOL AND ANDREW BARSS

University of Arizona

## Introduction

In this chapter, we review some of the research on the comprehension of two types of agreeing phrases: anaphors, which agree with an antecedent, and verbs, which agree with a subject. We restrict our attention to within-sentence agreement because of the intrinsic similarity in the constraints on antecedent-anaphor and subject-verb agreement. In addition, we focus on studies that probe comprehension as it is occurring, that is, *online* research. Finally, because languages differ with respect to the type and extent of agreement marking and the manifestation of pronominal forms, cross-linguistic processing data often have different characteristics. Due to length limitations, we restrict the scope of our review to studies of English.

We begin with observations about anaphora and agreement and their treatment within formal syntax. Then we turn to empirical research in these areas. We conclude that comprehension processes are similar for the two.

## Some facts about anaphora and agreement

In running discourse, referents are introduced, and repeated reference is made to them as the discourse unfolds. The traditional term *anaphora* is used when a phrase (typically a short, high-frequency *pronominal*) refers back to a previously mentioned entity by pairing up with another phrase (its *antecedent*). The antecedent's meaning determines that of the dependent pronominal—they *co*-refer to the same referent. Pronominals divide into two major subgroups: reflexives (*herself, himself,*

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

*itself, themselves*) and pronouns (*she/her, he/him, it, they/them*). The use of such words reduces redundancy and speaker effort (compare *The sister of the guy you met at the party talks about him a lot* and *The sister of the guy you met at the party talks about the guy you met at the party a lot),* but can create ambiguity that makes the listener or reader's task more complex.

## *Anaphor agreement*

Both pronominal subtypes express certain features of their antecedent, including person (first, second, third), number (singular versus plural), gender (masculine/feminine/neuter), and humanness *(he/she* versus *it*). The basic agreement effect with both reflexives and pronouns is that they must match their antecedents in these features. In addition, reflexives always pair with an antecedent in the same sentence, and have very precise limits on the position in which the antecedent may occur. For these reasons, they present the most interesting case to study from the point of view of asking how many of the grammatical constraints are deployed by the online processing systems, and at what point. So, reflexives are the primary focus of the discussion below.

For a speaker, choosing a reflexive is a way of unambiguously signaling that coreference is intended between two argument positions of the same clause. The hearer's task is to reconstruct the intended message. Recognizing the anaphor and linking it to an antecedent is a crucial part of that process, and in turn agreement is central to that search.

The mini-discourse in (1) introduces three referents: a dentist, a doctor, and a ballerina.

(1) The dentist said the doctor had carefully described the ballerina to himself

Who is the doctor describing the ballerina to? Not the ballerina, since *himself* and *the ballerina* are mismatched in gender. (And not *the dentist* in the higher clause, due to the clause-bounded nature of reflexive-antecedent pairs; see below). This leaves just *the doctor*, a noun unmarked for gender (and within the same clause) and therefore a grammatical antecedent for the masculine reflexive. Thus agreement (as well as locality) can be a disambiguator in the search for an antecedent.

The question arises as to what the sentence processing system does in such cases: does it temporarily consider all these NPs as possible antecedents? Or is the parser so narrowly guided by grammatical structure and principles that only the right NP is ever entertained?

Note that in (2), the reflexive is feature-congruent with both *the doctor* and *the ballerina*, leading to ambiguity:

(2) The dentist said the doctor had carefully described the ballerina to herself

When the antecedent is human and unspecified for gender (*doctor*, *customer, chemist*, *electrician*) speakers will choose whichever pronominal gender they feel is appropriate[1] (e.g., *herself or himself* for *doctor* above). But the listener often will not know the intended gender. Thus when a third-person singular reflexive is encountered, two sorts of antecedent phrases might have to be accessed—gender-matching ones, and gender-unspecified ones.

## *Pronoun distribution and agreement*

When we replace the reflexive in (2) with a pronoun (3), the antecedent possibilities change.

(3) The dentist said that the doctor had carefully described the ballerina to her

Here only *the dentist* can serve as antecedent of the pronoun. This reflects a much-studied division of labor between reflexive and pronoun forms: pronouns may take antecedents in precisely those positions where a reflexive can't (Chomsky, 1981, 1986; Reinhart, 2006). When a listener or reader encounters a pronoun, the search for an antecedent must cast its net wide: the antecedent can be at any distance, and in any position (other than those where a reflexive's antecedent could appear).

Reflexives universally require a linguistic antecedent (one cannot point to a person and say **I like himself*[2]), while pronouns often have antecedents (*John said he had to leave)*, but do not require them (one can point and say *I like him*). The location of the antecedent with respect to the reflexive is very highly constrained. Roughly speaking, reflexives require an antecedent that is a higher argument in the same clause. Typically the reflexive will be the object or indirect object of a verb, and the antecedent will be that verb's subject or object.[3] This description wraps together two properties of reflexive-antecedent pairs that have been discussed extensively in the linguistic literature: *locality* (the same-clause requirement) and *hierarchical superiority* (the antecedent must be higher than, or *c-command,* the reflexive; see Chomsky, 1981).

Consider the following sentence, which also appears in Figure 15.1:

(4) The spokeswoman said that the ballerina with the injuries was blaming herself for the accident.

The syntactic restrictions described above are illustrated in Figure 15.1 in the following way: *herself* is c-commanded by *the spokeswoman* and *the ballerina*, but not *the injuries*. But *the spokeswoman* cannot be the antecedent, due to locality.

So, two grammatical regulators—featural agreement and structural position—might come into play in real-time comprehension whenever a reflexive or pronoun is encountered. And they could do so from the beginning, by limiting the considered options to only the grammatically possible antecedents; or only play a role in later stages of processing, filtering down the entire larger set of previously encountered NPs to the grammatically possible ones.

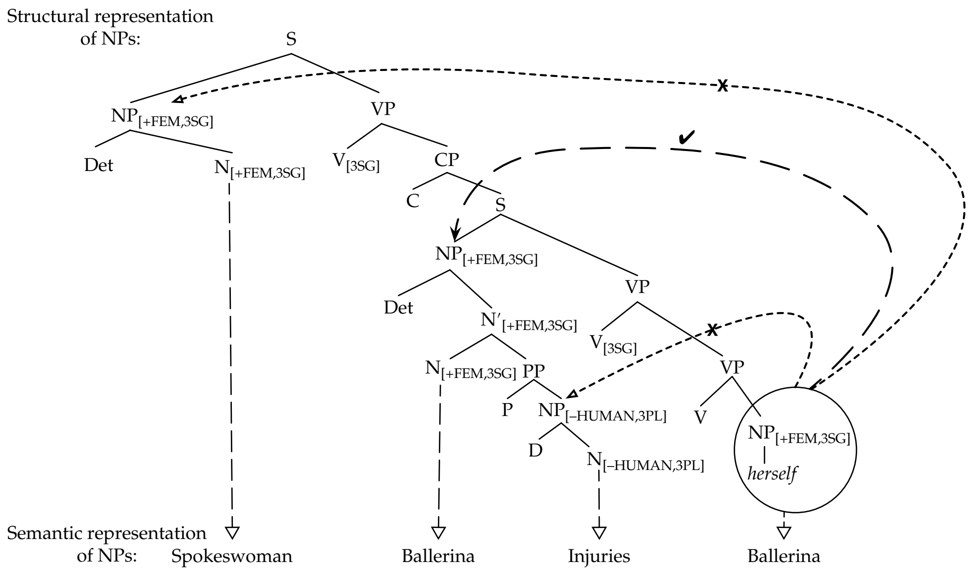



**Figure 15.1** Schematic representation of coreference processing, in which a proform initiates a process whereby the syntactic information (both structural and featural) associated with previously-mentioned NPs is used to eliminate inappropriate NPs (note the X in the dotted arrows to higher and non-c-commanding NPs). Once permissible NPs are identified as possible antecedents, their semantics may be retrieved. Here the dashed arrow indicates the link between the reflexive and the syntactically appropriate antecedent.

## *Subject-verb agreement*

English is a language with fixed word order, in which the subject NP always precedes the verb. NPs may be complex: they may contain other NP**s**, as in example (5).

(5) [$_{NP}$ The ballerina with [$_{NP}$ the injuries]] was …

The head of an NP is the unique noun around which the NP is built. In (5), the noun *ballerina* is the head of *the ballerina with the injuries*: it determines the overall properties of the NP, including the number, person, gender, and humannness specification of the NP, relevant to agreement.

Subject-Verb Agreement (SVA) is the appearance on a verb, in some tense-person-number combinations, of morphology mirroring the subject NP's agreement features. English marks subject-verb agreement only in indicative clauses, and only then partially and in two subcases: present tense, in which there is agreement on main verbs and on auxiliary *have* for third-person singular, expressed by *-s* or the form *has*; and the single verb *be*, which has three distinct forms in the present tense (*am*, *was*, and *are*) and two in the past tense (*was/were*). There is no agreement in indicatives when a modal is present (*I/you/he/they can do that easily*). And there is no agreement in non-indicatives, including small clauses (*The spectators are watching [John dance(*s)]*), infinitivals (*I arranged for [John to dance(*s)]*), and subjunctives (*The rules require that [John dance(*s)]*).

In English, the post-subject region up to the main verb can contain a variety of verbal heads (Chomsky, 1957), all optional, as in (6). Only the leftmost (and thus highest) of these may bear agreement (7):

(6) (modal) (have) (be) main verb
(7) a. She/they might have been dancing (no agreement due to modal)
b. She/they has/have been dancing
c. She/they is/are dancing
d. She/they dances/dance

A central argument in generative grammar is that agreement is between the verbal head and the structural head noun of the subject, that this is a local relation, and that the notion of locality is hierarchical, not linear (Berwick, Pietroski, Yankama, & Chomsky, 2011; Chomsky, 1957; Lasnik, 2000). The noun controlling verb agreement can be arbitrarily far away from the verb, with multiple linearly intervening nouns:

(8) The man near the judges who came from foreign countries is/*are dancing tonight

Once again the question arises as to how this static fact about grammatical patterns is utilized in comprehension. Do the intervening nouns matter? Or is the comprehension system precise in keeping track of the features of the head noun?

We end this section by considering some more intricate cases: collectives (*cast, government, group*), collective partitives (*a lot/number/bunch of*), and reference shift (*The ham sandwich at Table 5 is getting impatient*). Each case may complicate both producing and comprehending agreement.

Consider the common type of vivid example, *My neighbor was hit on the left front fender*, where the subject refers not to my neighbor but to his car. The NP's reference shifts to something associated with the referent. Note that both verb agreement and pronoun agreement in such sentences can be controlled by the [singular, human] conceptual features of the shifted referent, overriding the [plural, non-human] grammatical features of the subject NP (e.g., 9 & 10).

(9) The French fries is getting impatient and wants his bill right away (adapted from Nunberg, 1995)
(10) The ham sandwich left without paying his bill because it was inedible (Fauconnier, 1994)

A subtler version of this is one way to understand the unstable agreement patterns of collective nouns (*committee, government, team, class, group*) that denote single entities made up of multiple individuals. Collectives are used either to refer to the entity as a single unit, or to refer to its multiple member individuals. There is an inherent singular-plural duality at the conceptual level, which may be a systematic use of reference shift, from the collective (11) to its members (12).

(11) The team is entering the competition.
(12) The team are complaining about their accommodations.

This shift seems to be grammaticalized for many speakers in the case of collective partitives like *a lot/bunch/number of people are coming to the party*.

## Experimental research

We assume that during sentence parsing, the phrase structure of the input sentence is computed incrementally. As each word appears, lexical processes access information about grammatical category, which is then incorporated into the syntactic structure under construction, in accordance with phrase structure rules of the language. One research question addressed in the relevant literature is whether during real-time sentence processing the computation of the two agreement relations under consideration here is *constrained* (and if so, when) by phrase structure information. Another way to think of this is: does the sentence processor obey the constraints described above?

We first review some of the research on the processing of coreference between a reflexive and an antecedent, and then turn to the research on the computation of subject-verb agreement.

### *The comprehension of reflexives*

Various experimental techniques have been used to explore the processing of reflexives, including priming, reading times, and ERP. Priming paradigms have been used to probe the activation of word meanings as different words are encountered in a sentence context. Reading times and ERP have been used to examine the effects of non-antecedents on the coreference process. These are discussed in greater detail below.

In one priming technique—cross-modal priming—a sentence is presented auditorily, and at a critical point in the sentence, a probe is presented visually on a computer screen. The probe is either a real word or nonword, and participants are asked to decide if the probe is a word or not. (Response times to press a button indicating the participant's decision are recorded.) This technique was used initially to explore the activation of meanings of semantically ambiguous words like *bank* or *bug* (Swinney, 1979). A semantically ambiguous word would appear in a sentence context that might favor the dominant or subordinate meaning, and a probe item would appear at the offset of the ambiguous word (e.g., *The man was not surprised when he found several spiders, roaches, and other* bugs *in the corner of the room*.) This probe was either semantically related to the ambiguous word (e.g., *ANT or SPY* in the case of *bugs*), or was unrelated (e.g., *SEW*), but matched to the related probe(s) with respect to features like word length and frequency (variables that are well-known to affect response times in a decision task; note that nonword probes would be paired with filler sentences). Faster response times to semantically related probes (compared to unrelated probes)—"priming"—would be taken to indicate that the related word meaning was active in the minds of the listeners.

Nicol and Swinney (2003) describe the results of an experiment that used cross-modal priming to examine the processing of reflexives in sentences like (13):

(13) The queen told the ballerina that the actress in the play had injured herself during the performance.

At the offset of *herself*, either a word would appear that was either related to *queen*, *ballerina*, or *actress*, or was an unrelated control word. They assumed that if *actress* was active, faster response times should be observed for a related word like *script* than an unrelated word like *format*. Likewise, they assumed that *ballerina* was not active if response times to the related word *dance* were equal to response times to the unrelated word *learn*. Their results showed priming only for the structurally appropriate NP: *the actress*. Further, in sentences containing a pronoun instead of a reflexive, the reverse pattern was observed: there was priming for *queen* and *ballerina* but not for *actress*. Other pronoun experiments (Nicol & Swinney, 2003) that manipulated the genders of the sentence participants and pronouns (e.g., *The skier told the ballerina that the doctor for the team would blame her/him for the recent injury*) found priming only for *non-mismatching* antecedents: gender-neutral nouns such as *skier* and *doctor* were reactivated by a pronoun such as *her* or *him*, but gender-marked nouns such as *ballerina* were not reactivated by *him*. Manipulations of number also showed selective priming: number-mismatching antecedents were not reactivated. Other work, involving manipulations of animacy, showed a similar effect (Shillcock, 1982). Nicol and Swinney argued that the appearance of a pronominal element such as a reflexive or pronoun triggers the reactivation of potential antecedents, but membership in the set of potential antecedents is constrained by pronominal feature congruence and by their position within a sentence structure. They proposed that during parsing, a representation is computed in which NPs are encoded along with information about their position within a syntactic structure, and their pronominal features (e.g., [+feminine] [+plural]). It is this additional information that the coreference process pays attention to: only the meanings of NPs with the appropriate structural position and pronominal feature specifications will be reactivated. If more than one potential antecedent is included in the candidate set, other processes (that consider real-world information and discourse context) need to be brought into play, just as they would come into play to inhibit the irrelevant meaning of a semantically ambiguous word.

It is important to note that although Nicol and Swinney conceptualized the coreference process as involving a search within a tree-structure, tree-structure information can be translated into a series of descriptions. Thus, each NP could be stored in memory along with descriptive information about its structural position, and its pronominal features. Either way, it is this cluster of information that is considered prior to the reactivation of word meaning.

This proposal was challenged by Badecker and Straub (2002), who conducted a self-paced reading study. In this paradigm, parts of a sentence are presented sequentially, one segment at a time. Each segment is either a word, or a larger multi-word chunk. The participant controls the presentation rate via the press of a button. Reading times for each chunk are recorded. Typically, a *moving-window display* is used: as each segment appears, the previous one disappears. Badecker and Straub found an effect on reading times of a potential antecedent that appeared in the wrong structural position to serve as the antecedent of a reflexive. They tested sentences such as *Bill/Beth thought that John owed himself another opportunity to solve the problem*. Because neither *Bill* nor *Beth* can be the antecedent of the downstream reflexive, it should not matter that one of them matches the gender of the reflexive and the other does not. But their results show that it does: reading times for sentences with matrix subject NPs that match the reflexive in gender are read more slowly. This was interpreted to mean that the gender-matched NP is part of the candidate set of antecedents; with two potential antecedents (e.g., *Bill* and *John*), one must be eliminated, and this takes time. (Note that all that needed to be shown was a difference between the *Bill* and *Beth* conditions. A reading time difference could have gone in the other direction, with longer reading times for the incongruent sentences. An explanation for such an effect could be that *Beth* is initially part of the candidate set, is evaluated as a possible antecedent of *himself*, and the mismatch creates processing difficulty.)

Other research, however, supports the Nicol and Swinney (2003) claim. For example, Sturt (2003), used eye-tracking to measure reading times. Unlike self-paced reading, eye-tracking measures fixation times as participants read a sentence. There may be several fixations on a word during the initial reading of that word, and in addition, the word may be re-read after a later portion of the sentence is read. Hence, eye-tracking offers multiple measures of reading time: "early" measures, including the first fixation times as a participant reads from left to right, and additional times that arise during re-reading. Sturt examined reading times for regions of text in short discourses such as the following:

(14) a. Jonathan was pretty worried at the City Hospital. He remembered that the surgeon had pricked himself with a used syringe needle. There should be an investigation soon.
b. Jennifer was pretty worried at the City Hospital. She remembered that the surgeon had pricked himself with a used syringe needle. There should be an investigation soon.
c. Jonathan was pretty worried at the City Hospital. He remembered that the surgeon had pricked herself with a used syringe needle. There should be an investigation soon.

d. Jennifer was pretty worried at the City Hospital. She remembered that the surgeon had pricked herself with a used syringe needle. There should be an investigation soon.

In this set of materials, the embedded clause in the second sentence contained as its subject a character with a stereotypical gender. In this example, *surgeon* is stereotypically male. The reflexive (*himself* in (14a) and (14b) and *herself* in (14c) and (14d)) corefers with this embedded subject. NPs outside of this embedded clause (*Jonathan/Jennifer, he/she*) cannot corefer with the reflexive.

The second set of materials is similar to the first, except that the non-antecedent intervenes between the subject of the second clause and the reflexive.

(15) a. Jonathan was pretty worried at the City Hospital. The surgeon who treated Jonathan had pricked himself/herself with a used syringe needle. There should be an investigation soon.

b. Jennifer was pretty worried at the City Hospital. The surgeon who treated Jennifer had pricked himself/herself with a used syringe needle. There should be an investigation soon.

If syntactic position information constrains the coreference process, then the gender of the non-antecedent should not matter. And this is what was found. Results for early measures of reading time showed only a main effect of gender-congruence between a reflexive and the subject of the clause in which it appears (*surgeon…herself* was read more slowly than *surgeon…himself*). The gender of the subject NP in the first sentence had no effect at all on initial reading times. This is clearly consistent with the Nicol and Swinney findings.

However, it did have an effect on re-reading times, in the first set of materials. Specifically, re-reading times were longer for a discourse such as *Jennifer…She remembered that the surgeon had pricked himself…* than for *Jonathan…He remembered that the surgeon had pricked himself…* (Note that this non-antecedent effect is different from Badecker and Straub's non-antecedent effect, in that here, faster reading times were associated with the condition in which there were two NPs that gender-matched the reflexive.) Considering the early and late measures together, Sturt proposed that the initial candidate set is constrained by coreference constraints, but that later interpretive processes may override this initial process, if, for example, a non-antecedent has discourse prominence, as it does in the sentences in (14) (in which Jonathan or Jennifer is the referent of the subject in the second clause as well as the first). This idea was supported by the results of a follow-up study, which explicitly queried the participants about their interpretation of the reflexive. This study showed that when a reflexive matched the non-antecedent (e.g., *Jonathan/Jennifer*), participants selected that noun significantly more often than when the reflexive mismatched the non-antecedent.

Other eye-tracking studies have confirmed the finding that the coreference process is constrained by syntax. For example, Dillon *et al.* (2013) tested the comprehension of sentences such as (16), in which a non-antecedent intervenes between the reflexive and the true antecedent. This non-antecedent is either number congruent with the true antecedent, or not. Further, they tested both grammatical and ungrammatical versions (the latter to explore repair processes):

(16) a. The new executive who oversaw the middle manager apparently doubted himself/themselves on most major decisions.
b. The new executive who oversaw the middle managers apparently doubted himself/themselves on most major decisions.

They found no effect of the number specification of the non-antecedent *manager/managers* in either the grammatical or ungrammatical sentences.

Felser *et al.* (2009), examined the reading of short narratives that began with the sentence *John/Jane and Richard were very worried in the kitchen of the expensive restaurant*, and ended with the sentence *Kitchens can be dangerous places*. The middle sentence was one of the following:

(17) a. John/Jane noticed that Richard had cut himself with a very sharp knife.
b. It was clear to John/Jane that Richard had cut himself with a very sharp knife.

The difference between the (a) and (b) versions is that the non-antecedent appears in different structural positions. At issue is, again, whether the gender of the non-antecedent affects the processing of the reflexive. Results show no effect of the non-antecedent at all, at any point in the sentence.

However, such was not the case for second language (L2) learners. Felser *et al.* included a group of L2-English learners whose native language was Japanese (which allows some reflexives to corefer with a long-distance subject). For this group, the gender of the non-antecedent mattered: when the non-antecedent matched the reflexive's gender, they took longer to read the reflexive. It is possible that comprehension processes are simply less well established in L2, or that the L2 learners were misapplying their L1 coreference routines.

A subsequent eye-tracking experiment by Cunnings and Felser (2013) also showed variability in their experiment participants. They used materials that resembled those used by Sturt (2003), in which the antecedent of the reflexive has a stereotypical gender, and in which the non-antecedent initially appears in "focus" position in the introductory sentence. In addition, the distance between the antecedent and reflexive is manipulated in two sets of materials, as exemplified below. Note that in the second set of materials, there is an intervening non-antecedent.

(18) a. James/Helen has worked at the army hospital for years. He/She noticed that the soldier had wounded himself/herself while on duty in the Far East. Life must be difficult when you are in the army.
b. James/Helen has worked at the army hospital for years. The soldier that he/she treated on the ward wounded himself/herself while on duty in the Far East. Life must be difficult when you are in the army.

Participants were also given the Daneman and Carpenter (1980) reading span test, and divided into high and low working memory span groups. Results showed an effect of congruence between the stereotypical gender of the antecedent and the reflexive: incongruence produced slower reading. This was true for both sets of materials. But the results for the second set of materials showed that the low span group was affected by the congruence of the non-antecedent as well. Cunnings and Felser suggest that low span readers might have difficulty inhibiting a non-antecedent when it is a clausal subject and intervenes between the reflexive and antecedent. Because it is still active, it is considered as an antecedent of the reflexive.

In a similar vein, a pronoun study described by Nicol and Swinney (2003) shows variability in reading patterns tied to accuracy in answering comprehension questions. They describe a study by Nicol (1997), who used a self-paced reading paradigm to examine reading times for sentences like (19), in which the embedded subject is disallowed as the antecedent of a following pronoun:

(19) a. My aunt heard that the congresswoman/congressman would contact her about the complaint.
 b. My aunt heard that the congresswoman/congressman would contact me about the complaint.

Analysis of all items, whether properly understood or not, showed that in the region following the embedded object, there was a significant interaction between gender congruence (between the matrix and embedded subjects) and embedded object type (third person pronoun versus other): For a sentence such as (19a), reading times were slower when the embedded subject was *congresswoman* (versus *congressman*). However, when only the data for sentences that were understood correctly were analyzed, this effect disappeared. The gender of the non-antecedent had no effect.

Finally, we consider the results of event-related potential (ERP) research. This method involves presenting sentences either auditorily or visually (word-by-word with tempo under computer control) and recording electrophysiological activity via electrodes positioned on the scalp. Early work on the processing of visually-presented sentences containing reflexives indicated that a gender mismatch between a reflexive and its antecedent (e.g., *The woman congratulated himself…*) is associated with a *P600*, a relatively more positive-going wave peaking approximately 600 ms. after the appearance of the critical word (the reflexive in this case) (Osterhout & Mobley, 1995). In general, the P600 is associated with parsing failure or difficulty, and is distinct from the ERP associated with semantic difficulty. Later research by Xiang, Dillon, and Phillips (2009) explored ERP responses to visually presented sentences such as the following:

(20) a. The tough soldier that Fred treated in the military hospital introduced himself to all the nurses.
 b. The tough soldier that Katie treated in the military hospital introduced herself to all the nurses.

c. The tough soldier that Fred treated in the military hospital introduced herself to all the nurses.

In all cases, there is a non-antecedent that intervenes between the reflexive and the true antecedent (*the soldie*r). In both (20b) and (20c), the true antecedent of *herself* is incongruent from the point of view of stereotypicality: In this example, American participants are likely to think that soldiers are more often male than female. But (20b) and (20c) differ with respect to the gender of the non-antecedent, with the (20b) sentence offering a potentially misleading non-antecedent that matches the gender of the reflexive. Results showed that coreference between *herself* and *soldier* is difficult for participants: A P600 effect was observed for the conditions represented in (20b) and (20c). But there was no difference between them. The gender of the non-antecedent had no effect on how the reflexive was processed.

Overall, then, the research suggests that non-antecedents do not affect processing of the reflexive, and in general, they do not affect the processing of subsequent regions of text. But the coreference process may sometimes break down, especially for readers who have low memory capacity or difficulty inhibiting irrelevant information, and may operate differently in second-language learners.

## *The comprehension of subject-verb agreement*

Let us now turn to the comprehension of subject-verb agreement. Much of the research in this area has been relatively recent, and focused on the question of whether agreement is computed in a similar fashion during language production and comprehension. During the 1990s, there was a sudden increase in language-elicitation studies involving a *repeat-and-complete* task in which a sentence beginning was presented to a participant, who repeated it and completed the sentence. Bock and Miller (1991) had shown that speakers produce verb agreement errors at a surprisingly high rate (sometimes as high as 25%), in contexts in which there is an intervening number-marked NP, as in (21):

(21) The key to the cabinets are rusty.

As described above, subject-verb agreement in English is relatively straightforward. A verb must agree in number (and person) with the head NP within the subject NP. Although most of the time, speakers get this right, some portion of the time, they do not, and produce sentences such as (21). The production data also indicated an asymmetry with respect to number marking: a sentence beginning like *The keys to the cabinet* did not elicit significantly more verb agreement errors than *The keys to the cabinets*. This asymmetry was attributed to the markedness of the plural (Bock & Eberhard, 1993). Roughly, a speaker may make note of the plurality of an NP—even if it is not the head of the subject phrase—and make the ensuing verb agree. This has been formalized by characterizing the feature [+singular] as an unmarked default, and so an absence of marking would be construed as singular. Plurals would be explicitly marked as [+plural].

Several explanations for the verb errors were proposed in which the verb acquired number marking via a *copying operation* within a tree structure: the number feature on the head NP is copied to the verb (or an inflectional node). An error may arise due to the following: (1) The [+plural] number feature of the nonhead erroneously migrates "upward" within a hierarchical structure, and is then copied to the verb (Vigliocco & Nicol, 1998)); (2) The encoding process misconstrues the subject phrase as plural (Eberhard *et al.*, 2005). This could happen with subject phrases such as *The label on the bottles…,* since, despite the grammatical singularity of *label*, semantically, there are multiple physical tokens of the label, one on each bottle. (For additional discussion of agreement operations in language production, see Franck, Chapter 2, this volume.)

The research on agreement in production led to research examining whether there were parallel effects in comprehension. On the face of it, it seemed that there might not be because number-marking on a verb contributes nothing to the computation of phrase structure, and not much if anything to the overall meaning of the sentence. But if there were, this could indicate that a common mechanism is at play, either in the computation of syntactic structure or processing of agreement, or both.

The initial studies did show a similarity to production. For example, Nicol, Forster and Veres (1997) tested sentences such as those in (22) in two tasks. In the first, a "maze" task, sentences were presented word by word, with each word paired with another word that did not represent a good continuation of the sentence, and participants chose (with a button-press) the word that offered a good continuation. The second task presented sentences as a whole and required a judgment about whether the words were correctly ordered. Ungrammatical agreement sentences were not presented, and the tasks did not focus on agreement (ungrammatical filler sentences containing blatant word order violations were included).

(22) a. The author of the speech is here now.
b. The author of the speeches is here now.
c. The authors of the speeches are here now.
d. The authors of the speech are here now.

Both tasks showed the same asymmetry: slower processing times for (b) than (a), but no differences between (c) and (d), exactly analogous to the error data in production. This supports the idea that even in comprehension, the marked [+plural] number feature of a nonhead NP might sometimes be unstable, and erroneously copied to the verb, creating a feature clash and processing slowdown.

Similar results were found by Thornton and MacDonald (2003), who used a self-paced reading methodology, and by Pearlmutter, Garnsey, and Bock (1999). The latter used both self-paced reading and eye-tracking. In their self-paced reading study, they looked at similar constructions to those above, and also found a "mismatch asymmetry": the mismatch in noun number caused processing difficulty only when the head noun was singular. In two additional experiments, they raised the issue of how agreement in *un*grammatical sentences is processed, comparing singular-head sentences followed by a grammatical versus ungrammatical verb:

(23) a. The key to the cabinet was rusty from many years of disuse.
b. The key to the cabinets was rusty from many years of disuse.
c. *The key to the cabinet were rusty from many years of disuse.
d. *The key to the cabinets were rusty from many years of disuse.

This manipulation allowed them to determine whether any difficulty posed by the plural nonhead in a fully grammatical sentence was akin to the difficulty of encountering a verb that was ungrammatical. One experiment used self-paced reading, another used eye-tracking. Pearlmutter and colleagues' self-paced reading study showed that, relative to the control sentence (a): reading times were slower at the verb when the non-head was plural (b,d), and slower at subsequent words when the verb was ungrammatical (c,d). But in addition, there was a grammaticality by nonhead-number interaction after the verb: in sentences in which the nonhead was plural, reading times were slower in the grammatical case (b) and faster in the ungrammatical case (d) (compared to the singular-singular controls).

Their eye-tracking results were more complicated but overall they were consistent with the pattern seen in self-paced reading. Gaze data further showed that on early reading time measures the (a) sentences were read more quickly than each of the other three sentence types in the relevant sentence regions. This means that the mismatch effect for grammatical sentences showed up in the initial reading times. But the mismatch effect for ungrammatical sentences did not. Further, in later reading measures, the difference in reading times for the two types of grammatical sentence diminished over the course of the sentence, while the difference in reading times for the two ungrammatical sentences increased, with the type of sentence in (c) creating the greatest difficulty. Given the difference in the timing and direction of the effects, Pearlmutter *et al.* conclude that the number mismatch effect in grammatical sentences is distinct from the number mismatch effect in ungrammatical sentences.

But the number mismatch effect in grammatical sentences (Nicol *et al.*, 1997; Pearlmutter *et al.*, 1999)—which had pointed to a common agreement mechanism in production and comprehension—was not observed in later studies, such as those described next.

In a series of self-paced reading studies, Wagers *et al.* (2009) examined the processing of sentences containing mismatching head and nonhead NPs, including materials like those in (23), with an adverb was inserted before the verb (e.g., *The key to the cabinets unsurprisingly was rusty from many years of disuse*). They included adverbs in order to avoid potential spillover effects from the nonhead NP (e.g., *cabinets*) to the adjacent verb, which, they argue, could be why the previous studies had found a mismatch effect at the verb in grammatical sentences. Unlike Pearlmutter *et al.* (1999), they found no reading time differences in the region of the verb in the grammatical sentences. But consistent with Pearlmutter *et al.* (1999), they did find a mismatch effect in their ungrammatical sentences, with shorter reading times when the nonhead was plural (and therefore agreed in number with the verb) than when the nonhead was singular. Further, they tested sentences containing a mismatching nonhead in a sentence configuration in which it appeared before the head: *the musicians that the reviewer praise so highly…* and again found a nonhead effect, and again only in the ungrammatical sentences. They argue that these results are incompatible with the explanations described above because those have to do with instability of the number feature within a structural representation, or uncertainty concerning number marking within the subject NP and do not make reference to the grammaticality of the verb. This uncertainty should result in processing slowdown in the case of correctly inflected verbs and processing facilitation in the case of incorrectly inflected verbs. But the Wagers *et al.* findings did not show this.

They explain their results by appealing to a cue-based retrieval process that is triggered by the appearance of the inflected verb. The notion of retrieval (as used here) is based on the obviously true premise that as the words in a sentence are understood, they are stored in some kind of memory buffer so that they can be integrated with later-occurring words as the meaning of the whole sentence is computed (for more details about these notions as they relate to agreement and coreference, see Badecker & Kuminiak, 2007; Dillon, Mishler, Sloggett, & Phillips, 2013; Martin, Nieuwland, & Carreiras, 2012; Vasishth, Brüssow, Lewis, & Drenhaus, 2008). Features of the verb are used to search through working memory for the appropriate controlling NP. Wagers *et al.* (2009) outline two ways that this could work. One is completely "backward-looking": with the appearance of an inflected verb, a search is initiated. The other is backward-looking, but also has a predictive component: the number feature on the subject will lead to an expectation for the verb to match (or presumably, in the case of uninflected verbs, at least not mismatch). Only if there is a mismatch will there be a backward search, as part of a repair strategy. (Incidentally, it is not entirely clear what kind of entity is retrieved. In their discussion, Wagers *et al.* refer to the retrieval of a number feature and to the retrieval of an NP.)

The cues that are used for retrieval include grammatical number specification (e.g., plural) and role (e.g., subject). In mismatch-grammatical sentences such as *The key to the cabinets was rusty…* the appearance of the verb would lead to retrieval of the number-matching subject NP, and it would find only [$_{NP}$ *the key to the cabinets].* The plural NP contained within the subject, [$_{NP}$ *the cabinets]* would not be retrieved because it does not match any of the retrieval cues (number and grammatical role).

Now consider the ungrammatical cases. In the match-ungrammatical sentences like *The key to the cabinet were rusty*, *the key* meets the *subject* criterion but not the *plural* criterion, and *the cabinet* meets neither criterion. In the mismatch-ungrammatical sentences like *The key to the cabinets were rusty…* the appearance of the plural verb would lead to the retrieval of both *the key* (which meets the *subject* criterion), and *the cabinet*s, which meets the *plural* criterion. Wagers *et al.* suggest that "when neither of the NPs matches the combined cue, as in the ungrammatical sentences, the number-matching non-subject is sometimes the best match" (p. 28). They propose that this is why the mismatch versions are easier to process than the match versions.

Does this pattern of results replicate with other methodologies? It does. In a recent eye-tracking study. Dillon *et al.* (2013) examined the reading of sentences like (24), which contain an adverb before the critical region, the verb.

(24) a. The new executive who oversaw the middle manager apparently was/*were dishonest.
b. The new executive who oversaw the middle managers apparently was/*were dishonest.

They found the same thing: the presence of a plural nonhead has no effect on the processing of the verb in grammatical sentences, but it eases processing of the verb in ungrammatical sentences.

Tanner, Nicol, and Brehm (2014) report similar findings from ERPs to visually-presented sentences. They tested sentences like the following:

(25) a. The chemist with the test tube is conducting an experiment.
b. *The chemist with the test tube are conducting an experiment.
c. The chemist with the test tubes is conducting an experiment.
d. *The chemist with the test tubes are conducting an experiment.

Previous research showed that subject-verb agreement violations in English elicit the P600 effect (see Molinaro, Barber, & Carreiras, 2011, for a review). As mentioned earlier, the P600 may also be associated with processing difficulty (e.g., Kaan, 2002; Kaan & Swaab, 2003; Kaan *et al.*, 2000; Nevins *et al.*, 2007; Osterhout, Holcomb, & Swinney, 1994; Xiang, Dillon, & Phillips, 2009). Thus, it was predicted that there would be greater positivity at the verb in (b) than (a) and in (d) than (c), due to the ungrammaticality of (b) and (d). In addition, based on the Wager *et al.* results, it was predicted that there would be a smaller P600 effect for the ungrammatical mismatch condition, (d), than for the ungrammatical match condition, (b). It was also predicted that there would be no P600 effect for (c) versus (a) because although (c) contains a number mismatch, it also contains a grammatical verb.

Tanner *et al.* (2014) found a P600 effect in the ungrammatical sentences compared to their grammatical counterparts, with a smaller effect in mismatch sentences such as (25d). But the comparison of (25c) with (25a) showed no positivity. (This partly replicates some findings by Shen, Staub, and Sanders, 2013, who presented similar sentences auditorily.) The judgment task showed no difference in accuracy for the two types of grammatical sentence, but for the two types of ungrammatical sentence, there were significantly more errors for ungrammatical mismatch sentences than ungrammatical match sentences.

A second experiment was conducted in order to test the idea that the processing difficulty caused by a mismatch effect in the grammatical sentences might take time to emerge. The materials that were used were identical to those above, except for the addition of an adverb before the verb. This adverb was intended to create some distance between the nonhead and the verb. The results of this experiment were similar to the first one: the ERP data showed effects of ungrammaticality, with a smaller effect in the mismatch condition, and there was no effect of a mismatch in the grammatical sentences. The judgment task, however, did show a mismatch effect for grammatical sentences (as well as for ungrammatical sentences), with more errors for sentences like *The chemist with the test tubes probably is conducting an experiment* than for their number-matching counterparts. The authors suggest that this result could be due to the fact that preverbal placement of an adverb is relatively infrequent in English, lowering the overall number of "grammatical" judgments, and interacting somehow with the singular head-plural nonhead configuration.

The failure to find a mismatch effect on (or after) the verb in grammatical sentences is obviously problematic for the view that the computation of agreement during sentence comprehension involves the same mechanisms as in production. The explanation for the occurrence of subject-verb agreement production errors is that, within a hierarchical structure, the number of the nonhead feature "slips" and is erroneously copied to the verb, or that the entire subject NP is mistakenly valued as plural. If these processes also occur during comprehension, there should be observable mismatch effects within the following VP, whether or not the verb is grammatical.

In sum, the preponderance of evidence suggests that the computation of grammatical verb agreement during comprehension does not typically consider nonheads as controllers of agreement.

## Conclusion

Overall, it appears that in fully grammatical sentences—by far the norm—the processing of both reflexives and inflected verbs is carried out without considering nouns that appear in a syntactically inappropriate position within the sentence. In this sense, the two processes appear to be similar. And both processes precisely follow the grammatical constraints reviewed earlier, supporting the view that the parser tightly incorporates the grammar.

Differences emerge when there is an agreement violation. In the reanalysis process, a mismatching nonhead that agrees with the verb may offer a way to make things right, but a mismatching antecedent that agrees with a reflexive apparently does not. One reason for the difference could have to do with the frequency with which the comprehension system must deal with the different types of agreement violations. Inflected verbs are far more common than reflexives: counts of word frequency show that even if only the verbs *is/was* and *are/were* are considered (a subset of all inflected verbs), they are at least thirty times more frequent than *herself/himself* and *themselves* (the ratios are the same for the singular and plural forms). And it is not as though reflexives are more prone than verbs to error: studies of elicited production show similar rates of erroneous plural productions (about thirteen percent) when participants are asked to complete a sentence beginning such as *The actor in the soap operas…* with a verb and a sentence beginning like *The actor in the soap operas watched…* with a reflexive (Bock, Nicol, & Cutting, 1999). But if, in spontaneous speech situations, speakers produce relatively few utterances that call for a reflexive, the frequency with which comprehenders encounter reflexive errors is likely to be very very small. This could lead the comprehension system to develop different repair strategies for the two types of agreement violation.

## NOTES

1 If the gender of the referent is unknown to the speaker, and/or the NP is intended to be understood generically, as in *any customer expects that you will serve ___ promptly*) no pronoun fits perfectly – there is a gap in the English pronoun paradigm. Various work-arounds exist, including *he or she*; *he/she*; the prescriptive *he*; and the increasingly popular plural form *they* (which, as a consequence of its increasing usage, may be losing its marking as exclusively plural).

2 We here use the linguist's asterisk, which is prepended to a sentence to indicate ungrammaticality.

3 However, in a number of configurations, the reflexive will be one clause down, an argument of a lower predicate, as in(i)-(iii), where complement clauses are bracketed:
   i. John considers [himself (to be) handsome]
   ii. Mary arranged for [herself to win the election]
   iii. The new reporter watched [himself interview the senator on TV]

## REFERENCES


Badecker, W. & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language*, 56, 65–85. doi: 10.1016/j.jml.2006.08.004

Badecker, W. & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 748–769. doi: 10.1037/0278-7393.28.4.748

Berwick, R. C., Pietroski, P., Yankama, B., & Chomsky, N. (2011). Poverty of the stimulus revisited. *Cognitive Science*, 35(7), 1207–1242. doi: 10.1111/j.1551-6709.2011.01189.x

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, 8(1), 57–99. doi: 10.1080/01690969308406949

Bock, K., & Miller, C. A. (1991). Broken Agreement. *Cognitive Psychology*, 23, 45–97. doi: 10.1016/0010-0285(91)90003-7

Bock, K., Nicol, J., & Cutting, J. C. (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language*, 40(3), 330–346. doi: 10.1006/jmla.1998.2616

Chomsky, N. (1957). *Syntactic structures*. The Hague.

Chomsky, N. (1981). *Lectures on government and binding*. Foris, Dordrecht.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Greenwood Publishing Group

Cunnings, I. & C. Felser. (2013). *The role of working memory in the processing of reflexives.* Language and Cognitive Processes, 188–219. doi: 10.1080/01690965.2010.548391

Daneman M, Carpenter PA. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, 69, 85–103. doi: 10.1016/j.jml.2013.04.003

Eberhard, K., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. Psychological Review, 112, 531–558. doi: 10.1037/0033-295X.112.3.531

Fauconnier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge University Press. doi: 10.1017/cbo9780511624582

Felser, C., Sato, M., & Bertenshaw, N. (2009). The on-line application of Binding Principle A in English as a second language. *Bilingualism: Language and Cognition*, 12*(04),* 485–502. doi: 10.1017/S1366728909990228

Kaan, E. (2002). Investigating the effects of distance and number interference in processing subject-verb dependencies: An ERP study. *Journal of Psycholinguistic Research*, 31, 165–193. doi: 10.1023/A:1014978917769

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15, 159–201. doi: 10.1080/016909600386084

Kaan, E., & Swaab, T. Y. (2003). Repair, revision and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15, 98–110. doi: 10.1162/089892903321107855

Lasnik, H. (2000). *Syntactic structures revisited: Contemporary lectures on classic transformational theory*. MIT Press.

Martin, A. E., Nieuwland, M. S., & Carreiras, M. (2012). Event-related brain potentials index cue-based retrieval interference during sentence comprehension. *NeuroImage*, 59, 1859–1869. doi:10.1016/j.neuroimage.2011.08.057

Molinaro, N., Barber, H., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, 47, 908–930. doi: 10.1016/j.cortex.2011.02.019

Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research*, 1164, 81–94. doi: 10.1016/j.brainres.2007.05.058

Nicol, J. (1997). *Processing pronouns in text: Effects of feature congruence of potential antecedents*. Unpublished manuscript, University of Arizona, Tucson, AZ.

Nicol, J., Forster, K., & Veres, C. (1997). Subject-Verb Agreement Processes in Comprehension. *Journal of Memory and Language*, 36, 569–587. doi: 10.1006/jmla.1996.2497

Nicol, J. L., & Swinney, D. A. (2003). The psycholinguistics of anaphora. In A. Barss (Ed.) *Anaphora: A reference guide* (pp. 72–104). John Wiley & Sons. doi: 10.1002/9780470755594.ch3

Nunberg, G. (1995). Transfers of meaning. *Journal of semantics*, 12(2), 109–132. doi: 10.1093/jos/12.2.109

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 786–803. doi: 10.1037/0278-7393.20.4.786

Osterhout, L., & Mobley, L. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773. doi: 10.1006/jmla.1995.1033

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427–456. doi: 10.1006/jmla.1999.2653

Reinhart, T. (2006). *Interface strategies*. Cambridge, MA: MIT Press.

Shen, E. Y., Staub, A., & Sanders, L. D. (2013). Event-related brain potential evidence that local nouns affect subject – verb agreement processing. *Language and Cognitive Processes*, 28, 498–524. doi: 10.1080/01690965.2011.650900

Shillcock, R. (1982). The on-line resolution of pronominal anaphora. *Language and Speech*, 25(4), 385–401.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48, 542–562. doi: 10.1016/S0749-596X(02)00536-3

Swinney, D. A. (1979). Lexical access during sentence comprehension:(Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*,18(6), 645–659. doi: 10.1016/S0022-5371(79)90355-4

Tanner, D., Nicol, J., & Brehm, L. (2014) The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language*, 72, 195–215. doi: 10.1016/j.jml.2014.07.003

Thornton, R. & MacDonald, M.C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language*, 48, 740–759. doi: 10.1016/S0749-596X(03)00003-2

Vasishth, S., Brüssow, S., Lewis, R. L., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science*, 32(4), 685–712. doi: 10.1080/03640210802066865

Vigliocco, G., & Nicol, J. L. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition*, 68, B13–B29. doi: 10.1016/S0010-0277(98)00041-9

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61, 206–237. doi: 10.1016/j.jml.2009.04.002

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, 108, 40–55. doi: 10.1016/j.bandl.2008.10.002

# 16 Prosody in Sentence Processing

ELIZABETH PRATT

The Graduate Center, City University of New York

## Introduction

*Prosody* is a broad term used to refer to the rhythm and intonation of spoken sentences. The study of prosody concerns the features of speech occurring above the phoneme or "segment" level, and how those features contribute to processing and interpretation. These *suprasegmental* features include measures of fundamental frequency ($F_0$), intensity, and duration, which combine to signal the organization of words into groups or phrases, or into more prominent and less prominent constituents.

Taken together, these features serve both linguistic and paralinguistic functions, indicating emotional expression, illocutionary force, new or contrastive focus, syntactic relations, and phrase structure. Beginning at the earliest stages of language development, prosody interacts with almost all levels of language processing. As part of the acquisition process, infants and children use prosodic information to identify individual words and phrases in the continuous speech stream (Gleitman & Wanner, 1982; Jusczyk, Houston, & Newsome, 1999; Morgan, 1996; Morgan & Demuth, 1996).

The prosodic contours of a language may be the first exposure a child has with language, possibly beginning before birth (Altman, 1997). Infants as young as four days old demonstrate a novelty response to foreign prosody, indicating that they find their native prosody familiar (Mehler *et al.*, 1988). They also demonstrate sensitivity to acoustic indicators of prosodic phrase boundaries, suggesting that they use those indicators early in the acquisition process (Christophe *et al.*, 1994; Christophe, Mehler, & Sebastián-Gallés, 2001).

The prosodic structure of a sentence may provide an initial framework for syntactic and semantic parsing (Speer, Shih, & Slowiaczek, 1989; Schafer, 1997). Prosodic boundaries, as often marked by pauses between phrases (among other

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

indicators), play a considerable role in interpretation, and prosody in general has been shown to influence multiple aspects of auditory processing and comprehension: beginning with segmentation of the speech stream, and continuing further to the disambiguation of syntactic structures (Koriat, Greenberg, & Kreiner, 2002; Kreiner, 2005; Kjelgaard & Speer, 1999), such as those below (Price *et al.*, 1991):

(1) a. Parentheticals: Mel knew (,) by the way (,) you were driving.
  b. Apposition: Only one remembered (,) the lady in red.
  c. Conjunction vs. Subordination: Mary was amazed and/Ann Dewey was angry.
  d. Tags: Dave will never know why he's enraged Willy/will he?

In relation to memory, prosody has been shown to mediate the effect of syntactic and semantic disruption on recall (Stine & Wingfield 1987), particularly when utterances are long or complex (Rosner *et al.*, 2004). Even further, the effect of prosody may extend beyond speech to both oral and silent reading. The Implicit Prosody Hypothesis (Fodor, 1998, 2002) proposes that readers project a "default" prosody onto text, which influences the processing of attachment and dependency relations, particularly across long distances. The prosody projected during reading may also affect syntactic processes such as agreement, where the segmentation of text into non-constituent chunks reduces readers' ability to detect subject-verb agreement errors (Kreiner, 2005). These effects may be related to multiple factors, such as prosodic alignment with syntactic constituents, which eases structurally based computations, or to facilitation of processing by way of scaffolding memory.

This chapter describes the prominent characteristics of prosodic features and structure, and provides an overview of how prosody interacts with processing for both production and comprehension, including reading and prosody within a second language.

## Prosodic structure and features

Early evidence of the validity of syntactic and prosodic constituency was found using *click displacement* studies (Fodor & Bever, 1965; Garrett, Bever, & Fodor, 1966). In this paradigm, participants were presented with spoken sentences in one ear, and clicks in the other ear, presented at specific locations in the sentence. According to listeners' reports, clicks presented in the middle of a syntactic constituent tended to instead be perceived as occurring at a syntactic boundary. For example, a click presented at | in (2) was often perceived as occurring after *happy*:

(2) That he was | happy was evident from the way he smiled.

This effect is robust enough that in later studies, listeners have reported hearing clicks at constituent boundaries, even when no clicks were actually presented (Reber & Anderson, 1970).

These studies provided early support for the centrality of the clause in processing; however, to disentangle syntactic effects from auditory effects, Wingfield and Klein (1971) presented listeners with ambiguous sentences where syntactic and prosodic boundaries either aligned, or where sentence were cross-spliced and the boundaries conflicted. They found that listeners falsely perceived clicks at both syntactic and prosodic boundaries. And interestingly, in the conflicting sentences, they were more likely to recall the sentences with an interpretation following the prosodic structure rather than the given syntactic structure.

In its most general definition, prosody is the segmentation of elements within a sentence or utterance, and the relative prominence of those segments to each other. In speech, these segments are associated with a particular pitch, and the overall sequence of pitch variations with an utterance, or *pitch contour*, can indicate illocutionary force. For example, the rising-falling contour of (3a) indicates statement intonation while the rising contour of (3b) indicates question intonation.

(3) a. 
Mary hired a brilliant lawyer.

b. 
Mary hired a brilliant lawyer?

One of the building blocks within prosodic structure is the pitch accent. A pitch accent is a tone or sequence of tones often associated with metrical prominence, or occurring in a lexically specified position, making it an accented syllable. Pitch accents are associated with focal material or new information, and in English, can be either a single tone (H*, L*), or bitonal (e.g., L+H*) A H* tone signifies a maximal rise in fundamental frequency ($F_0$) relative to the remainder of the utterance, a L* tone indicates a minimum $F_0$, while L+H* indicates a drop to a low $F_0$ followed by a rapid rise to a high $F_0$. Although not rigidly applied, H* often signals new information status (4), while L+H* tends to indicate a contrastive interpretation (5) (Watson, Tanenhaus, & Gunlogson, 2008; Pierrehumbert & Hirschberg, 1990).

(4) Who invited Mary?
**Bill** invited Mary.
H*

(5) Did Bill invite Sarah?
Bill invited **Mary**.
L+H*

Pitch accents may also affect the accessibility of a phrase during processing, which in turn influences attachment and dependency-resolution decisions (Carlson, Frazier, & Clifton, Jr., 2009).

The prosodic hierarchy (6) has sometimes been formalized by way of the Strict Layering Hypothesis, which states that a unit at each level is comprised exclusively of units from the next level down in the hierarchy (Selkirk, 1984; Selkirk, 2003; Nespor & Vogel, 1986; Hayes, 1989), although exceptions to the Hypothesis have led to its reformulation as a set of strong, yet violable constraints (see Ito & Mester, 1992, or Shattuck-Hufnagel & Turk, 1996, for an overview). The hierarchy can be used to define the domain in which certain phonological and prosodic phenomena are observed to take place.

(6) Prosodic hierarchy model (adapted from Nespor & Vogel, 1986)

Utterance (U)
|
Intonational Phrase (IP)
|
Phonological Phrase (φP)
|
Clitic Group (CG)
|
Prosodic Word (ω)
|
Foot (F)
|
Syllable (σ)

In this framework, an utterance is composed of at least one intonational phrase (*IP*). An IP is associated with a unique intonational contour and consists of at least one pitch accent and a boundary tone (L%, H%) (Pierrehumbert & Hirschberg, 1990). Larger boundaries are accompanied by final syllable lengthening, a larger pause, and pitch reset for following segments (Wightman *et al.*, 1992; Beckman & Pierrehumbert, 1986). IPs often align with syntactic units, but can be determined by semantic and discourse-related factors as well. Selkirk (1984) proposed that IPs are subject to a semantic constraint termed the Sense Unit Condition. A "sense unit" is comprised of either a single or multiple constituents that, in the semantic interpretation of the sentence, share a modifier or argument relation. The condition does not directly dictate which elements must form an IP, but prohibits elements that do not form a sense unit from occupying the same IP (Selkirk, 1984).

Parenthetical clauses, nonrestrictive modifiers, and other forms are typically contained within their own IP (Nespor & Vogel, 1986). Each IP consists of one or more phonological phrases. Phonological phrases (*φP*) are roughly equivalent to syntactic maximal projections, and are characterized at least one pitch accent, followed by a phrase accent (L-, H-). This hierarchical level between the intonational phrase and the prosodic word has also been termed a Major (*MaP*) or Minor Phrase (*MiP*) (Selkirk & Tateishi, 1988), Accentual Phrase (*AP*), or Intermediate Intonational Phrase (*ip*) (Beckman & Pierrehumbert, 1986).

A clitic group (*cg*) contains no more than one content word, and optionally any adjacent monosyllabic clitics, or function words. The prosodic word (*ω*) pertains to a morphosyntactic word, which may be defined as either a content or function word (Nespor & Vogel, 1986; Hayes, 1989), or may be restricted to content words only (Selkirk, 2003). Finally, foot (*F*) refers to a unit containing at most one stressed syllable, followed by any number of weak syllables dominated by the same node (Nespor & Vogel, 1986).

## Parsing the speech stream

Prosody plays a role in comprehension at multiple points in the parsing process; it facilitates segmentation of the speech stream, guides word recognition, and is used to provide and interpret acoustic cues to information structure. It can also disambiguate between syntactic structures, and support memory during parsing.

Young children and infants are sensitive to prosodic patterns in their language of exposure, and there is evidence that they are able to discriminate varying patterns at both the phrase and word levels (Christophe, Guasti, & Nespor, 1997; Christophe, Mehler, & Sebastián-Gallés, 2001; Bull, Eilers, & Oller, 1984; Bull, Eilers, & Oller, 1985) and that they use that ability as part of their acquisition process.

The *prosodic bootstrapping* hypothesis holds that prosody allows infants to segment fluent speech and identify the critical elements of a sentence and/or utterance (Nazzi & Ramus, 2003). Soderstrom, Seidl, Kemler Nelson, and Jusczyk (2003) found that infants as young as six months old are able to use prosodic cues to identify and segment phrasal units in connected speech. Other researchers have found that infants as young as two months old use prosodic grouping of speech into clauses to organize the input and encode information from the speech signal into memory (Mandel, Jusczyk, & Kemler Nelson, 1994; Hirsh-Pasek *et al.*, 1987).

Prosodic information is also used by adults during auditory word recognition. Grosjean and Gee (1987) cite evidence that (stressed) syllabic saliency in the speech stream guides word recognition (Cutler, 1976; Ladd, 1980), and that pausing patterns during oral reading align with prosodic structures, but not necessarily syntactic ones (Gee & Grosjean, 1983). Slowiaczek (1990) presented words with either a correct or incorrect stress pattern (no**TOR**ious/*notor**I**ous), finding that response times were faster for correct patterns than for incorrect patterns, and that word/non-word identification was faster for stimuli presented with correct lexical stress.

However, incorrect stress patterns have been found to be more highly disruptive to word recognition in Dutch and German than in English (Koster & Cutler, 1997; Friedrich, 2003). Since vowel reduction is less common in Dutch and German, and lexical pairs more frequently differentiated by stress patterns alone, this suggests that prosodic cues such as lexical stress are more informative in languages where vowel reduction is less likely to occur.

## Prosody in syntactic processing

Beyond the use of prosodic acoustic features during parsing, the structure of the prosodic hierarchy allows for phrasal and grouping effects. These effects also influence interpretation and have been shown to informatively interact with ambiguous syntactic configurations.

### *Coordination*

A basic function of prosody within syntactic disambiguation is coordination, or bracketing within a phrase. A sentence such as (7) is ambiguous between interpretation a) and interpretation b) (Lehiste, 1973):

(7) a. Old men and women sat on the bench.
b. [[Old men] and [women]] sat on the bench.
c. [Old [men and women]] sat on the bench.

In speech, the two forms can be disambiguated prosodically: in interpretation (8a), the word *men* tends to be longer in duration than in (8b), with characteristic $F_0$ differences that allow hearers to differentiate between the two. Katz *et al.* (1996) examined this effect in production with both adults and children. Participants were presented with three colored blocks (Pink, Green, & White) that were grouped into one of three configurations as in (8):

(8) a. [P]     [G] [W]
b. [P] [G]     [W]
c. [P]   [G]   [W]

Participants were asked to describe the arrangement of blocks using the phrase "pink and green and white" in a way that a blindfolded or absent person could determine the grouping. The adults were able to reliably produce word and pause durational cues to indicate groupings, although the children were not. However, in a closely related perceptual study by Beach, Katz, and Skowronski (1996), both children and adults were able to utilize prosodic cues (duration and pitch contour, manipulated independently) to disambiguate similar stimuli in comprehension. Nonetheless, there were notable differences between the participant populations; specifically the children's data, which was noisier than that of the adults, lead Beach *et al.* to conclude that "the mapping between prosody and phrasal interpretation [was] not fully mastered" (p. 1156) in five-year-olds and seven-year-olds.

### *Garden-paths and attachment ambiguities*

In studies of monolingual speakers, as well as of bilinguals and L2 learners, prosody has consistently been shown to influence syntactic interpretations, as well as perform a wide range of other communicative functions (Cutler, Dahan, & van Donselaar, 1997). Oral prosody influences the parsing strategies of readers, affecting their interpretation of ambiguous constructions such as (9), where *who was on the balcony* can be interpreted as modifying either *servant* (high attachment) or *actress* (low attachment) (% indicates a prosodic boundary).

(9) a. Someone shot the servant % of the actress who was on the balcony.
 b. Someone shot the servant of the actress % who was on the balcony.

In the absence of other cues, English speakers generally tend to prefer low attachment (Cuetos & Mitchell, 1988). However, the insertion of a prosodic phrase break after the first noun *servant* (9a) biases the interpretation toward low attachment, while a prosodic phrase break after the second noun *actress* (9b) biases toward high attachment (Fernández, 2007; Fodor, 1998; Maynell, 1999; Maynell, 2000).

In a language such as Korean, a head-final and pro-drop language, the potential for both global and temporary ambiguities is much higher than in English. Despite these typological differences, the alignment of prosodic and syntactic boundaries in Korean confers similar processing advantages. For Korean, a prosodic boundary following the initial NP biases listeners against interpreting it as the subject of the immediately following verb in both temporarily (10) and globally (11) ambiguous relative clause constructions (Kang & Speer, 2002, 2005).

| | | | | |
|---|---|---|---|---|
| (10) | *Caywon-ika* | *wulkois-nun* | *ai-rul* | *namwura-sse.* |
| | Caywon-NOM | cry-REL | child-ACC | scold-PAST |

“Caywon lightly scolded the child who was crying.”

| | | | | |
|---|---|---|---|---|
| (11) | *Caywon-ika* | *wulkoiss-nun* | *iywu-rul* | *chwuchukha-ysse.* |
| | Caywon-NOM | cry-REL | reason-ACC | guess-PAST |

a. “Caywon guessed the reason why (*pro*) was crying.”
b. “(*pro*) guessed the reason why Caywon was crying.”

In such participle constructions such as (12), where *wusu-myonse* “smiling” can modify either Yumi-NOM or reporter-ACC, a prosodic boundary after the ambiguous predicate, that is, *Yumi-ka wusu-myonse* %, biases interpretation toward the matrix subject (12a), while one placed after the initial NP biases interpretation toward the matrix object (12b).

| | | | | | |
|---|---|---|---|---|---|
| (12) | *Yumi-ka* | *wusu-myonse* | *cilmumha-nun* | *kica-rul* | *onghoga-ysse.* |
| | Yumi-NOM | smile-CONT | ask-REL | reporter-ACC | support-PAST |

a. “Yumi, smiling, supported the reporter who was asking questions.”
b. “Yumi supported the reporter who was smiling and asking questions.”

## *Prosody and parsing preferences*

Auditory processing studies have examined the role of prosody in both global and local ambiguities, and its ability to signal structural variations to the parser. Studies investigating globally ambiguous sentences indicate that prosody can be used to favor one syntactic interpretation over another (Nicol & Pickering, 1993; Schafer, 1997; Schafer, Speer, & Warren, 2005). Studies involving locally ambiguous sentences have explored the online effect of prosody as the listener is processing incoming structure and whether prosody could be manipulated to override processing preferences.

Marslen-Wilson, Tyler, Warren, Grenier, and Lee (1992) used the ambiguity of a verb taking either a direct object or a clause complement to test the effect of prosody on the Minimal Attachment principle. According to this principle, in the absence of other cues, the parser prefers to attach incoming material as a direct object of the current clause, rather than beginning a new subordinate clause Frazier, 1987; Frazier & Fodor, 1978). Thus, in a sentence preamble as in (13a), the parser initially analyzes the NP *several solutions* as the object of *knew*, thus preferring continuation (13b) to continuation (13c), leading to a garden-path effect.

(13) a. The pupils knew several solutions to the problem
b. …in Physics 100.
c. …would be quite possible.

Participants were presented the first part of the sentence auditorily, with prosody either favoring clause complement continuation (sharp $F_0$ fall on verb and upstep on NP) or a direct object continuation (continuing $F_0$ declination across verb and NP). Their reaction time was then measured during oral reading of a visual probe that displayed the first word of the clause complement version (*would* in (13c)). Reaction times were faster in the clause prosody condition than the direct object prosody condition, indicating that prosody is not only used by the parser to resolve ambiguities while structure building, but is also able to override parsing preferences such as Minimal Attachment.

Speer *et al.* (1996) examined the effect of prosody on the Late Closure principle (Frazier & Rayner, 1982), which calls for an incoming element to be attached to the current clause. Speer *et al.* presented sentences as in (14) with cooperating, neutral, or conflicting prosody in both a comprehension task and a cross-modal naming task. In the cooperative condition, the prosody and syntax matched, and in the conflicting condition, they did not. The neutral condition had no prosodic boundary in the critical region, and was considered equally appropriate for either interpretation (below, % indicates a prosodic boundary, and/indicates a syntactic boundary).

(14) a. Cooperating, Late Closure: Whenever the guard checks the door %/it's locked.
b. Cooperating, Early Closure: Whenever the guard checks %/the door is locked.

c. Conflicting, Late Closure: Whenever the guard checks % the door/it's locked.
d. Conflicting, Early Closure: Whenever the guard checks the door % is locked.
e. Neutral, Late Closure: Whenever the guard checks the door it's locked.
f. Neutral, Early Closure: Whenever the guard checks the door is locked.

In the neutral condition of the naming task, there was an advantage for the (14e) interpretation, following the Late Closure principle. In the cooperating condition, overall reaction times were faster, while in the conflicting condition, reactions times were slower. Crucially, there was no advantage in the cooperating late closure condition, suggesting that prosodic cues do not necessarily further facilitate processing of preferred parses.

These results indicate that the presence of a prosodic boundary at a point of ambiguity can influence interpretation, and, as in the Marslen-Wilson *et al.* (1992) study, can override parsing preferences. Speer *et al.* (1996) propose that a prosodic boundary serves to close the current constituent, allowing the parser to assume a potentially structurally dispreferred interpretation. Kjelgaard and Speer (1999) further explored this interaction between prosody and closure strategies with equivalent results, concluding that the facilitation and interference effects of prosody indicate that listeners use prosodic information online, and at an early stage, during parsing.

## *Prosodic boundary size*

The presence and location of prosodic boundaries in an utterance factor significantly in interpretation, and thus play a critical role in the comprehension process. In the most general sense, elements are more likely to be processed together when they are contained within a prosodic unit than if they are separated by an intonational boundary.

Along with the location of a boundary, the relative size of a boundary may provide additional information to the parser. Carlson *et al.* (2001) proposed that the effect of a prosodic boundary is not determined by its absolute size, but by its size in relation to other relevant boundaries in the utterance. In a sentence such as (15) containing an ambiguous adjunct, the final PP *after John visited* can either attach high in the matrix clause to *learned* or low in the embedded clause to *telephoned*. They found that if the boundary following *learned* was larger than the boundary following *telephoned*, low attachment was preferred. However, if the boundary following *telephoned* was larger, high attachment was preferred.

(15) Susie learned | that Bill telephoned | after John visited.

Clifton, Jr. *et al.* (2002) extended these findings to other ambiguous constructions such as coordination structures, possessive phrases, relative clauses, and adverbial modification, building evidence that the parser takes advantage of global prosodic representation during comprehension.

Since much of the work on prosody involves comprehension and interpretation of structures as presented by trained speakers, Schafer *et al.* (2000) explored how naive speakers may use prosody to disambiguate forms in a more naturalistic setting. Focusing on the early/late closure ambiguity, they examined the production of sentences as in (16) in a structured discourse task. In the early closure sentence (16a), *moves* is an intransitive verb, and *square* is the subject of the second clause. In the late closure sentence (16b), *moves* is transitive, and *square* is its object.

(16) a. When that moves the square will…
 b. When that moves the square it…

Schafer *et al.* found that beyond boundary presence and location, the parser most likely takes into account overall prosodic structure in disambiguation. For example, a boundary after *moves* biases toward an early closure interpretation only if the boundary following *square* is smaller. If the boundaries are equal, other prosodic features may come into play, and there is greater speaker variation in the features used to disambiguate structures.

Interestingly, Schafer and colleagues also found that prosody was used to disambiguate a structure already disambiguated by the discourse context, which strongly suggests that prosody performs a critical function in processing and comprehension beyond clarification of ambiguous contexts.

Adding to Schafer *et al.*'s demonstration that prosody is used and understood productively by naive speakers in more naturalistic contexts, subsequent research by Snedeker and Trueswell (2003) suggests that efficiency considerations may interact with when prosodic differentiation is emphasized. Snedeker and Trueswell (2003) report that speakers may not use prosody to disambiguate unless they are aware of the ambiguity, and if the context does not already disambiguate.

## *Prominence and pitch accent effects*

Pitch accents also influence sentence comprehension, indicating that a word or phrase is focused with either new or contrastive information status. The location of a pitch accent may also affect attachment preferences in ambiguous relative clauses. Schafer, Carter, Clifton, and Frazier (1996) make a compelling case for the use of prosodic information during parsing, in particular, its interaction with information structure, and the alignment of prosodic features with semantic interpretation. In the absence of strong intonational boundary cues, listeners are more likely to attach a relative clause to the noun that received a pitch accent. Additionally, relative clauses are more likely to be attached to contrastively accented NPs than to focally accented NPs. Maynell (1999) found a significant effect of a contrastive L + H* accent, where placement of the accent on the NP1 favored NP1 attachment of the relative clause. These findings further support arguments that the information status of incoming elements is inferred online, and that listeners have access to full prosodic representations during parsing.

## Prosody in extended domains

### *Prosody and memory*

Early research on working memory maintained that items held in memory are stored and processed in specialized components based on information type—the visuospatial sketchpad for visual field items, and the phonological loop for spoken and written material (Baddeley & Hitch, 1974).

Related to the phonological loop is the concept of an auditory buffer, which forms part of the selective listening and memory model of Broadbent (1958, 1971). As elaborated by Frankish (1989, 1995), the auditory buffer privileges memory for last items in a list, even if multiple lists are presented. This suggests that the grouping of spoken utterances into phrases may take advantage of this feature in the auditory buffer. The temporal phrasing of speech may increase the efficiency of auditory memory, and thus play a significant role in comprehension.

The rapid decay of the elements held in memory can be slowed by rehearsal, such as repeating a telephone number until it can be dialed. Slowiaczek and Clifton (1980) demonstrated that during silent reading, rehearsal takes the form of subvocalization, assisting readers in building a mental representation of the sentences being read. When subvocalization was disrupted by the readers performing a verbal task (e.g., repeating syllable strings silently), comprehension was impaired. Prosody is hypothesized to be a critical component of the subvocalization routine, and thus presents consequences for the storage and processing of material in the phonological loop.

Prosody has also been shown to mediate the effect of syntactic and semantic disruption on recall (Stine & Wingfield, 1987), and interestingly, when prosody and syntax conflict, prosody "wins," leading to potential errors in recall (Wingfield, 1975). Beyond the effect of prosody on syntactic computations, there is a long history of research supporting its effect on memory for speech. Epstein (1961) and subsequent work by O'Connell, Turner, and Onuska (1968), Leonard (1973), and Harriman and Buxton (1979) demonstrated that memory for nonsense syllables was increased by adding morphosyntactic structure, and that the addition of sentence prosody alone could improve memory performance.

The effect of prosody on memory is clearly demonstrated in studies showing that synthetic speech produced without prosodic cues adversely affects memory and comprehension (Paris *et al.*, 2000), and natural prosody in speech and oral reading is more effective in aiding memory than monotone reading (Koriat, Greenberg, & Kreiner, 2002). Rosner *et al.* (2004) further show that the facilitating effect of prosody is seen most prominently in contexts which are more difficult to process, in particular with long or complex utterances.

While prosody alone may not be a primary information source during parsing, it can provide a level of ancillary support from the input that aids processing. When prosody is disrupted or removed completely, greater strain is placed on memory and memory-based tasks, an effect seen even more robustly in complex and demanding parsing contexts.

## *Prosody in reading*

Given the relative absence of explicit prosodic cues and boundaries in text, examining the role of prosody during reading requires a modified approach. The ability to project appropriate prosody in reading requires that the reader correctly assign syntactic roles to the sentential elements, demonstrating a grasp not only of the structure, but of the general message of the sentence (Chafe, 1988). Studies show a strong correlation between phrasing ability and comprehension, and it has been contended that the ability to appropriately phrase textual material into meaningful units is fundamental to fluency in both oral and silent reading (Clay & Imlach, 1971; Dowhower, 1991; Paige *et al.*, 2014). Appropriately chunking text into syntactically and semantically related groups reflects cognitive restructuring of the input and leads to successful encoding into memory. Once this level of reading skill has been achieved, simultaneous improvements in comprehension are often observed (Rayner & Pollatsek, 1989).

Early studies in reading fluency noted the correlation of fluency with comprehension, also noting that greater reading fluency is associated with more "appropriate" prosodic phrasing and contours (Clay & Imlach, 1971; Dowhower, 1987). Moreover, studies have shown that chunking text in such a way as to preserve major syntactic and prosodic boundaries improves oral fluency (LeVasseur *et al.*, 2006; LeVasseur, Macaruso, & Shankweiler, 2008; Rasinski *et al.*, 1994). Prosody may be seen as an intermediary between fluency and comprehension, such that individuals who demonstrate appropriate prosody are more likely to exhibit better comprehension as well (Paige *et al.*, 2014).

## *Prosodic phrasing and reading fluency*

Much of the research on prosody and reading comprehension focuses on automaticity and the developing skill sets of children and other early readers. However, what is the role of prosody for more advanced native readers? Is prosody simply transference of auditory processing techniques to reading?

The ability to group words into appropriate units is a key aspect of fluent reading (see Allington, 2006, among others), and non-fluent readers tend to either read word-by-word or group words differently than typical oral speech patterns (Kuhn & Stahl, 2003). Prosody in oral language can be used as a cue to syntactic and semantic information. Children are particularly attentive to prosody in the auditory input (Schreiber, 1980) and must transition this skill to reading by using punctuation and grammatical cues to appropriately segment text. Even with advanced word decoding skills, fluency is not achieved unless this ability to segment text is developed (Schreiber, 1991).

Pedagogical research has investigated how manipulation of text presentation format may enhance reading skill and support the development of reading fluency. Pre-segmentation of text into meaningful phrasal chunks has been shown to improve reading performance in both children (LeVasseur *et al.*, 2006; O'Shea & Sindelar, 1983) and less skilled adult readers (Cromer, 1970). Skilled readers seem to be more resistant to imposed text segmentation, suggesting that their own phrasing skills override cues from the input.

In an early study, Cromer (1970) investigated the effect of text segmentation on the comprehension of good and poor reader groups. Sentences were presented in one of four formats: regular sentence (17), single word (18), phrase (19), or fragmented group (20).

(17) The cow jumped over the moon.
(18) The | cow | jumped | over | the | moon.
(19) The cow jumped | over the moon.
(20) The cow | jumped over the | moon.

Assuming that skilled readers chunk text into phrases, while less-skilled readers read word-by-word, Cromer predicted that guiding the less skilled groups to read in phrases would improve their comprehension (i.e., make them look like more skilled readers), and guiding the skilled readers to read word-by-word would disrupt their comprehension (i.e., make them look like less skilled readers).

All participants were matched for IQ, but were grouped based on their performance on comprehension and vocabulary scores. Participants were first divided based on comprehension score into "good" and "poor" reader groups. The "poor" readers were further divided into a "difference" group, which had low comprehension scores but were matched with the good readers in vocabulary scores, and a "deficit" group, which had both low comprehension and vocabulary scores. Cromer found that comprehension varied based on both reading skill level and presentation format. "Good" readers were unaffected by presentation format, and comprehended equally well in all conditions. "Difference" readers were disrupted in the word and fragment conditions, but improved in the phrase condition. "Deficit" readers comprehended best in the word condition and were not significantly affected by any of the other conditions.

Cromer concluded that while "good" and "deficit" readers may be more or less impervious to text manipulations, either due to the strength of their own phrasing skills ("good" readers), or deficit in their reading fluency ("deficit" readers), "difference" readers may benefit from textual phrasing as an aid to reading fluency. This suggests that (i) reading comprehension involves additional components beyond general IQ and vocabulary knowledge, and (ii) as long as sufficient vocabulary skills have been acquired, facilitating text presentation may directly influence reading comprehension.

## *Summary*

Prosody is sometimes omitted from structurally based parsing models, perhaps due to difficulty in defining the relevant features of prosody within those models, and their scope of application. However, the research overwhelmingly demonstrates that prosody is used by the parser to clarify syntactic variations, identify clausal and phrasal boundaries, as well as facilitate online memory for structure and memory for recall. This effect extends even to silent reading, where implicit prosody is shown to affect ambiguity resolution and agreement processes, particularly in high processing or memory load contexts.

## *Prosody in silent reading*

Explicit prosody has been shown make use of phrasing (breaks) and intonational cues (pitch accents) to disambiguate between syntactic representations (Schafer *et al.*, 1996). In the absence of explicit prosodic cues, is there evidence that phonological features such as phrasing and intonation are projected, and can similarly influence interpretation? Early evidence from Baddeley and Hitch (1974) and Slowiaczek and Clifton (1980) suggests that if subvocalization is blocked, comprehension during silent reading is impaired; thus, it appears that rehearsal involves projection of phonological/prosodic information, and that this information contributes to the processing and comprehension.

In subsequent years, there has been additional evidence that fluent readers are not only able to produce prosody during oral reading, but while reading silently as well (Bader, 1998; Fodor, 2002). The experience of hearing an "inner voice" during silent reading has long been anecdotally attested, however, more recent research also supports this experience theoretically and empirically. Because prosody is not consistently or always explicitly indicated in written language, it is often unclear what role it plays in reading and whether that role is critical to reading comprehension. As in production and comprehension of oral prosody, there is some controversy as to whether prosody is a direct reflection of syntactic processes, or whether it makes a unique contribution to syntactic (and other) analyses during parsing.

There is, of course, concern as to whether phrasing effects during reading directly correspond to similar prosodic effects in auditory comprehension. Previous studies investigating presentation formats have often attributed processing differences to either purely syntactic factors, for example, whether phrasal breaks are given at syntactic boundaries or not, or to the disruptive effect of word-by-word presentation, which prevents typical reading behaviors such as parafoveal preview and regressive eye movements to earlier material (Rayner *et al.*, 2012). However, many studies have closely linked presentation format with prosodic phrasing and oral fluency during reading (see LeVasseur *et al.*, 2006, 2008; Rasinski *et al.*, 1994). Further, Staub (2007) and Hirotani, Frazier, and Rayner (2006) present evidence that the appropriate insertion of commas in both simple and complex sentences facilitates reading, suggesting that they act as cues to implicit prosodic boundaries. Taken together, it is not unreasonable to assume that manipulation of text format may directly disrupt the projection of a prosodic contour during silent reading.

Koriat *et al.* (2002) proposed the Structural Precedence Hypothesis, which claims that during reading, readers establish an early structural frame for a phrase or sentence based on function words and morphosyntactic cues that may indicate general phrase structure. Prosody during oral reading reflects this early processing and may be used to help maintain a structure in memory while further processing takes place. While some semantic influences may play an early role, syntactic and prosodic processing precede complete semantic analysis, as evidenced by the ability to project appropriate prosody onto nonsense sentences as long as morpho-syntactic cues remain intact.

Other researchers have suggested that prosody more directly influences both early syntactic and reanalysis processes. Bader (1998) shows that prosody is able to affect the ease of reanalysis during reading of ambiguous sentences: he claims that during reading, both a prosodic and a syntactic structure are produced. If revision of syntactic structure is necessary, it is made more difficult if the prosodic structure must be revised as well (Prosodic Constraint on Reanalysis, Bader, 1998, p. 8). The Implicit Prosody Hypothesis takes this claim further by stating that a default prosodic contour is projected onto text during silent reading, and this projection directly affects interpretations such as ambiguity resolution (Fodor, 1998, 2002).

Related claims have been made for lexical stress variations (Breen & Clifton, Jr., 2011), rhythmic stress patterns (Ashby & Clifton, Jr., 2005; Ashby & Martin, 2008; Kentner, 2012), and prosodic phrase lengths (Hirose, 2003; Hwang & Schafer, 2009; Hwang & Steinhauer, 2011), suggesting that many prosodic features contribute to processing during reading.

Disruption of implicit prosody has been shown to affect agreement processing during reading, where errors in subject-verb agreement are less likely to be detected if natural reading rhythm is impeded. Kreiner (2005) proposed that natural reading prosody facilitates online syntactic integration, allowing subject-verb mismatches to be more easily detected. However, if this integration is disrupted in some way, mismatches will be more difficult to detect.

Testing this hypothesis, Kreiner found no significant effect of prosody when the subject and verb were adjacent. However, when the subject and verb were separated by a relative clause, agreement errors were detected only when participants read with natural prosody, not when prosody was disrupted. These results suggest that natural prosody does facilitate subject-verb agreement processing, particularly when processing and/or working memory load is greater.

Relatedly, it has been found that rapid serial visual presentation (RSVP) of text has the potential to accelerate reading rates, but seems to result in reduced comprehension rates (Bernard, Chaparro, & Russell, 2001; Kang & Muter, 1989). This effect is attributed to the fixed presentation rate of the materials, suggesting that the invariable pace interferes with the projection of prosody onto the text (Fernández, 2007; Castelhano & Muter, 2001).

## Prosody in second language processing

Second language (L2) prosody in processing has not been as extensively studied, nor is there a consensus as to how L2 learners may use prosodic information. While many unanswered questions remain in this field, several studies have set the groundwork for research in this area.

Harley, Howard, and Hart (1995) investigated age effects in the use of prosodic cues to sentence structure, to determine whether learners would favor prosodic or syntactic cues to structure when prosody and syntax were in conflict. Learners were presented with sentences as in (21) and (22) with either natural prosody (a), or conflicting prosody (b). The conflicting prosody items were created by splicing the first part of (a) sentences with the continuation of the (b) counterpart, and vice versa.

(21) a. The new teacher's watch | has stopped.
     b. The new teacher's | watch has stopped.

(22) a. The new teachers | watch baseball on TV.
     b. The new teachers watch | baseball on TV.

L1 and L2 participants ranging from 7 to 23 years of age were asked to identify the subject noun phrase of each sentence. The oldest native group had the highest accuracy, demonstrating the ability to use syntactic information to override conflicting prosodic cues. On the other hand, the younger native groups and all L2 groups were significantly influenced by prosodic cues. The authors attribute these results to the primacy of prosodic information, particularly at earlier stages of language development. They suggest that since prosodic cues are often salient and reliable in the input, overriding those cues requires an advanced skill level not available to younger or less experienced speakers.

Dekydtspotter, Donaldson, Edmonds, Fultz, and Petrush (2008) examined L2 relative clause attachment ambiguities using both aural and written stimuli. Beginning and intermediate L2 French learners were tested on their resolution of ambiguous constructions as in (23).

(23) a. Nous adorons | le secrétaire | du psychologue qui se promène | (au centre ville).
     b. Nous adorons | le secrétaire du psychologue | qui se promène | (au centre ville)."We adore the secretary of the psychologist who takes a walk (downtown)."

Experiment 1 was a silent reading task manipulating relative clause (RC) length, where items contained either a long RC (the modifier in parentheses), or a short RC (no modifier). In Experiment 2, identical items were presented both aurally and in written form. Both RC length and intonation were manipulated such that the RC either formed a constituent with the NP2 (23a), or the NP1 and NP2 formed a constituent and the RC its own constituent (23b). Experiment 3 was a self-paced reading task where contextually disambiguating information was provided for each of the items.

The results for Experiment 1 indicated that beginner learners were not sensitive to RC length effects, although intermediate learners were. In Experiment 2, overall, the learners did not show an effect of intonation; however, a subset of intermediate learners did consistently make use of the prosodic cues. And finally, in Experiment 3, all learner groups demonstrated a Minimal Attachment bias, where, despite the disambiguating contexts, response times were shorter when the RC modified NP2 than when it modified NP1.

Overall, these results demonstrate that not only can learners make use of prosodic cues, but they are also capable of deploying and integrating syntactic and prosodic strategies during parsing. Even further, in some cases, L2 speakers may rely more heavily on prosodic rather than syntactic cues, particularly when syntax and prosody are misaligned (Harley, Howard, & Hart, 1995).

## *L2 prosody and reading*

Looking at major prosodic features crosslinguistically, languages perhaps most notably differ in intonation and lexical stress patterns (see discussion in Cutler, 2012). Languages may also differ in typical or acceptable prosodic phrase lengths, and may use prosody differently to indicate variations in information structure. For example, French speakers use different syntactic strategies than English speakers to indicate topic and focus elements in spontaneous discourse. This use of distinct syntactic arrangements (e.g., clefting, dislocation) is, in turn, reflected in the prosodic pattern of the output in each language (Holmes, 1995).

However, general prosodic phrasing patterns appear to be more universal, at least when those patterns align with syntactic constituents. In several notable studies, naive listeners were able to correctly identify pauses occurring at constituent boundaries, even with no previous exposure to the test language (Wakefield, Doughtie, & Yom, 1974; Pilon, 1981; Endress & Hauser, 2010). L2 learners also demonstrate an ability to use both auditory and written prosodic cues to disambiguate structures, and revise initial parses.

Evidence would suggest, then, that where the performance of L1 and L2 speakers may diverge is not necessarily in prosodic phrasing itself, but in its relation to overall fluency and the availability of processing resources. Tasks manipulating the prosodic projection environment are thus further complicated in reading, adding the processing burden of decoding and automatized word recognition in a second language.

Most L2 reading fluency and comprehension studies focus on processing speed in word recognition, leaving the role of prosody unexplored. This is perhaps partially due to the difficulty in accurate and consistent measures of prosodic patterns. However, since prosodic phrasing is typically indicated by pauses in oral production, measures of pause frequency and location can be used to measure development of native-like patterns in L2 speakers.

For native speakers of a language, the clause typically operates as a planning unit in speech, and pauses tend to occur at clause boundaries (Butterworth, 1980; McDaniel, McKee, & Garrett, 2010; Garrett, 1982). However, for low-fluency L2 learners, there is no evidence of the clause as a planning unit (Temple, 2000). Pauses are distributed across the utterance, and fewer slowdowns at clause boundaries than in native speech. However, as the development of fluent L2 speech progresses, pauses at clausal boundaries increase, and may begin to converge with native speech patterns. Thus, despite crosslinguistic similarities in prosodic and syntactic phrase alignment, something specific in the development of L2 fluency is needed before native-like pausing patterns can emerge. If implicit prosody is the projection of prosodic contours onto text, it would follow that similar effects of fluency would be found for reading as well.

Liljestrand Fultz (2009) examined the effect of prosodic phrasing on L2 ambiguity resolution, and found differences based on syntactic structure and complexity. While the L2 learners consistently used prosodic information to disambiguate conjunct modification and PP-attachment constructions, no clear preference was found for relative clause ambiguities. She suggests three possible explanations for this result: (i) learners cannot perceive the prosodic cues for certain structures, and so the parser cannot incorporate this information during processing, (ii) learners can perceive the prosodic cues, but the parser cannot incorporate the information for independent reasons, or (iii) learners perceive and incorporate the cues, but for certain structures are not able to use that information in conjunction with other information from the parse. The first two possibilities do not fit her results: the L2 learners were able to both perceive and incorporate prosodic cues for conjuncts and PP constructions. This leaves the third possibility: that learners do perceive and use prosodic cues, but the parser may not be able to integrate that information in complex computations required by relative clauses. At earlier stages of proficiency, learners may be able to effectively integrate all information when parsing a simpler structure, but not a complex one. As proficiency increases and processing routines develop, the parser is able to more efficiently integrate information from multiple cues, even when the computation is complex.

It may be that the L2 participants are generally able to perceive and utilize prosodic cues in certain circumstances. However, when complex computations are required, those cues, may not be effectively integrated at the point of interpretation.

Looking at Dutch L2 speakers of English, Anema (2008) found evidence supporting that more native-like prosody, as indicated by frequency and location of pauses, correlated with higher scores in reading comprehension. Two experimental participant groups were recruited for this study: Dutch-English bilinguals living for at least one year in the United States, and Dutch-English bilinguals in the Netherlands.

The two groups, immersed and non-immersed bilinguals, did not differ significantly in working memory span, and both performed at near-native like levels in a list-generation task. However, they performed significantly differently in both pausing patterns and reading comprehension. The participants who had spent at least one year immersed in the L2 environment performed better on reading comprehension measures than the non-immersed group. This group also displayed more native-like pausing patterns in oral reading, pausing less frequently at inappropriate positions, and pausing less frequently overall.

In addition to strengthening the evidence for a link between phrasing and comprehension, this study also suggests a strong role of exposure to the development of prosodic phrasing strategies and related reading fluency.

## *Text segmentation in L2 reading*

While L2 reading fluency has not been given the same attention as L1 reading fluency, there is a clear connection between reading fluency and comprehension (Fuchs, Fuchs, & Maxwell, 1988; Nathan & Stanovich, 1991; Jenkins *et al.*, 2000; Fuchs *et al.*, 2001), which suggests that comprehension improves with the development of fluent reading patterns and phrasing skills. Variations of Cromer's

(1970) study have been adapted for exploration of text segmentation and reading comprehension in an L2.

In a series of studies with Japanese learners of English, Kadota and colleagues examined the effect of text segmentation on reading comprehension. Kadota (1982) and Kadota and Tada (1992) found that text segmented into phrasal units improved comprehension and recall rates over sentence or word unit presentations (see Yamashita & Ichikawa, 2010). In subsequent work, Kadota, Yoshida, and Yoshida (1999) presented text in three modes: word-by-word (24), phrase-by-phrase (25), and clause-by-clause (26).

(24) A | glacier | is | a | river | of | ice. | It | may | be | ten | to | thirty | miles | long | and | one | or | two | miles | wide.
(25) A glacier is a river of ice. | It may be ten to thirty miles long | and one or two miles wide.
(26) A glacier | is a river | of ice. | It may be | ten to thirty | miles long | and one or two | miles wide.

Comprehension was higher and reading times faster in both the phrase and clause conditions than in the word condition.

Yamashita and Ichikawa (2010) further expanded on this design, presenting narrative texts to Japanese learners of English in four modes: Whole, Word-by-Word, Chunk, and Fragment. In the Chunk presentation mode, boundary positions roughly corresponded with phrasal boundaries (27), while in the Fragment presentation mode, boundary positions deliberately violated grammatical units (28).

(27) The origin of Australian Rules Football | is unclear. Some people say | it might have developed | from an ancient game | in which a ball made of kangaroo skin | was kicked around.
(28) The origin of Australian | Rules Football is unclear. | Some people say it might | have developed from an | ancient game in | which a ball made of kangaroo | skin was kicked around.

Yamashita and Ichikawa predicted that lower proficiency learners' comprehension would be facilitated by appropriately chunked text (Chunk mode), and disrupted by inappropriately chunked text (Fragment mode), while advanced learners would not be affected by presentation mode. Due to ceiling effects in the comprehension measure, they were unable to confirm the facilitative effect of appropriate chunking in either test group; however, in the Fragment mode, comprehension was significantly lower for the lower proficiency group than in any other mode. There was no effect of mode on comprehension for the advanced learners, although reading times in the Word-by-Word mode were significantly longer.

The results suggest that the advanced learners' typical phrasing patterns may override the effect of text presentation, but lower proficiency learners' underdeveloped phrasing patterns make their reading more susceptible to both disruptive or facilitative effects.

Recent research by Pratt (2015) explores structural and prosodic effects during reading, examining their influence on agreement processing and comprehension in native English (L1) and Spanish-English bilingual (L2) speakers. The experimental design manipulated text presentation to influence implicit prosody, using sentences designed to induce subject-verb agreement attraction errors. Materials included simple and complex relative clauses with head nouns and verbs that were either matched or mismatched for number. Participants read items in one of three presentation formats (whole sentence, word-by-word, or phrase-by-phrase), rated each item for grammaticality, and responded to a comprehension probe. Results indicated that presentation format differentially affected both measures in the L1 and L2 groups. For the L1 participants, facilitating the projection of phrasal prosody onto text (phrase-by-phrase presentation) enhanced performance in agreement processing, while disrupting prosodic projection via word-by-word presentation decreased comprehension accuracy. For the L2 participants however, phrase-by-phrase presentation was not significantly beneficial for agreement processing, and also resulted in lower comprehension accuracy. More work is needed to specify exactly why L1 and L2 speakers demonstrate different patterns of interference and facilitation, however, these results indicate a universally significant role of prosodic phrasing in at least some aspects of syntactic processing and memory retrieval, regardless of language background and profile.

## Conclusions

The wealth of research into the function and role of prosody over the last few decades has greatly advanced the study of sentence processing in general. Prosody is an integral part of language that is intertwined with practically all aspects of processing, from the first stages of acquisition, to syntactic analysis, memory, and advanced reading fluency.

The research summarized here demonstrates the importance of prosodic representations in both comprehension and grammatical computations. However, much work still remains in clarifying how prosodic factors interact with other linguistic elements and cognitive factors, and how this information may be applied productively to native speakers and learners of all languages.

The future of the field will likely involve broader processing models coordinating syntactic, semantic, phonological, and discourse features, and prosody is staged to be a uniting link in the study of language structure and its cognitive correlates. Further work is needed to more comprehensively describe the prosodic system, and how it interacts with language components at all stages of processing.

The growing intersection of language and technology will also draw upon elements within the prosodic realm. From greater reliance on voice recognition and computer-generated speech, to increasing prevalence of small-display text presentation, the application of prosody to modern communication will become increasingly compelling.

## REFERENCES


Allington, R. L. (2006). *What Really Matters for Struggling Readers: Designing Research-Based Programs*. 2nd Edition. Boston: Allyn & Bacon.

Altman, G. T. M. (1997). *The Ascent of Babel*. Oxford: Oxford University Press.

Anema, I. (2008). The relationship between fluency-based Suprasegmentals and comprehension in oral and silent reading in Dutch speakers of English. City University of New York.

Ashby, J. & Clifton, Jr., C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, 96(3). doi: 10.1016/j.cognition.2004.12.006

Ashby, J. & Martin, A. E. (2008). Prosodic phonological representations early in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 34(1), 224–236. doi: 10.1037/0096-1523.34.1.224

Baddeley, A. D. & Hitch, G. (1974). Working Memory. *Psychology of Learning and Motivation*, 8, 47–89. doi: 10.1016/S0079-7421(08)60452-1

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in Sentence Processing*, (pp. 1–46). Dordrecht: Kluwer. doi: 10.1007/978-94-015-9070-9_1

Beach, C. M., Katz, W. F., & Skowronski, A. (1996). Children's processing of prosodic cues for phrasal interpretation. *Journal of the Acoustical Society of America*, 99(2), 1148–1160. doi: 10.1121/1.414599

Beckman, M. E. & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255–309. doi: 10.1017/S095267570000066X

Bernard, M. L., B. Chaparro, B. S., & Russell, M. (2001). Examining automatic text presentation for small Screens. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 45 (6), 637–639. doi: 10.1177/154193120104500613

Breen, M. & Clifton, Jr., C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language* 64(2). Elsevier Inc., 153–170.

Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon Press.

Broadbent, D. E. (1971). *Decision and Stress*. London: Academic Press. doi: 10.1016/B978-1-4832-0079-8.50009-8

Bull, D., Eilers, R. E., & Kimbrough Oller, D. (1984). Infants' discrimination of intensity variation in multisyllabic stimuli. *The Journal of the Acoustical Society of America*, 76, 13–17. doi: 10.1121/1.391110

Bull, D., Eilers R. E., & Kimbrough Oller, D. (1985). Infants' discrimination of final syllable fundamental frequency in multisyllabic stimuli. *The Journal of the Acoustical Society of America*, 77, 289–295.

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language Production: Speech and Talk* (pp. 155–176). London: Academic Press.

Carlson, K., Clifton, Jr. C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, 45, 58–81. doi: 10.1006/jmla.2000.2762

Carlson, K., Frazier, L., & Clifton, Jr. C. (2009). How prosody constrains comprehension: A limited effect of prosodic packaging. *Lingua*, 119, 1066–1082. doi: 10.1016/j.lingua.2008.11.003

Castelhano, M. S. & Muter, P. (2001). Optimizing the reading of electronic text using rapid serial visual presentation. *Behavior & Information Technology*, 20(4), 237–47. doi: 10.1080/01449290110069400

Chafe, W. (1988). Punctuation and the prosody of written language. *Written Communication*, 5, 395–426. doi: 10.1177/0741088388005004001

Christophe, A., Dupoux, E., Bertonici, J., & Mehler, J. (1994). Do infants percieve word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, 95(3), 1570–1580. doi: 10.1121/1.408544

Christophe, A., Guasti, T., & Nespor, M. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, 12(5-6), 585–612. doi: 10.1080/016909697386637

Christophe, Anne, Jacques Mehler, and Núria Sebastián-Gallés. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, 2(3), 385–394. doi: 10.1207/S15327078IN0203_6

Clay, Marie M, and Robert H Imlach. (1971). Juncture, pitch, and stress as reading behavior variables. *Journal of Verbal Learning and Verbal Behavior*, 10, 133–39. doi: 10.1016/S0022-5371(71)80004-X

Clifton, Jr., C., Carlson, K., & Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech*, 45(2), 87–114. doi: 10.1177/00238309020450020101

Cromer, W. (1970). The difference model: A new explanation for some reading difficulties. *Journal of Educational Psychology*, 61, 471–483. doi: 10.1037/h0030288

Cuetos, F. & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition*, 30(1), 73–105. doi: 10.1016/0010-0277(88)90004-2

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*, 20(1), 55–60. doi: 10.3758/BF03198706

Cutler, A. (2012). *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge, MA: MIT Press.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40, 141–201.

Dekydtspotter, L., Donaldson, B., Edmonds, A. C., Liljestrand Fultz, A., & Petrush, R. A. (2008). Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by english-french learners. *Studies in Second Language Acquisition*, 30(4), 453. doi: 10.1017/S0272263108080728

Dowhower, S. L. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Reading Quarterly* 22(4), 389–406. doi: 10.2307/747699

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory Into Practice*, 30(3), 165–175. doi: 10.1080/00405849109543497

Endress, A. D. & Hauser, M. D. (2010). Word segmentation with universal prosodic cues. *Cognitive Psychology*, 61(2), 177–199. doi: 10.1016/j.cogpsych.2010.05.001

Epstein, W. (1961). The influence of syntactical structure on learning. *American Journal of Psychology*, 74(1), 80–85. doi: 10.2307/1419827

Fernández, E. M. (2007). How might a rapid serial visual presentation of text affect the prosody projected implicitly during silent reading? In *Conferências Do V Congresso Internacional Da Associaçao Brasiliera de Lingüistica*, 5, 117–54.

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27(2), 285–319.

Fodor, J. D. (2002). Psycholinguistics cannot escape prosody. In *Proceedings of the 1st International Conference on Speech Prosody*, 83–88.

Fodor, J. A. & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 4(5), 414–420. doi: 10.1016/S0022-5371(65)80081-0

Frankish, C. R. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(3), 469–479. doi: 10.1037/0278-7393.15.3.469

Frankish, C. R. (1995). Intonation and auditory grouping in immediate serial recall. *Applied Cognitive Psychology*. 9(7), S5–S22. doi: 10.1002/acp.2350090703

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII*, (pp. 559–586). Hillsdale, NJ: Erlbaum.

Frazier, L. & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291–325. doi: 10.1016/0010-0277(78)90002-1

Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178–210. doi: 10.1016/0010-0285(82)90008-1

Friedrich, C. (2003). Prosody and spoken word recognition: Behavioral and ERP correlates. Max Planck Institute of Cognitive Neuroscience.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis." *Scientific Studies of Reading*, 5(3), 239–256. doi: 10.1207/S1532799XSSR0503_3

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, 9, 20–28. doi: 10.1177/074193258800900206

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and Pathology in Cognitive Functions*, (pp. 19–76). London: Academic Press. doi: 10.3758/BF03207817

Garrett, M. F., Bever, T. G., & Fodor, J. A. (1966). The active use of grammar in speech perception. *Perception & Psychophysics*, 1, 30–32.

Gee, J. P. & Grosjean, F. (1983). Performance structures: a psycholinguistics and linguistic appraisal. *Cognitive Psychology*, 15, 411–458. doi: 10.1016/0010-0285(83)90014-2

Gleitman, L. R. & Wanner, E. (1982). Language acquisition: The state of the state of the art. In L. R. Gleitman & A. A. Wanner (Eds.), *Language Acquisition: The State of the Art*, (pp. 3–48). Cambridge, MA: Cambridge University Press.

Grosjean, F. & Gee, J. P. (1987). Prosodic structure and spoken word recognition. *Cognition*, 25, 135–155. doi: 10.1016/0010-0277(87)90007-2

Harley, B., Howard, J., & Hart, D. (1995). Second language processing at different ages: Do younger learners pay more attention to prosodic cues to sentence structure?" *Language Learning*, 45(1), 43–71. doi: 10.1111/j.1467-1770.1995.tb00962.x

Harriman, J. & Buxton, H. (1979). The influence of prosody on the recall of monaurally presented sentences. *Brain and Language*, 8(1), 62–68. doi: 10.1016/0093-934X(79)90040-3

Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (Eds.), *Rhythm and Meter*, (pp. 201–260). New York: Academic Press. doi: 10.1016/b978-0-12-409340-9.50013-9

Hirose, Y. (2003). Recycling prosodic boundaries. *Journal of Psycholinguistic Research*, 32(2), 167–195. doi: 10.1023/A:1022448308035

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, 54(3), 425–343. doi: 10.1016/j.jml.2005.12.001

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for prelinguistic infants. *Cognition*, 26, 269–286. doi: 10.1016/S0010-0277(87)80002-1

Holmes, V. M. (1995). A crosslinguistic comparison of the production of utterances in discourse. *Cognition*, 54(2), 169–207. doi: 10.1016/0010-0277(94)00635-X

Hwang, H. & Schafer, A. J. (2009). Constituent length affects prosody and processing for a dative NP ambiguity in Korean. *Journal of Psycholinguistic Research*, 38(2), 151–175. doi: 10.1007/s10936-008-9091-1

Hwang, H. & Steinhauer, K. (2011). Phrase length matters: The interplay between implicit prosody and syntax in Korean "garden path" sentences. *Journal of Cognitive Neuroscience*, 23(11), 3555–3575. doi: 10.1162/jocn_a_00001

Ito, J. & Mester, A. (1992). Weak layering and word binarity. LRC-92-09. Linguistics Research Center.

Jenkins, J. R., Fuchs, L. S., Espin, C., van den Broek, P., & Deno, S. L. (2000). Effects of task format and performance dimension on word reading measures: Criterion validity, sensitivity to impairment, and context facilitation. In *Paper Presented at Pacific Coast Research Conference*. San Diego.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159–207. doi: 10.1006/cogp.1999.0716

Kadota, S. (1982). Some psycholinguistic experiments on the process of reading comprehension. *Journal of Assumption Junior College*, 9, 49–70.

Kadota, S. & Tada, M. (1992). Eibun Oyobi Nihonbun No Dokkai to Shoritani [Reading comprehension and processing units in English and Japanese]. *Annual Bulletin of Research Institute for Social Science*, 22, 137–153.

Kadota, S., Yoshida, S., & Yoshida, H. (1999). Dokkai Niokeru shoritani—Eibun No Teijitani Garikaido Oyobi Shorijikan Ni Oyobosu Eikyo [Processing units in EFL reading: An effect of presentation units on comprehension rate and time]. *Annual Review of English Language Education in Japan*, 10, 61–71.

Kang, S. & Speer, S. R. (2002). Prosody and clause boundaries in Korean. In B. Bel & I. Marlien (Eds.), *Speech Prosody 2002*, (pp. 419–421). Aix-en-Provence: Laboratoire Parole et Langage.

Kang, S. & Speer, S. R. (2005). Effects of prosodic boundaries on syntactic disambiguation. *Studia Linguistica*, 59(2–3), 244–258. doi: 10.1111/j.1467-9582.2005.00128.x

Kang, T. J. & Muter, P. (1989). Reading dynamically displayed text. *Behavior & Information Technology*, 8(1), 33–42. doi: 10.1080/01449298908914536

Katz, W. F., Beach, C. M., Jenouri, K., & Verma, S. (1996). Duration and fundamental frequency correlates of phrase boundaries in productions by children and adults. *The Journal of the Acoustical Society of America*, 99(5), 3179–3191. doi: 10.1121/1.414802

Kentner, G. (2012). Linguistic rhythm guides parsing decisions in written sentence comprehension. *Cognition*, 123(1), Elsevier B.V., 1–20.

Kjelgaard, M. M. & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 40(2), 153–194. doi: 10.1006/jmla.1998.2620

Koriat, A., Greenberg, S. N., and Kreiner, H. (2002). The extraction of structure during reading: Evidence from reading prosody. *Memory & Cognition*, 30(2), 270–280. doi: 10.3758/BF03195288

Koster, M. & Cutler, A. (1997). Segmental and suprasegmental contributions to spoken-word recognition in Dutch. In G. Kokkinakis, N. Fakotakis, & E. Dermatas (Eds.), *Proceedings of Eurospeech '97* (pp. 2167–2170). Patras, Greece: WCL, University of Patras.

Kreiner, H. (2005). The role of reading prosody in syntactic and semantic integration: Evidence from eye movements. In *International Symposium on Discourse and Prosody as a Complex Interface*, 1–17.

Kuhn, M. R. & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95(1), 3–21. doi: 10.1037/0022-0663.95.1.3

Ladd, D. R. (1980). *The Structure of Intonational Meaning*. Bloomington, IN: Indiana University Press.

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa*, 7, 107–122. doi: 10.1121/1.1982702

Leonard, L. B. (1973). The role of intonation in the recall of various linguistic stimuli. *Language and Speech*, 16, 327–335.

LeVasseur, V. M., Macaruso, P., Conway Palumbo, L., & Shankweiler, D. (2006). Syntactically cued text facilitates oral reading fluency in developing readers. *Applied Psycholinguistics*, 27, 423–445. doi: 10.1017/s0142716406060346

LeVasseur, V. M., Macaruso, P., & Shankweiler, D. (2008). Promoting gains in reading fluency: A comparison of three approaches. *Reading and Writing*, 21, 205–230. doi: 10.1007/s11145-007-9070-1

Liljestrand Fultz, A. (2009). Prosody in lexical and syntactic disambiguation in English-French interlanguage. Indiana University.

Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, 53(2), 155–180. doi: 10.1016/0010-0277(94)90069-8

Marslen-Wilson, William D, Lorraine K Tyler, Paul Warren, Philippe Grenier, and Catherine S Lee. (1992). Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology*, 45, 73–87. doi: 10.1080/14640749208401316

Maynell, L. (1999). Effect of pitch accent placement on resolving relative clause ambiguity in English. In *Poster Presented at the 12th Annual CUNY Conference on Human Sentence Processing*. New York, NY.

Maynell, L. (2000). Prosodic effects on relative clause attachment. In *Poster Presented at the 13th Annual CUNY Conference on Human Sentence Processing*. La Jolla, CA.

McDaniel, D., McKee, C., & Garrett, M. F. (2010). Children's sentence planning: Syntactic correlates of fluency variations. *Journal of Child Language*, 37(1), 59–94. doi: 10.1017/S0305000909009507

Mehler, J., Peter Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143–178. doi: 10.1016/0010-0277(88)90035-2

Morgan, J. L. (1996). A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language1* 35, 666–688. doi: 10.1006/jmla.1996.0035

Morgan, J. L. & Demuth, K. (1996). Signal to syntax: An overview." In J. L. Morgan & K. Demuth (Eds.), *From Signal to Syntax: Bootstrapping From Speech to Grammar in Early Language Acquisition* (pp. 1–22). Mahwah, NJ: Erlbaum.

Nathan, R. G. & Stanovich, K. E. (1991). The causes and consequences of differences in reading fluency. *Theory Into Practice*, 30(3), 176–184. doi: 10.1080/00405849109543498

Nazzi, T. & Ramus, R. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41(1), 233–243. doi: 10.1016/S0167-6393(02)00106-1

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.

Nicol, J. & Pickering, M. J. (1993). Processing syntactically ambiguous sentences: Evidence from semantic priming. *Journal of Psycholinguistic Research*, 22(2), 207–237.

O'Connell, D. C., Turner, E. A., & Onuska, L. A. (1968). Intonation, grammatical structure, and contextual association in immediate recall. *Journal of Verbal Learning and Verbal Behavior*, 7, 110–116. doi: 10.1016/S0022-5371(68)80173-2

O'Shea, L. J. & Sindelar, P. T. (1983). The effects of segmenting written discourse on the reading comprehension of low- and high-performance readers. *Reading Research Quarterly*, 18(4), 458–465. doi: 10.2307/747380

Paige, D. D., Rasinski, T. V., Magpuri-Lavell, T., & Smith, G. S. (2014).

Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research*, 46(2): 123–156. doi: 10.1177/1086296X14535170

Paris, C. R., Thomas, M. H., Gilson, R. D., & Kincaid, J. P. (2000). Linguistic cues and memory for synthetic and natural speech. *Human Factors*, 42(3), 421–431. doi: 10.1518/001872000779698132

Pierrehumbert, J. B. & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in Communication*, 271–311. Cambridge, MA: MIT Press.

Pilon, R. (1981). Segmentation of speech in a foreign language. *Journal of Psycholinguistic Research*, 10(2), 113–122. doi: 10.1007/BF01068032

Pratt, E. (2015). Is cue-based memory retrieval "good-enough?": Agreement, comprehension, and implicit prosody in native and bilingual speakers of English. City University of New York.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, 90(6), 2956–2970. doi: 10.1121/1.401770

Rasinski, T. V., Padak, N., Linek, W., & Sturtevant, E. (1994). Effects of fluency development on urban second-grade readers. *The Journal of Educational Research*, 87(3), 158–165. doi: 10.1080/00220671.1994.9941237

Rayner, K. & Pollatsek, A. (1989). *The Psychology of Reading*. Englewood Cliffs, NY: Prentice-Hall.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, Jr., C. (2012). *Psychology of Reading*. 2nd Edition. New York: Psychology Press.

Reber, A. S. & Anderson, J. R. (1970). The perception of clicks in linguistic and nonlinguistic messages. *Perception and Psychophysics*, 8(2), 81–89. doi: 10.3758/BF03210179

Rosner, B. S., Grabe, E., Nicholson, H. B. M., Owen, K., & Keane, E. L. (2004). Prosody, memory load, and memory for speech. *Oxford University Working Papers in Linguistics*, Philology and Phonetics.

Schafer, A. J. (1997). Prosodic parsing: The role of prosody in sentence comprehension. University of Massachusetts, Amherst.

Schafer, A. J., Carter, J., Clifton, Jr. C. & Frazier, L. (1996). Focus in relative clause construal. *Language and Cognitive Processes*, 11(1–2), 135–164. doi: 10.1080/016909696387240

Schafer, A. J., Speer, S. R., & Warren, P. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In M. Tanenhaus & J. Trueswell (Eds.), *Approaches to Studying World Siutation Langauge Use: Psycholinguistic, Linguistic, and Computational Perspectives on Bridging the Product and Action Tradition* (pp. 209–225). Cambridge, MA: MIT Press.

Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29(2), 169–182. doi: 10.1023/A:1005192911512

Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of Literacy Research*, 12(3), 177–186. doi: 10.1080/10862968009547369

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory Into Practice*, 30(3), 158–164. doi: 10.1080/00405849109543496

Selkirk, E. O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, E. O. (2003). The prosodic structure of function words. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, (pp. 187–213). Mahwah, NJ: Erlbaum.

Selkirk, E. O. & Tateishi, K. (1988). Constraints on minor phrase formation in Japanese. In G. Larson & D. Brentari (Eds.), *Proceedings of the 24th Annual Meeting of the Chicago Linguistics Society* (pp. 316–336). Chicago: Chicago Linguistics Society.

Shattuck-Hufnagel, S. & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2), 193–247. doi: 10.1007/BF01708572

Slowiaczek, L. M. (1990). Effects of lexical stress in auditory word recognition. *Language and Speech*, 33(1), 47–68.

Slowiaczek, M. L. & Clifton, Jr., C. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior*, 19(5), 573–582. doi: 10.1016/S0022-5371(80)90628-3

Snedeker, J. & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 48(1), 103–130. doi: 10.1016/S0749-596X(02)00519-3

Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: evidence from prelinguistic infants. *Journal of Memory and Language*, 49(2), 249–267. doi: 10.1016/S0749-596X(03)00024-X

Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, 25(2), 249–271. doi: 10.1007/BF01708573

Speer, S. R., Shih, C.-L., & Slowiaczek, M. L. (1989). Prosodic structure in language understanding: Evidence from Tone Sandhi in Mandarin. *Language and Speech*, 32(4), 337–354. doi: 10.1121/1.2027247

Staub, A. (2007). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 550–569. doi: 10.1037/0278-7393.33.3.550

Stine, E. L. & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging*, 2(3), 272–279. doi: 10.1037/0882-7974.2.3.272

Temple, L. (2000). Second language learner speech production. *Studia Linguistica*, 54(2), 288–297. doi: 10.1111/1467-9582.00068

Wakefield, J. A., Doughtie, E. B., & Yom, B.-H. L. (1974). Identification of structural components of an unknown language. *Journal of Psycholinguistic Research*, 3, 262–269. doi: 10.1007/bf01069242

Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science*, 32(7), 1232–1244. doi: 10.1080/03640210802138755

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91(3), 1707–1717. doi: 10.1121/1.402450

Wingfield, A. (1975). Acoustic redundancy and the perception of time-compressed speech. *Journal of Speech and Hearing Research*, 18, 96–104. doi: 10.1044/jshr.1801.96

Wingfield, A. & Klein, J. F. (1971). Syntactic structure and acoustic pattern in speech perception. *Perception and Psychophysics*, 9(1(A)), 23–25.

Yamashita, J. & Ichikawa, S. (2010). Examining reading fluency in a foreign language: Effects of text segmentation on L2 readers. *Reading in a Foreign Language*, 22(2), 263–283.

# 17 Semantic-Pragmatic Processing

PETRA B. SCHUMACHER

University of Cologne

## Introduction

This chapter will review some of the language comprehension research on sentence- and text-level phenomena. A central question is how we arrive at a meaningful interpretation and more specifically at the interpretation intended by the speaker. A starting point for this is Grice's (1989) theory of meaning, in which he distinguishes two notions of meaning: sentence meaning versus speaker's meaning. The latter is a reconstruction the speaker's potential intentions and often requires the hearer to draw inferences, enrich interpretation, and consider alternative expressions the speaker could have used. Accordingly, sentence and text comprehension pose specific demands on the processing system. For instance during reference processing, one of the core tasks of the processor is to keep track of referents. A referent is a person, object, or concept that is designated by a linguistic expression and to which this expressions refers. For instance, in *Henriette is happy,* the expression *Henriette* refers to a particular individual whose name is Henriette. In ongoing discourse, hearers are confronted with many different referential expressions and the question arises how the processing system manages referents—for example, Henriette may be referred to by *Henriette, she, the girl,* and so on—and which cues guide the process of reference resolution—for example, in the case of referential ambiguity such as in *Tim pinched Tom after he woke up*, where the pronoun *he* can refer to Tim or Tom. Another task of the language processor is to determine what is meant by a particular utterance: the information a speaker intends to communicate to a hearer may not be explicitly articulated, requiring the hearer to perform additional computations to extract the speaker's meaning. For example, a speaker uttering *Some cookies are burned* may in fact imply that *some but not all cookies are burned.* The comprehender's task is to decipher these implicit meaning aspects and determine what the speaker actually meant. In the following examples, reference resolution and

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

pragmatic processing will be showcased to illustrate the kinds of computations and inferences required during language comprehension.

What these phenomena have in common is that they can be viewed within a model of communication where speakers and hearers cooperate with each other on the basis of conversational principles (cf. Grice, 1989). Speakers do not only rely on lexically coded meaning and grammatical principles when they produce an utterance, and hearers are aware of the subtle conversational principles that guide speakers' choices. In Grice's seminal work, speaker's meaning is obtained from the general principles of how speakers and hearers utilize language for successful communication. A speaker may use a particular form to be optimally informative, to avoid redundancy or to lower processing effort. A hearer in turn draws inferences to arrive at the speaker's meaning.

## Reference resolution

In human communication, referents are introduced and referred back to continuously. For example in the following fable by Aesop in (1) below, the expression *an Arab Camel-driver* introduces a new referent into the mental representation, as well as the expressions *the lading* and *his Camel*.

(1) An Arab Camel-driver having completed the lading of his Camel, asked him which he would like best, to go up hill or down hill. The poor beast replied, not without a touch of reason: "Why do you ask me? Is it that the level way through the desert is closed?" (Aesop, translated by Worthington, 2009)

In subsequent discourse, other referential expressions, such as pronouns (*he, him*) or definite descriptions (*the poor beast*), can be used to reactivate and refer back to these referents. While pronouns or repetitions reactivate a previously introduced referent, expressions like *the poor beast* also activate additional properties about the respective referent. This distinction between new and given information will be discussed in more detail in the next section. Furthermore, speakers choose among different referential forms (pronouns, full noun phrases, etc.) with the aim of optimizing information transfer. They indicate the relative prominence of a referent within the current discourse (by using a particular form, prosodic realization or syntactic structure) and make assumptions about the cognitive status of a referent in the hearer's mind. Typically, a pronoun or null form (if available in the language's repertoire) is used when the referent is prominent and easily accessible for the hearer, and more specified forms are used for less prominent referents (Gundel *et al.*, 1993; Almor, 1999; Ariel, 2001). We return to the issue of referential prominence further below.

### *Information status*

Let's first look at the distinction between new and given referents and its consequences for language comprehension. Traditionally, the given-new distinction has been viewed as a dichotomy (cf. e.g., Chafe, 1976). In languages that have definiteness marking, the introduction of new referents has been correlated with indefinite marking (e.g., *a camel-driver*) and the reactivation of given referents with definiteness (e.g., *the camel-driver*) (see also Heim, 1982). However, such a rigid differentiation does not account for definite expressions like *the lading* in (1) where the use of the definite determiner indicates that the referent is somehow known to the hearer but the referent also represents a discourse-new entity. It is apparent that *the lading* is semantically associated with *camel-driving* so that an indirect semantic link between the *camel-driver* and the *lading* can be established. It is thus indirectly activated in the mental representation of the hearer and this semantic link licenses the definiteness marking of the referential expression. This latter type of definite expression has obtained numerous labels in the theoretical literature such as bridging, inferrables, associative, and indirect anaphors (Clark, 1975; Prince, 1981b; Heim, 1982; Schwarz-Friesel, 2011). These partly new and partly given entities indicate that givenness must be viewed as a continuum and they represent an important test case for the investigation of the comprehension of information status.

Using a self-paced sentence-by-sentence reading task, Haviland and Clark (1974) showed that target sentences containing a directly given (coreferential) expression (*the beer* in (2a)) were read faster than sentences containing an indirectly related expression (2b). These faster reading times demonstrate a processing advantage of given over new/inferred information. Similar results were obtained using a cued recall task, with faster recall rates for given referents (McKoon & Ratcliff, 1980). This processing advantage does not arise from mere repetition of the head noun but reflects facilitation from a coreferent available in the mental representation, which is also supported by studies testing synonymous, coreferring expressions (Yekovich & Walker, 1978).

(2) a. We got some beer out of the trunk. <u>The beer</u> was warm.
    b. We checked the picnic supplies. <u>The beer</u> was warm.

A nuanced understanding of the processing of information status can be obtained using online methods such as event-related brain potentials (ERPs). ERPs represent small voltage fluctuations that reflect spontaneous neural activity that arises in response to cognitive, motor or sensory events. They are time-locked to a stimulus (e.g., a referential expression) and provide a high temporal resolution of the underlying cognitive processes. The activity is measured in a non-invasive manner by means of electrodes positioned on the scalp. Now, ERP studies on direct and indirect anaphors indicate that information status contributes to two discrete processes. Comparing the reading comprehension of a definite expression (*the professor* in (3)) following different context sentences, the indirectly related expression (3b) elicited more processing effort than the coreferential expression (3a).

(3) a. Zoe visited a professor in Berlin. She said that <u>the professor</u> was very inspiring.
    b. Zoe visited a lecture in Berlin. She said that <u>the professor</u> was very inspiring.
    c. Zoe met Patricia in Berlin. She said that <u>the professor</u> was very inspiring.

Time-locked to the onset of the referential expression (*the professor*), this effort was reflected in two effects, a pronounced negative deflection with a peak latency around 400 ms after the onset of the critical expression (so-called N400) and an enhanced Late Positivity between roughly 500–800 ms post-onset (Burkhardt, 2006). These data indicate that the givenness advantage observed in previous offline measures is generated by two discrete cognitive processes. The N400 difference is associated with the effort required for accessing the referent, which is more enhanced when extra inferential work is needed to link *professor* with the previously mentioned *lecture.* The Late Positivity is attributed to processing demands arising from the introduction of a new mental representation. The biphasic pattern and the discreteness of the two processes is corroborated by a third condition in which the definite expression cannot be linked or bridged to previous discourse and is thus discourse- and hearer-new (3c). The ERPs for this condition showed an even more pronounced N400 (3a < 3b < 3c)—that is, more demands when trying to access a referent for the definite expression—and a Late Positivity that has the same amplitude as the indirect anaphor (3a < 3b/c)—suggesting that mental modal updating costs equally accrue for indirect anaphors and new referents.

Chafe (1976) and others observed that information status is also marked prosodically. In West-Germanic languages, given referents are typically deaccented and new referents receive a pitch accent to indicate their information status to the hearer. Again, intermediary prosodic realizations have been reported for different types of indirect anaphors (e.g., part-whole relations or scenario-based relations) (for English and German accent types, respectively, see Pierrehumbert & Hirschberg, 1990; Baumann & Grice, 2006). Building on the production study of Baumann and Grice (2006) who showed that a high falling accent with a low target on the accented syllable (H + L*) is the most appropriate accent type for whole-part relations (e.g., on *the sole* in (4)), it was shown that deviations from this accent pattern were penalized and resulted in processing cost (Schumacher & Baumann, 2010).

(4) Sabine repairs an old shoe. In doing so, she cuts <u>the sole.</u>

## *Referential prominence*

Givenness is one aspect that the processing system uses to determine the relative prominence of a referent. During discourse processing, the system is confronted with an increasing amount of information but has limited resources for storing. It is therefore assumed that referents are stored in a certain ranked order. The notion of prominence may then serve as a relational property between referents that singles out one referent from the set of referents (i.e., the most prominent referential candidate). Two crucial questions arise from this. First, which features contribute to a referent being prominent in discourse? Second, how does the speaker refer to referents of different prominence in the ongoing discourse?

An inverse relation between the prominence of a referent in the mental representation and the explicitness of a particular form used by the speaker has been proposed in linguistic research (Gundel *et al.*, 1993; Almor, 1999; Ariel, 2001). Accordingly, the speaker chooses an unstressed pronoun or null form when referring to the most prominent referent in the current set and uses a more complex form when referring to a less prominent entity (for an overview over speakers' choices in language production see Arnold, 2010). The hearer in turn uses these form-specific constraints during reference tracking.

A theory that seeks to implement some of these insights is Centering Theory, which predicts that the most prominent entity in prior discourse is picked up by a pronoun, where prominence is a function of grammatical role in English (Grosz *et al.*, 1995). This prediction is borne out by a self-paced sentence-by-sentence reading study, where the use of the proper name (*George*) evoked longer reading times in (5) in contrast to the use of the less explicit pronominal form (Gordon *et al.*, 1993). Processing costs for the more specified expression were also observed for other forms such as repeated definite descriptions vs. pronouns and in ERP studies (e.g., Almor, 1999; Swaab *et al.*, 2004).

(5) George jumped out from behind a tree and frightened Debbie. He was surprised at her hysterical reaction. He/George never thinks about how others might feel.

Findings like these make personal pronouns viable candidates for assessing the role of different factors on reference resolution. Such an endeavor allows for three different kinds of outcomes: i) a single feature may be identified that determines referential prominence, ii) multiple weighted features may interact with each other, or iii) form-specific mappings may apply. The first scenario can be discarded on the basis of language processing data that reveal interactions between prominence-lending features. Numerous features have been reported to contribute to a referent's prominence, including morphosyntactic features (see also the contribution by Nicol and Barss, this volume), grammatical function, linear order, distance, agentivity, animacy, topicality, givenness, coherence structure (Clark & Sengul, 1979; Chambers & Smyth, 1998; Arnold *et al.*, 2000; Kehler *et al.*, 2008; Schumacher *et al.*, 2016, among many others). One challenge for this line of research is that some of the features are tightly connected, for instance subject, topic, given, animate, and agent are often aligned, which renders the disentanglement of features rather difficult. Moreover, the question of which features influence reference resolution and how they interact with each other has only partially been answered since experimental research cannot assess the entire inventory of factors in a single study. As a result, we only have a partial understanding of prominence constraints. Note however that most of this research has been carried out with contexts that make available two competing referential candidates. In more elaborate texts, topics appear to be privileged candidates for prominence (e.g., Kaiser & Trueswell, 2008). There also seems to be an asymmetry in the interpretive preferences for different types of referring expressions as suggested by form-specific accounts. This is certainly the case for reflexives (versus pronouns) that are subject to locality constraints (e.g., Burkhardt, 2005; Kaiser *et al.*, 2009) but it has also been proposed for personal vs. demonstrative pronouns in English and Finnish (Brown-Schmidt *et al.*, 2005; Kaiser & Trueswell, 2008).

In addition to the choice of referential forms, the speaker can use prosodic cues to give the hearer clues about the status of a referent. We have already discussed one example for this in (4) with regard to different prosodic realizations of the degree of givenness. Another example are personal pronouns that may come in stressed or unstressed form, with the former being the more specified form, which indicates that the privileged referent is not the intended target. This was confirmed by a referent identification study that presented pronouns with and without contrastive stress in contexts such as (6) (Balogh, 2004).

(6) Excited by their costumes for the Halloween play, some of the third graders started rough-housing back stage. An alien pinched an acrobat just behind the curtain and a ghost pinched her/HER near the backdrop. Soon the whole audience heard the giggling back stage.

In the unmarked, unaccented form, the pronoun was taken to refer to the *acrobat* in the majority of cases (which is explained by effects of parallel sentence structure between the two conjoined clauses; e.g., Chambers & Smyth, 1998; Streb *et al.*, 1999). When the pronoun received a pitch accent, interpretive preferences shifted to *the alien.*

Balogh (2004) followed up on this finding with a cross-modal lexical decision priming task which probes referential processes in real-time. In this paradigm, participants perform two tasks: they listen to texts like (6) and make a lexical decision to a visually presented probe word (word vs. non-word decision). This word is presented right after pronoun-offset and is either semantically related to one of the antecedents (e.g., *space* for *alien, circus* for *acrobat*) or unrelated (length and frequency controlled *union* or *subway* for (6)). Priming describes a mechanism by which the activation of a related word shows facilitation, that is, faster lexical decision times. Applied to pronoun resolution, the preferred referential candidate should evoke a priming effect (Nicol & Swinney, 1989). The data showed that for (6) with an unaccented pronoun, only the probe word related to *acrobat* showed faster reaction times; in contrast, with the accented pronoun, a priming effect occurred for the word related to *alien*. This confirmed the referential choices made offline and further revealed that referential decisions are made immediately at the pronominal expression.

Differences in interpretive preferences have also been observed for other comparisons of referential forms such as unmarked personal pronouns vs. more specified demonstrative pronouns. Characterizations typically state that the demonstrative is the marked choice and excludes the most prominent referent (e.g., Comrie, 1997). Brown-Schmidt and colleagues (2005) investigated instructions like (7) and found that the personal pronoun (*it*) showed an interpretive bias for the initially introduced entity (*the cup*), while the demonstrative *that* was preferable resolved toward the composite (*the cup on the saucer*).

(7) Put the cup on the saucer. Now put it/that over by the lamp.

In German, demonstrative pronouns can also refer to animate referents and are typically used to indicate that the most prominent referent does not qualify as a referent. Previous research has suggested an anti-subject or anti-topic bias for the demonstrative in German (e.g., Bosch & Umbach, 2007). Schumacher *et al.* (2016) elicited the referential preferences for the personal pronoun *(er)* and the demonstrative pronoun (*der*) in sentence completion and referent selection tasks. They used two types of verbs in order to disentangle the contribution of grammatical role and thematic role on reference resolution. Active accusative verbs as in (8) canonically unite the highest grammatical and thematic role on one argument (*the magician* is the subject and the agent). Dative experiencer verbs as in (9) cross these two features, that is, *the singer* represents the highest thematic role (here the experiencer role) but is the grammatical object, while *the dancer* holds the highest grammatical role (subject) but the lower thematic role (stimulus). Note also that (9) reflects the basic argument order in connection with dative experiencer verbs.

(8) Der-NOM Zauberer wollte den-ACC umarmen. Aber er/der war viel zu klein.
The magician wanted to hug the doctor. But he-PPRO/he-DEM was way to small.

(9) Dem-DAT Sänger ist der-NOM Tänzer aufgefallen. Aber er/der will die Feier sehen.
The singer has noticed the dancer. But he-PPRO/he-DEM wants to watch the ceremony.

The investigation of interpretive preferences following contexts with dative experiencer verbs makes it possible to determine whether grammatical function or thematic role is the guiding cue during pronoun resolution. The studies registered robust preferences of the personal pronoun for the highest thematic role (*the magician* in (8) and *the singer* in (9)) and a complementary preference of the demonstrative for the lower thematic role (*the doctor* in (8) and *the dancer* in (9)).

Referential prominence can be further mediated by syntactic structure. Concerning information status, cross-linguistic data reveal that speakers tend to place given information before new information (given-new ordering). In languages with flexible word order (e.g., Finnish), the given-new linearization preference may give rise to object-initial constructions. Thus when confronted with an initial, previously introduced object, hearers subsequently expect a new referent. Measuring eye movements on a visual display during auditory comprehension (in Finnish), participants showed anticipatory looks toward the new referent even before the referential expression had been articulated (Kaiser & Trueswell, 2004). For example for (10), there was a visual display showing a doctor, a nurse and a patient.

(10) On the hospital reception desk are leaning a doctor and a nurse, and it is almost two o'clock. After a moment,

a. doctor–*object* glances-at patient–*subject*. (=The patient glanced at the doctor.)
b. doctor–*subject* glances-at patient–*object*. (=The doctor glanced at the patient.)
This patient is holding a pair of scissors.

Crucially, the context sentence mentioned only two referents (*doctor, nurse*) rendering these two discourse-given. Using the object-initial structure (which is clearly indicated by case morphology in Finnish) (10a) may be understood as a means to express given-new ordering and accordingly a discourse-new entity (*patient*) should be expected rather than a previously given entity (*nurse*). Indeed, anticipatory looks to the unmentioned entity (*patient*) were observed after verb offset and crucially before the acoustic information of the second nominal expression could have been processed. This suggests that hearers use the given-new ordering to make incremental decisions as the sentence unfolds. In the canonical subject-initial ordering (10b), no such predictions were derived.

Other ordering constraints following from the prominence cues introduced above affect referent linearization as well (cf. e.g., Bornkessel-Schlesewsky & Schlesewsky, 2009). Topicality (i.e., what an utterance is about; Reinhart, 1981) is one of the features influencing word order, because topics most often occur sentence-initially. Topicality typically presupposes the givenness of the respective referent but may be more restrictive in certain languages, like Mandarin Chinese, where the ideal topic combines features of givenness, agentivity, and animacy (Givón, 1983). A series of ERP studies revealed position-specific effects during the processing of referential expressions, indicating that entities in topic and non-topic positions are subject to distinct constraints. For Mandarin Chinese, referents in sentence-medial position confirmed the effects of givenness reported above for German (biphasic N400-Late Positivity for indirectly related referents compared to given referents; Burkhardt, 2006). Importantly, sentence-initial referents were processed on the basis of constraints on topicality, with a penalty for inanimate referents as well as for referents that induce a topic shift (Hung & Schumacher, 2012, 2014). In these studies, differences in the Late Positivity were observed when a less prominent referent occurred in topic position. This effect is taken to reflect the assessment of constraints on information packaging.

So far, this section has looked at the "backward" orientation of referential expressions, that is a pronoun or definite expression links up with a referent in prior discourse. However, referential expressions also carry a forward potential. The speaker can signal the hearer whether a certain referent is maintained in future discourse or whether a shift in the referential prominence ranking will occur. Demonstratives are powerful means to indicate such a shift and certain indefinites also have an anticipatory function (for so-called indefinite *this* see Prince, 1981a). By using these referential forms, the speaker provides a cue about the future status of the respective referent, that is, whether it is likely to be rementioned in the particular discourse or even to assume topic status. This forward-looking function has been tested via a story completion task (Gernsbacher & Shroyer, 1989). Participants heard the beginning of a prerecorded story like (11) and were asked to continue this story for 20–30 seconds.

(11) I went to the coast last weekend with Sally. We'd checked the tide schedule and we'd planned to arrive at low tide 'cuz I just love beachcombing. Right off, I found three whole sand dollars. So then I started looking for agates, but I couldn't find any. Sally was pretty busy too. She found this egg/an egg….

A regular indefinite (*an egg*) evoked only few resumptions of the particular referent but indefinite *this* elicited frequent referrals in subsequent discourse, which were also realized by less specified referential forms. Comparable data were reported for German indefinite *this* (Deichsel & von Heusinger, 2011) and for the forward-orientation associated with the demonstrative pronoun *der* in German (Schumacher *et al.*, 2015). This indicates that referential choices also convey information about the developing discourse.

This section tried to demonstrate how manifold the speaker's referential decisions are and how sensitive the hearer's mind is to these reference tracking cues. Prosodic, syntactic, semantic, and form-specific features contribute to the management of referents in ongoing discourse and they are used to guide the identification of the coreferent as well as provide cues for the makeup of the unfolding discourse. In the next section, we turn to examples of meaning constitution that involve implicit meaning constituents.

## Inferences

In Gricean pragmatics, speaker's meaning reflects the reconstruction of the speaker's (potential) intentions. The speaker attempts to make a contribution that is true, informative, relevant, and perspicuous and very often exploits these principles to implicate meaning. The hearer is expected to detect the exploitation or flouting of a conversational principle and infer the speaker's meaning. There is often a trade-off between saying as much as one can and saying no more than required (cf. Horn's Q- and R-Principle; Horn, 1984). A waiter uttering for instance (12) (from Nunberg, 1979) does certainly not mean that the food item is sitting at the table but rather intends to refer to the ham sandwich orderer or eater. The waiter only uses this shortcut to refer to the customer when he assumes that the addressee is able to draw the relevant inference. We will return to this case of meaning shift in the next section.

(12) The ham sandwich is sitting at Table 20.

Similarly, the speaker of (13) may implicate that some but not all students fell asleep. By using the scalar term *some*, she signals that she has no reason to use the stronger, more informative term *all* (Horn, 1972). Note however that scalar implicatures can be canceled (*Today some students fell asleep, in fact all of them did.*). The processing of implicatures will be reviewed below.

(13) Today some students in my class fell asleep.

## *Meaning extension*

The example from (12) is a case of referential transfer and therefore serves as a nice link between research on reference and inferences. It essentially describes a communicative situation in which the principles of perspicuity and brevity are in conflict. The waiter could also say (12′) but in the microcosm of the restaurant setting this amount of information is not required since what a customer has ordered emerges as a salient property and therefore (12) is a more economical way of referring.

(12′) The woman who ordered the ham sandwich is sitting at Table 20.

What are the processing consequences of using a salient property to refer to a person? An ERP study revealed that the comprehension of an expression that indicates reference transfer (*the hepatitis in* (14a)) was more costly than its unshifted control (14b) (Schumacher, 2011). Specifically, it evoked a Late Positive potential, which has already been discussed above for referent management. This is taken to indicate that the representation of *hepatitis* is reconceptualized into a mental representation denoting the person who has hepatitis, and this results in processing costs.

(14) a. The doctor asks his assistant who had called. The assistant responds that <u>the hepatitis</u> had called.
b. The doctor asks his assistant what it is that concerns so many people. The assistant responds that <u>the hepatitis</u> concerns so many people.

(15) a. The ham sandwich wants a coke and #has gone stale.
b. The ham sandwich wants a coke and asks for the check.

The notion of reconceptualization is supported by the coordination diagnostic which demonstrates that the original denotation is no longer available once the shift has taken place. In (15a) (from Copestake & Briscoe, 1995), *the ham sandwich* is shifted to the ham sandwich eater in the first clause and the person-denotating referent is then picked up in the second conjunct yielding an infelicitous utterance (*the ham sandwich eater has gone stale*). Coordination is however possible in (15b) where the person-denotation is intended in both clauses (*The ham sandwich eater wants a coke and the ham sandwich eater asks for the check.*)

These observations have been extended to other cases of meaning transfer (also termed metonymy) including container-for-content (*Tim drank the goblet* versus *Tim dropped the goblet*) and animal-for-statue (*the stone lion* versus *the tired lion*), which also engendered a Late Positivity (Schumacher, 2013). Interesting, not all kinds of meaning shift give rise to such a positivity. Content-for-container (*Ann put the beer in the fridge* versus *Ann drank the beer in the fridge*) and producer-for-product alternations (*Luise read Goethe* versus *Luise met Goethe*) do not show any ERP effects (Schumacher, 2013; Weiland-Breckle & Schumacher, 2017). The different processing profiles among metonymic expressions may be explained on the basis of the conventionality of the meaning shift. The latter two types of alternations occur more frequently in daily conversation and may therefore have become part of lexically coded meaning, while the former cases require more contextual support and a higher amount of inferencing to arrive at the intended meaning. This is supported by eye-tracking research that also reports no differences for producer-for-product alternations and other conventional uses such as place-for-event (*A lot of Americans protested during Vietnam* versus *visited Vietnam*) or place-for-institution (*That woman answered to the convent* versus *These businessmen purchased the convent*) (Frisson & Pickering, 1999, 2007).

Another kind of meaning extension has been investigated widely in the psycholinguistic research where the selectional restriction of the predicate requires a type shift from an object-denoting entity to an event. (16) illustrates this so-called complement coercion, where the predicate (*began*) requires a complement of the type event but is confronted with an entity.

(16) a. The girl began the book.
b. The author began the book.
c. The girl read the book.

This feature mismatch is resolved by the hearer by inferring some kind of activity that can be performed with the book, which typically elicits *reading the book* in (16a) and *writing the book* in (16b). Reading time and eye-tracking studies indicate that the computation of complement coercion evokes extra processing demands compared to a control condition (16c) (McElree *et al.*, 2001, 2006; Traxler *et al.*, 2002). This cost is attributed to structure building processes associated with the type shift. This is further supported by type shifting demands at the phrasal level: Frisson and colleagues (2011) report processing demands for the comprehension of *the difficult mountain* (versus *the difficult exercise*), where the event-modifying adjective triggers a type shift from the entity *mountain* to an event-denotation (such as *the mountain that is difficult to climb*).

Where does the extra event information come from that is required for the type shift? Some researchers have proposed the lexicon as the source, others more general inferences and world knowledge. The lexical approach assumes that fine-grained representations specifying formal and functional properties of concepts guide interpretive processes (cf. Pustejovsky, 1995 on qualia representations that detail the makeup of lexical entries). Accordingly, the lexical entry of *book* specifies its formal and constitutive properties, its function (*being read*, telic role) and its genesis (*being written*, agentive role). Some initial support for the lexical view is provided by corpus data that indicated that the majority of complement coercion cases could be resolved by telic or agentive role information of the complement noun (Lapata *et al.*, 2003). However, a follow-up sentence completion study registered over 35% of responses that could not be accounted for in terms of qualia structure indicating that qualia structure alone cannot account for the interpretive processes. This suggested that context had an impact on the choice of inferred event. The pragmatic approach thus assumes that general inferencing leads to event recovery (de Almeida & Dwivedi, 2008; Katsika *et al.*, 2012).

Another kind of coercion process requires the extension of the aspectual domain. In so-called aspectual coercion, a feature mismatch arises from the punctuality of the verb (*jump*) and the durativity expressed by a temporal modifier (*for an hour*). The implied meaning of (17a) is that the tiger jumped repeatedly because *jumping* is an activity with an inherent punctual beginning and end point, which should be incompatible with a temporal modifier that expresses a time span. Reading time studies reported processing costs for this aspectual extension in comparison to a control condition with a temporally unbounded verb (17b) (Piñango *et al.*, 1999; Brennan & Pylkkänen, 2008).

(17) a. The tiger jumped for an hour.
b. The tiger slept for an hour.

An alternative explanation of the increased reading times for (17a) could be that the processing of iterative events is generally more costly. This was tested by Todorova and colleagues (2000) who compared coercion cases with singular objects (*a large check,* 18a) with a case in which bare plural objects (*large checks,* 18b) overtly license iteration and are hence exempt from coercion. Using a stop-making-sense task, rejection rates as well as reading times were larger for (18a) than (18b) at the temporally unbounded modifier (*for many years*). This suggests that the observed costs are not due to iteration per se.

(18) a. Even though Howard sent a large check to his daughter for many years, she refused to accept his money.
b. Even though Howard sent large checks to his daughter for many years, she refused to accept his money.

The phenomena reviewed so far in this section can be described in terms of the interplay of the conversational principles of brevity and clarity according to which speakers do not say more than they must. Other cases of the extension of an entity or event can be found in rhetorics, most notably metaphor, which is viewed as an exploitation of the principle of truthfulness conjoined with brevity. When Shakespeare's Romeo utters *Juliet is the sun*, the hearer infers some kind of mapping of the involved concepts to arrive at the intended meaning. Another case is approximation: when we here that *France is a hexagon*, we need to minimally adjust the concept of a hexagon to arrive at the shape of France. The psycholinguistic literature on metaphor is extensive and will not be reviewed here. But it should be pointed out that the different types of meaning extensions have typically been investigated in isolation, with a few notable exceptions. Comparison of complement coercion and producer-for-product metonymy revealed processing demands for the former only (McElree *et al.*, 2006). Different processing profiles were observed for metaphor and metonymy (e.g., Gibbs, 1990; Weiland *et al.*, 2014) and for metaphor, metonymy and approximation (Bambini *et al.*, 2013). Future research should systematically highlight commonalities and differences and identify core mechanisms shared during meaning constitution across different phenomena.

## *Implicature*

In the previous section, we looked at implied meaning arising during composition at the syntax-semantics interface. This section will now focus on inferences at a more global level where the sentences satisfy grammatical constraints and the speaker's intention must be inferred from the conversational setting. Scalar implicatures like (13) (*Today some students in my class fell asleep.*) are the most extensively investigated cases of generalized conversational implicatures, that is, pragmatic inferences arising from a general property such as the availability of a scale like <some, all>. This scale reflects the degree of informativeness of the terms and it is assumed that a speaker who chooses the weaker form (*some*) implicates that she has not sufficient information to use the stronger form (Horn, 1972). As a consequence, (13) is taken to implicate *Some but not all of the students in my class fell asleep.*

How are these implicatures processed? Bott and Noveck (2004) found that pragmatic inferencing results in processing costs. They tested sentences like (19) and instructed participants in one session to interpret *some* as *some but not all* (pragmatic condition) and in another session to interpret it as *some and possibly all* (logical condition).

(19) Some elephants are mammals.

Response latencies were longer and accuracy was lower for the pragmatic condition. In another study without this instruction, some participants responded pragmatically (i.e., considered (19) false) and some responded logically (i.e., evaluated (19) true). Their responses also confirmed the processing difference between pragmatic and logical interpretations. Chevallier *et al.* (2008) generalized these findings to another type of scalar implicature by investigating the interpretation of *or*, which comes with an exclusive reading (*flowers or champagne → either flowers or champagne, but not both*) or an inclusive reading (*flowers or champagne → flowers and champagne*). The more informative (exclusive) meaning evoked more processing effort.

Are scalar inferences drawn automatically? Some pragmatic theories assume that the scalar implicature arises automatically (i.e., *some means not all*) but can be canceled subsequently if required by the context (*some, in fact all*) (Levinson, 2000; Chierchia, 2004). Others argue that only the contextually relevant meaning is computed (Sperber & Wilson, 1986). This makes contrary predictions for the time-course of implicature processing. Take for instance example (20) from Breheny *et al.* (2006).

(20) a. Mary asked John whether he intended to host all his relatives in his tiny apartment. John replied that he intended to host some of his relatives. The rest would stay in a nearby hotel.

b. Mary was surprised to see John cleaning his apartment and she asked the reason why. John told her that he intended to host some of his relatives. The rest would stay in a nearby hotel.

In (20a) the context makes available the entire set (*all his relatives*) and therefore strengthens the scalar implicature (*some but not all*). In turn (20b) does not evoke the scalar relation <some, all> explicitly, hence whether the reading *some maybe all* or *some but not all* is the intended meaning is not immediately relevant. The Neo-Gricean account (Levinson, 2000; Chierchia, 2004) predicts that scalar inferences are drawn automatically and their annulation during context updating exerts costs. Hence, reading times of the scalar term (*some of his relatives)* should be slower for (20b). Relevance Theory (Sperber & Wilson, 1986) predicts costs for contextually required scalar implicature, hence slower reading times for (20a). Self-paced reading time measures demonstrate that the scalar term is read significantly slower in (20a) compared to (20b) (Breheny *et al.*, 2006). This was further confirmed by reading times of the following anaphoric expression (*the rest*), which showed longer latencies for (20b), indicating that the underlying scale has not been activated yet but must be computed at the anaphor to find a referent for the set relation. This suggests that scalar inferences are not generated automatically and consume processing resources when contextually required. For similar findings on scalar inferences see Huang and Snedeker (2009); but there is also growing evidence from time-sensitive eye-tracking showing rapid emergence of inferences and an intricate interaction with other factors (Sedivy *et al.*, 1999; Grodner *et al.*, 2010; Breheny *et al.*, 2013).

Another type of implicature, particularized conversational implicature, has received only scarce attention in the psycholinguistic literature. Particularized implicatures arise in certain contexts and rely on common ground between the interlocutors. For example, the speaker's meaning of *It's hard to give a good presentation* in (21) differs dramatically as a function of context (21a-c).

(21) Nick and Paul are taking the same history class. Students in this class have to give a 20 minute presentation to the class on some topic. Nick gave his presentation
 a. and then decided to ask Paul what he thought of it.
 b. and it was truly terrible. He decides to ask Paul what he thought of it.
 c. and it was excellent. He decides to ask Paul what he thought of it.
 Nick: What did you think of my presentation?
 Paul: It's hard to give a good presentation.

In indirect reply (Paul's response), the speaker exploits the principle of relevance and the hearer's task is to identify the contextually relevant meaning of the indirect reply. Holtgraves (1998) investigated the interpretation of indirect answers in stories like (21) and first asked participants to paraphrase what Paul meant. The answers revealed a majority of indirect interpretations, that is, the implicature was generally drawn, with many negative connotations (e.g., *Paul did not like the presentation.*). The negative bias of the inferences can be explained with reference to the principle of politeness from which it can be inferred that an indirect reply is motivated by the speaker's attempt to be considerate and to save the face of the conversational partner (Brown & Levinson, 1987). This in turn predicts that the indirect answer is not warranted in the positive context (21c) and should therefore create some kind of mismatch. This was confirmed by a second study that measured reading times on the indirect reply as well as on a subsequent paraphrase of Paul's intention, which were reliably longer for the positive context (21c) relative to the other two contexts (21a/b).

## Conclusion

Meaning has different facets and mastering lexical meaning alone does not suffice for successful communication. Rather, deciphering the communicative intention of the speaker is a core task of the hearer who is faced with numerous unarticulated meaning constituents and assumptions. "Reading between the lines" is therefore an essential skill of the hearer, which is mediated by general conversational principles that guide speaker-hearer interaction. The hearer thus assumes that the speaker seeks to be informative, relevant, clear, truthful for the course and purpose of the conversation and obeys social principles (i.e., politeness/face management). The principle of informativeness is central to many of the phenomena discussed in this chapter, such as referential choices or scalar inferences. Informativeness is an instantiation of the principle of economy and least effort (cf. speaker's and hearer's economy in Zipf, 1949) and may be derived from a more general cognitive constraint. As experimental research of semantic and pragmatic processing advances, researchers should strive to develop a general model of meaning constitution that looks at the different phenomena of semantic and pragmatic processing in a unified manner.

## REFERENCES

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, 106(4), 748–765. doi: 10.1037/0033-295X.106.4.748

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 29–87). Amsterdam/Philadelphia: John Benjamins.

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, 4(4), 187–203. doi: 10.1111/j.1749-818X.2010.00193.x

Arnold, J. E., Eisenband, J. G., Brown Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76(1), B13–B26. doi: 10.1016/S0010-0277(00)00073-1

Balogh, J. E. (2004). *Pronouns, prosody, and the discourse anaphora weighting approach*. Doctoral dissertation, University of California, San Diego.

Bambini, V., Ghio, M., Moro, A., & Schumacher, P. B. (2013). Differentiating among pragmatic uses of words through timed sensicality judgments. *Frontiers in Psychology*, 4(938), 1–16. doi: 10.3389/fpsyg.2013.00938

Baumann, S., & Grice, M. (2006). The intonation of accessibility. *Journal of Pragmatics*, 38(10), 1636–1657. doi: 10.1016/j.pragma.2005.03.017

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and linguistics Compass*, 3(1), 19–58. doi: 10.1111/j.1749-818X.2008.00099.x

Bosch, P., & Umbach, C. (2007). Reference determination for demonstrative pronouns. ZAS Papers in Linguistics, 48, 39–51. Available online at: http://www.zas.gwz- berlin.de/185.html.

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language*, 51(3), 437–457. doi: 10.1016/j.jml.2004.05.006

Breheny, R., Ferguson, H. J., & Katsos, N. (2013). Investigating the timecourse of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes*, 28(4), 443–467. doi: 10.1080/01690965.2011.649040

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, 100(3), 434–463. doi: 10.1016/j.cognition.2005.07.003

Brennan, J., & Pylkkänen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of Aspectual Coercion. *Brain and Language*, 106(2), 132–143. doi: 10.1016/j.bandl.2008.04.003

Brown-Schmidt, S., Byron, D. K., & Tanenhaus, M. K. (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language*, 53(2), 292–313. doi: 10.1016/j.jml.2005.03.003

Brown, P., & Levinson, S. (1987). *Politeness: Universals in language usage*. Cambridge: Cambridge University Press.

Burkhardt, P. (2005). *The syntax-discourse interface: Representing and interpreting dependency*. Amsterdam/Philadelphia: John Benjamins.

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, 98(2), 159–168. doi: 10.1016/j.bandl.2006.04.005

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–55). New York: Academic.

Chambers, C. G., & Smyth, R. (1998). Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory and Language*, 39(4), 593–608. doi: 10.1006/jmla.1998.2575

Chevallier, C., Noveck, I. A., Nazir, T., Bott, L., Lanzetti, V., & Sperber, D. (2008). Making disjunctions exclusive. *The Quarterly Journal of Experimental Psychology*, 61(11), 1741–1760. doi: 10.1080/17470210701712960

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In *3* (pp. 39–103). Oxford: Oxford University Press.

Clark, H. H., & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition*, 7(1), 35–41.

Clark, H. H. (1975). Bridging. In B. Nash-Webber & R. Schank (Eds.), *Theoretical issues in natural language processing* (pp. 188–193). Yale University Mathematical Society Sciences Board, Cambridge, MA.

Comrie, B. (1997). Pragmatic binding demonstratives as anaphors in Dutch. In M. L. Juge & J. L. Moxley (Eds.), *Proceedings of the 23rd Annual Meeting of the Berkeley Linguistics Society: General session and parasession on pragmatics and grammatical structure* (Vol. 23, pp. 50–61). Berkeley: Berkeley Linguistics Society.

Copestake, A., & Briscoe, T. (1995). Semi-productive polysemy and sense extension. *Journal of Semantics*, 12, 15–67. doi: 10.1093/jos/12.1.15

de Almeida, R. G., & Dwivedi, V. D. (2008). Coercion without Lexical Decomposition: Type-Shifting Effects Revisited. *Canadian Journal of Linguistics*, 53(2–3), 301–326. doi: 10.1353/cjl.0.0026

Deichsel, A., & von Heusinger, K. (2011). The cataphoric potential of indefinites in German. In I. Hendrickx, S. Lalitah Devi, A. Branco, & R. Mitkov (Eds.), *Anaphora Processing and Application* (pp. 144–156). Heidelberg: Springer.

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1366–1383. doi: 10.1037/0278-7393.25.6.1366

Frisson, S., & Pickering, M. J. (2007). The processing of familiar and novel senses of a word: Why reading Dickens is easy but reading Needham can be hard. *Language and Cognitive Processes*, 22(4), 595–613. doi: 10.1080/01690960601017013

Frisson, S., Pickering, M. J., & McElree, B. (2011). The difficult mountain: enriched composition in adjective-noun phrases. *Psychonomic Bulletin & Review*, 18(6), 1172–1179.

Gernsbacher, M. A., & Shroyer, S. (1989). The cataphoric use of the indefinite this in spoken narratives. *Memory & Cognition*, 17(5), 536–540. doi: 10.3758/BF03197076

Gibbs, R. W., Jr. (1990). Comprehending figurative referential descriptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(1), 56–66.

Givón, T. (1983). *Topic continuity in discourse: a quantitative cross-language study*. Amsterdam/Philadelphia: John Benjamins.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17(3), 311–347. doi: 10.1207/s15516709cog1703_1

Grice, H. P. (1989). *Studies in the way of words*. Cambridge: Harvard University Press.

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, 116(1), 42–55. doi: 10.1016/j.cognition.2010.03.014

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2), 203–225.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69(2), 274–307. doi: 10.2307/416535

Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 512–521.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. Doctoral Dissertation, University of Massachusetts, Amherst.

Holtgraves, T. (1998). Interpreting indirect replies. *Cognitive Psychology*, 37(1), 1–27. doi: 10.1006/cogp.1998.0689

Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature in context. In D. Schiffrin (Ed.), *Meaning, form and use in context : linguistic applications* (pp. 11–42). Washington, D.C: Georgetown University Press.

Horn, L. R. (1972). *On the Semantic Properties of Logical Operators*. Doctoral dissertation, UCLA.

Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers:

Insight into the semantics–pragmatics interface. *Cognitive Psychology*, 58(3), 376–415. doi: 10.1016/j.cogpsych.2008.09.001

Hung, Y.-C., and Schumacher, P. B. (2012). Topicality matters: position-specific demands on Chinese discourse processing. *Neuroscience Letters*, 511(2), 59–64. doi: 10.1016/j.neulet.2012.01.013

Hung, Y.-C., and Schumacher, P. B. (2014). Animacy matters: ERP evidence for the multi-dimensionality of topic-worthiness in Chinese. *Brain Research*, 1555, 36–47. doi: 10.1016/j.brainres.2014.01.046

Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, 23(5), 709–748. doi: 10.1080/0169096070 1771220

Kaiser, E., Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition*, 112(1), 55–80. doi: 10.1016/j.cognition.2009.03.010

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2), 113–147. doi: 10.1016/j.cognition.2004.01.002

Katsika, A., Braze, D., Deo, A., & Piñango, M. M. (2012). Complement coercion: type-shifting vs. Pragmatic inferencing. *Mental Lexicon*, 7(1), 58–76. doi: 10.1075/ml.7.1.03kat

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and Coreference Revisited. *Journal of Semantics*, 25(1), 1–44. doi: 10.1093/jos/ffm018

Lapata, M., Keller, F., & Scheepers, C. (2003). Intra-sentential context effects on the interpretation of logical metonymy. *Cognitive Science*, 27(4), 649–668. doi: 10.1016/S0364-0213(03)00035-1

Levinson, S. C. (2000). *Presumptive meanings*. Cambridge: Bradford Book.

McElree, B., Frisson, S., & Pickering, M. J. (2006). Deferred interpretations: Why starting Dickens is taxing but reading Dickens isn't. *Cognitive Science*, 30(1), 181–192. doi: 10.1207/s15516709cog0000_49

McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched composition. *Cognition*, 78(1), B17–B25. doi: 10.1016/S0010-0277(00)00113-X

McKoon, G., & Ratcliff, R. (1980). The Comprehension Processes and Memory Structures Involved in Anaphoric Reference. *Journal of Verbal Learning and Verbal Behavior*, 19, 668–682.

Nicol, J. L., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, 18(1), 5–19.

Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy. *Linguistic and Philosophy*, 3, 143–184.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen & J. M. M. Pollack (Eds.), *Intentions in communication* (pp. 271–323). Cambridge: MIT Press.

Piñango, M. M., Zurif, E., & Jackendoff, R. (1999). Real-time processing implications of enriched composition at the syntax-semantics interface. *Journal of Psycholinguistic Research*, 28(4), 395–414.

Prince, E. F. (1981a). On the inferencing of indefinite-*this* NPs. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 231–250). Cambridge: Cambridge University Press.

Prince, E. F. (1981b). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical pragmatics* (pp. 223–255). Academic, New York.

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge/London: MIT Press.

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 27(1), 53–94.

Schumacher, P. B. (2011). The hepatitis called…: Electrophysiological evidence for enriched composition. In J. Meibauer & M. Steinbach (Eds.), *Experimental*

*Pragmatics/Semantics* (pp. 199–219). Amsterdam: John Benjamins.

Schumacher, P. B. (2013). When combinatorial processing results in reconceptualization: Towards a new approach of compositionality. *Frontiers in Psychology*, 4(677), doi: 10.3389/fpsyg.2013.00677

Schumacher, P. B., & Baumann, S. (2010). Pitch accent type affects the N400 during referential processing. *Neuroreport*, 21(9), 618–622. doi: 10.1097/WNR.0b013e328339874a

Schumacher, P. B., Backhaus, J. & Dangl, M. (2015). Backward- and forward-looking potential of anaphors. *Frontiers in Psychology*, 6, 1746. doi: 10.3389/fpsyg.2015.01746

Schumacher, P. B., Dangl, M., & Uzun, E. (2016). Thematic role as prominence cue during pronoun resolution in German. In A. Holler, & K. Suckow (Eds.), *Experimental perspectives on anaphora resolution* 121–147. Berlin: de Gruyter.

Schwarz-Friesel, M. (2011). Text comprehension as the interface between verbal structures and cognitive memory processes: The case of resolving direct and indirect anaphora. In C. Zelinsky-Wibbelt (Ed.), *Relations between Language and Memory. Organization, Representation, and Processing*. (pp. 293–330). Frankfurt: Lang.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71(2), 109–147. doi: 10.1016/S0010-0277(99)00025-6

Sperber, D., & Wilson, D. (1986). *Relevance: communication and cognition*. Oxford: Basil Blackwell.

Streb, J., Rosler, F., & Hennighausen, E. (1999). Event-related responses to pronoun and proper name anaphors in parallel and nonparallel discourse structures. *Brain and Language*, 70(2), 273–286. doi: 10.1006/brln.1999.2177

Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence for reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience*, 16(5), 717–726. doi: 10.1023/B:JOPR.0000027961.12577.d8

Todorova, M., Straub, K., Badecker, W., & Frank, R. (2000). *Aspectual coercion and the online computation of sentential aspect*. Proceedings of the twenty-second annual conference of the Cognitive Science Society, Philadelphia.

Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language*, 47(4), 530–547. doi: 10.1016/S0749-596X(02)00021-9

Weiland, H., Schumacher, P. B., & Bambini, V. (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Frontiers in Human Neuroscience*, 8, 583. doi: 10.3389/fnhum.2014.00583

Weiland-Breckle, H., & Schumacher, P. B. (2017). Artist-for-work metonymy: Type clash or underspecification? *Mental Lexicon*, 12(2).

Worthington, R. (2009). *Aesop's Fables*. Auckland, N.Z.: The Floating Press Limited.

Yekovich, F. R., & Walker, C. H. (1978). Identifying and using referents in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 17, 265–278.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.

# 18 Comprehension in Older Adult Populations: Healthy Aging, Aphasia, and Dementia

JET M. J. VONK[1], EVE HIGBY[2], AND LORAINE K. OBLER[1]

[1] The Graduate Center, City University of New York
[2] University of California, Riverside

## Introduction

Language comprehension is a much more complex process than it might appear. For most children and adults comprehension is an effortless, even automatic, process that occurs at an astoundingly rapid pace. Yet the efficiency with which the various elements involved in this process converge to result in successful language comprehension can change with advancing age and with changes to the brain's neural networks due to stroke or other pathology.

"Comprehension" has many meanings, of course, including reference to different linguistic levels (word, sentence, or discourse), different modalities (oral or written, for our purposes), and different depths of processing required (literal or metaphoric). Moreover, the problems that arise in comprehension differ rather substantially across the three groups of older adults we report on in this chapter: those who have no apparent brain damage (considered "healthy elderly" here), those with aphasia, and those with dementia, especially probable Alzheimer's disease and primary progressive aphasia.

In healthy aging—unlike in aphasia and dementia—there are virtually no reports of degradation of linguistic representations per se or of the architecture of the linguistic system; however, the cognitive underpinnings for language processes are known to decline with healthy aging, so the research literature has focused on their interaction with language-comprehension performance. The literature on comprehension in dementia, too, has explored the relation between the cognitive underpinnings of language and the linguistic representations, particularly in the semantic realm. In the aphasia literature, by contrast, there is much less reference

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

to the cognitive abilities that may have started to decline before an aphasia-producing incident in an older adult, but it is the more strictly linguistic factors that have interested researchers and neurolinguistic theoreticians (but see Cahana-Amitay & Albert, 2015). Each of the sections of the chapter, thus, is structured to convey a sense of what has been treated in various studies to date.

A definitional question is that of what is meant by "older" adults. Generally, in the literature we review, "older" adults include individuals aged 60 to 85, though some researchers extend the lower end of the range to 55 or even 50 and some the upper end as far as 90 or 95. Some divide older adults into "young-old" (e.g., 55 to 70) and "old-old" (e.g., 71 to 90). It is worth noting that in the aphasia literature, the ages included in the studies we review tend to be in the "young-old adult" range, due to the increased incidence of stroke starting in the 50s, whereas in the studies of dementia (i.e., Alzheimer's disease) the participants tend to be in the "old-old" group, for analogous reasons of common age of onset of the diseases associated with Alzheimer's disease. "Young" adults are generally the control group for studies of language in aging. When studied by researchers based at schools with undergraduates, they are predictably in the 18-to-21-year-old range; when studied by researchers at other institutions, their ages may extend to around 35.

## Comprehension in healthy aging

The literature on language processing in healthy elderly adults and how it differs from that seen in younger adult populations has been characterized by inconsistent findings—a sort of "now you see it, now you don't" problem. Some of the variation in the findings appear to arise due to the use of different tasks. For example, older adults typically do worse than young adults on tasks of written sentence comprehension when tested after the sentences are presented, but not necessarily when online measures are used (e.g., DeDe, Caplan, Kemtes, & Waters, 2004). Similarly, declines in auditory word recognition performance are found in elderly participants in the absence of any hearing loss even though early, automatic lexical and semantic processing are frequently found to be robust across the lifespan (e.g., Federmeier, Van Petten, Schwartz, & Kutas, 2003).

Another reason why we observe aging effects on some language tasks and not on others may be that language tasks typically include a combination of linguistic and non-linguistic factors, including sensory processes (e.g., hearing or vision), cognitive processes (e.g., working memory abilities or the selection of a word among possible candidates), and linguistic processes (e.g., the strength of semantic connections or the size of the lexicon). Each of these domains may be differently affected by the aging process. Thus, in order to obtain a clear picture of performance on language comprehension tasks in healthy older adults, it is necessary to consider the role of each of these domains (sensory, cognitive, and linguistic), including how they differentially affect performance on the task and how they may interact.

## *Sensory, cognitive, and linguistic factors in language comprehension*

Consider the sensory processes required for comprehension. A number of sensory processes decline in healthy older adults, with hearing and visual acuity having particular relevance for different types of language processes. Declines in sensory processes are likely to have repercussions on subsequent stages of processing, potentially masking any age-related differences in linguistic processing per se (Pichora-Fuller, 2003). For instance, hearing loss affects spoken word recognition abilities, particularly in situations of a degraded speech signal, such as speech in noise (e.g., Benichov, Cox, Tun, & Wingfield, 2012), speech at low amplitudes (Tun, Benichov, & Wingfield, 2010), and "glimpsed" speech, in which words contain multiple, short intervals of silence (Kidd & Humes, 2012). In addition, hearing loss affects the ability to recognize spoken words based only on their initial segments (word-onset gating) (Elliott, Hammer, & Evan, 1987). Even older adults without detectable hearing impairments exhibit slowed automatic sensory processing compared to younger adults (Federmeier *et al.*, 2003; Giaquinto, Ranghi, & Butler, 2007). Researchers have had difficulty explaining certain age-related effects on word recognition processes in the absence of apparent hearing loss. This may be due in part to more covert declines in sensory mechanics such as the loss of cochlear nerve connectivity that could potentially have widespread effects prior to the detection of any sensory hearing loss by standard audiometric thresholds (Sergeyenko, Lall, Liberman, & Kujawa, 2013). Similarly, visual processing in older adults may affect performance on visual word recognition tasks at the level of pre-lexical word encoding in the absence of documented visual impairments (Allen, Madden, Weber, & Groth, 1993).

Declining sensory processing in elderly individuals does not affect all types of language tasks uniformly, but rather it interacts with linguistic conditions. When presented in contexts with high semantic predictability, the difference in performance between younger and older adults on spoken word recognition in noise disappears, even for mildly hearing-impaired adults (Benichov *et al.*, 2012; Cahana-Amitay *et al.*, 2016; Lash, Rogers, Zoller, & Wingfield, 2013; Obler, Nicholas, Albert, & Woodward, 1985). The ability to predict an upcoming word through semantic context reduces the reliance on hearing acuity for successful language comprehension. Hearing loss also taxes the cognitive system in older adults, which may be exacerbated in the presence of background or competing noise (Stenfelt & Rönnberg, 2009). Degraded hearing ability in older adults has been associated with increased cognitive decline, potentially leading to both direct and indirect effects of sensory decline on language processing (Lin *et al.*, 2013). Thus, it is important to consider not only how differences in hearing and visual acuity can affect language performance in older and younger adults, but also how sensory processes can interact with language and cognitive processes.

As research on the independence and interdependence of different types of executive functions has evolved in recent years (e.g., Adrover-Roig, Sesé, Barceló, & Palmer, 2012; Miyake *et al.*, 2000), researchers have begun to investigate the relationship between specific cognitive abilities and language processing. Certain cognitive skills, such as executive functions (processes engaged in controlling mental activities such as goal selection and maintenance, planning, sequencing, inhibition, and other supervisory processes) and working memory (the ability to hold recent input in mind and work with it) decline with age (e.g., Baddeley, Baddeley, Chincotta, Luzzi, & Meikle, 2005; Goral *et al.*, 2011; Higby *et al.*, submitted; Park & Hedden, 2001). While it is not always clear precisely which cognitive processes are involved in specific language tasks, correlations between performance on non-linguistic cognitive tasks and certain types of language processing (e.g., Kwong See & Ryan, 1995; Sommers & Danielson, 1999) indicate that cognitive abilities play an integral role in language performance. For example, set-shifting abilities, typically defined as the ability to switch attention between two task sets as required by a task, are a significant predictor of sentence comprehension ability, in particular those with more complex structures like object-relative clauses and sentences containing more than one negative clause (Goral *et al.*, 2011). Furthermore, several studies have implicated working memory storage and processing capacity as crucial components for successful sentence comprehension (Daneman & Merikle, 1996; Kemper & Sumner, 2001; Payne *et al.*, 2014). Van der Linden *et al.* (1999) reported that the significant relationship found between age and language comprehension was mediated by speed of processing, resistance to interference, and working memory, and that working memory mediated the relationship between the other two cognitive processes and language performance.

It is unclear, however, whether the type of working memory capacity that is needed for successful sentence comprehension is the same as that measured by traditional working memory span tests. Waters and Caplan have proposed a division in working memory processes engaged during sentence comprehension into processes involved in online interpretation and integration and those that operate post-interpretively (e.g., Caplan & Waters, 1999; Waters & Caplan, 1996; 2001; 2004). Their proposal aims to explain findings in the aging literature showing that older adults do less well on post-processing sentence comprehension probes compared to younger adults but show no difference in their ability to comprehend sentences when measured using online tools (DeDe *et al.*, 2004; Kemper, Crow, & Kemtes, 2004; Stine-Morrow, Milinder, Pullara, & Herman, 2001).

While the role of working memory has been studied primarily in the sentence comprehension literature, research at the lexical level has considered how declines in inhibitory control processes may impact language processes (Kane, Hasher, Stoltzfus, Zacks, & Connelly, 1994; Kramer, Humphrey, Larish, & Logan, 1994; Sommers & Danielson, 1999; Zacks, 1989). For example, Sommers and Danielson (1999) compared older and younger adults on their ability to recognize words in noise by employing words with different phonological neighborhood densities (the number of words that differ by only one phoneme with the target word). Older adults exhibited greater difficulty than younger adults for words with a higher phonological neighborhood density. Moreover, the performance for both age groups was related to individual differences on two additional inhibition tasks, suggesting that poor inhibition abilities contribute to decreased performance on tasks that elicit more lexical competitors. However, in the same study, Sommers and Danielson (1999) report that when the high neighborhood-density words appeared in a highly constraining semantic context, older adults performed just as well as younger adults. Thus, the relationship between cognitive abilities and language performance is far from transparent, as it appears to depend on several factors, including specific task demands (Caplan & Waters, 2003).

While most of the studies relating inhibitory control abilities and language comprehension have focused on word-level processing, the role of inhibition has recently also been implicated in successful sentence comprehension performance in older adults as well. Yoon *et al.* (2015) reported that older adults with good inhibitory control performed similarly to younger adults when processing complex sentences, in particular, semantically implausible sentences with two clauses containing negation markers, while older adults with poor inhibitory control performed worse than younger adults. Interestingly, overall performance between the two age groups for these complex sentences did not differ; thus, rather than age, it was individual differences in inhibitory control that contributed to good and poor performance on comprehension of sentences that were both syntactically and semantically complex. What this important study indicates is that inhibition may play an important role at both the lexical and sentence levels.

Considering cognitive changes in aging is certainly necessary for understanding how they interact with language functioning. Nevertheless, it is plausible that changes also occur strictly within the linguistic system independent of cognitive factors. One such proposal, the Transmission Deficit Hypothesis (Burke, MacKay, & James, 2000), considers how transmission of information through levels of language processing may be compromised due to weakened connections among linguistic representations. Certain types of changes in older adults' performance, such as an increase in tip-of-the-tongue states[1] during lexical retrieval, point to a deficit in the link between semantic and phonological representations (James & Burke, 2000), whereas semantic processing itself remains quite strong for older adults (e.g., Burke, White, & Diaz, 1987; Federmeier *et al.*, 2003; Laver & Burke, 1993; Madden, Pierce, & Allen, 1993; c.f., Barresi, Nicholas, Connor, Obler, & Albert, 2000). To date, the Transmission Deficit Hypothesis has been tested only for lexical production. It is not implausible that a transmission deficit between phonological and semantic levels of representation could have consequences for language comprehension processes as well. However, evidence from auditory and orthographic priming studies in older adults showed that the priming of semantics, phonology, and orthography facilitates performance similarly for younger and older adults (MacKay, Abrams, & Pedroza, 1999; Taylor & Burke, 2002). Such findings suggest a lack of a deficit in transmission from phonological word-forms to semantics during comprehension processes in older adults.

Even if the linguistic architecture is unaffected by age, there are additional language-use differences that may impact how older and younger adults perform on language tasks. Older adults typically score higher than younger adults on measures of vocabulary size (Verhaeghen, 2003). They also have more years of language experience, which has been argued to influence language-specific factors like lexical frequency effects (Gollan, Montoya, Cera, & Sandoval, 2008). Some studies also report that older adults read more than younger adults, which is related to overall language exposure and has implications for performance on tasks such as visual word identification and the frequency with which individuals encounter certain syntactic constructions associated with written, more than oral, language (Payne *et al.*, 2014). These language-specific factors are likely to contribute to the overall pattern of language processing differences found between younger and older adults, both at the single word and the sentence levels of comprehension.

## *Automatic versus controlled processing*

To account for the patterns seen in language comprehension in healthy older adults, some researchers have proposed that older adults are not impaired on automatic processes but do show declines for controlled processes (Federmeier *et al.*, 2003; Grieder *et al.*, 2012; Harley, Jessiman, & MacAndrew, 2011; Jennings & Jacoby, 1993; Wlotko, Lee, & Federmeier, 2010). This pattern can be seen across various levels of language processing. Tasks that tap into the strength of connections in the semantic network have typically shown that automatic spreading activation of semantic information is preserved in older age (Balota & Duchek, 1988; Gold, Andersen, Jicha, & Smith, 2009; Grieder *et al.*, 2012), as is automatic morphological decomposition (Goral & Obler, 2003; Kavé & Levy, 2005). Early automatic lexical access has been found to be similar for younger and older adults as well (Allen *et al.*, 1993; Federmeier *et al.*, 2003; Giaquinto *et al.*, 2007; Stern, Prather, Swinney, & Zurif, 1991; but see Balota & Duchek, 1988). Additionally, when processing continuous speech, older and younger adults appear to segment the speech stream in similar ways and react similarly to syntactically more complex speech (Wingfield & Lindfield, 1995). Automatic processing of simple syntactic violations, such as number disagreement, is also just as effective in older as in younger adults (Kemmer, Coulson, De Ochoa, & Kutas, 2004).

While many automatic processes indeed show little or no decline with aging, a handful of studies have found that older adults do not necessarily engage certain language processes that are considered quite automatic in younger adults, even though they can demonstrate their ability to do so when prompted to. During comprehension, listeners automatically predict what types of words and structures are likely to occur following a given set of phrases or words. Depending on the type of information provided, some contexts are considered highly constraining, which means that upcoming word predictions tend to be quite specific, or constrained, by the context, while others are low constraining, meaning predictions may be general or encompass a large number of possibilities. This predictive mechanism aids comprehension by pre-activating lexical and phonological information about probable upcoming words. Federmeier, Kutas, and Schul (2010) found that older adults were less likely than younger adults to engage predictive mechanisms to given contextual information such as "a kind of tree." A production task using the same type of contextual information as prompts, however, showed that older adults were able to generate a highly predictive response just as fast as younger adults. Additionally, individuals who showed better predictive patterns in the comprehension task performed better on the word generation task, suggesting that efficiency of access to lexical information may underlie performance in both the production and comprehension tasks. The automaticity of predictive processes in younger adults, then, may decrease with age together with a decrease in the ability to efficiently access lexical items in long-term memory. As automaticity decreases, more cognitive resources may be required and called upon only as needed. Additional support for this idea comes from eye-tracking research showing a reduced ability by older adults to utilize word-onset information during word recognition (Ben-David *et al.*, 2011), a process that is thought to be quite automatic in younger adults.

In contrast to the preservation of many automatic processes in older age, higher-order language processes, such as the construction of message-level meanings during auditory comprehension, take more time in older adults as compared to younger adults (Federmeier *et al.*, 2003). The comprehension of complex syntactic structures, such as object-relative embedded clauses (e.g., Goral *et al.*, 2011) and cleft object constructions, and the resolution of ambiguity also show worse performance by older adults than younger adults (e.g., Caplan, DeDe, Waters, Michaud, & Tripodis, 2011). When encountering ambiguity, older adults appear to have a stronger bias than younger adults toward one resolution possibility (Payne *et al.*, 2014) and are less able to suppress that bias in order to revise an initially formed meaning (Lee & Federmeier, 2012; Stites, Federmeier, & Stine-Morrow, 2013). These findings point to a reduced ability for older adults to perform language functions that require the engagement of control mechanisms, such as switching and inhibition. The decline in language processes requiring more top-down influence is in line with research demonstrating a decline in executive functions in older adults (e.g., Fisk & Sharp, 2004; Salthouse, Fristoe, McGuthry, & Hambrick, 1998). These executive control functions are understood to involve frontal lobe brain networks, which are known to change with aging (e.g., Brickman *et al.*, 2006; Foster, Black, Buck, & Bronskill, 1997; Harley *et al.*, 2011).

Despite evidence suggesting that older adults' efficiency of engaging control mechanisms is compromised, some neuroimaging research has reported that older adults recruit frontal lobe networks in the brain to perform some tasks for which younger adults do not show frontal lobe activation. This suggests that control processes may be required to compensate for a deficit in other aspects of processing. Differentially greater recruitment of frontal lobe brain areas, which have been associated with a wide variety of higher-order cognitive processes, as well as more bilateral hemispheric involvement in older adults as compared to the more typical lateralized neural activity found for younger adults (e.g., Cabeza, 2002), is observed in older adults more often than younger adults for tasks such as visual word recognition (Gold *et al.*, 2009), sentence comprehension (Grossman et al., 2002), and syntactic processing (Tyler *et al.*, 2010).

The need for greater recruitment of executive control processes to compensate for the detrimental effects of decreases in sensory or other cognitive processes may cause a disproportionate processing load on older adults compared to younger adults. For example, Stine-Morrow *et al.* (2010) proposed that older adults may process and integrate incoming information earlier than younger adults in order not to overload their working memory systems. A generalized effect of processing load has also been found for older adults when memory load and processing complexity are increased, whereas younger adults showed more selective levels of interference (Kemper & Herman, 2006). Thus, older adults' increased use of compensatory control mechanisms may lead to overall decreased processing resources being available to them. Such effects are only evident when processing load reaches a certain threshold.

### *Summary of comprehension in healthy older adults*

Modern theories of language processing in healthy aging must account for the independent effects of sensory, cognitive, and linguistic changes as well as their interactions. Differential patterns of decline and maintenance for a variety of language tasks point to a dissociation between age-related changes in controlled and automatic processes of language comprehension as well as the involvement of certain compensatory mechanisms that at times can lead older adults to perform just as well as younger adults and at other times do not result in the same level of processing efficiency as younger adults.

## Comprehension in aphasia

Aphasia is an acquired disorder in which damage to the brain results in a disturbance of language. Depending on the brain regions and networks that have been damaged, the aphasia takes on different forms. One gross distinction arises between non-fluent and fluent aphasias, the latter being traditionally associated with comprehension problems. In this chapter we focus on two common aphasia types: Broca's (non-fluent) aphasia and Wernicke's (fluent) aphasia, as the comprehension difficulties reported in these two types of aphasia have been studied the most.

### *Broca's aphasia*

In 1861, Paul Broca first described a severe form of the type of aphasia that later took on his name. His patient, Leborgne, produced extremely little speech, but was relatively unimpaired in language comprehension, and exhibited brain damage in the posterior inferior frontal gyrus of the left hemisphere of the brain (Broca, 1861), thereafter named Broca's area.

A fascinating subset of patients with the production difficulties of Broca's aphasia have relatively good lexical production but make errors in producing words and affixes that convey grammatical aspects of sentence structure, a phenomenon that Arnold Pick labeled "agrammatism" (Pick, 1902). Only over time did it become clear that patients with agrammatic production perform relatively well on comprehension tasks except when all semantic and pragmatic cues to meaning were minimized and comprehension of the sentence was solely dependent on syntactic structure. For example, these patients comprehend passive voice constructions relatively well when real-world knowledge leads them to a single plausible interpretation (e.g., *The gazelle was killed by the tiger. Who died?*). However, when knowledge of syntax is the only cue that can be used to lead to the correct interpretation, in so-called reversible passives (e.g., *The lion was killed by the tiger*), performance by agrammatic patients declines to near chance levels (e.g., Bastiaanse & Van Zonneveld, 2006; Berndt, Mitchum, & Haendiges, 1996; Caramazza & Zurif, 1976; Luzzatti *et al.*, 2001; Meyer, Mack, & Thompson, 2012; Schwartz, Saffran, & Marin, 1980). Grodzinsky (1986; 1995) has hypothesized that the mechanism used to connect a sentence element that was moved at the syntactic level (e.g., *lion* in the latter example) to the "trace" left by that movement in its original syntactic position is impaired in Broca's aphasia (Trace Deletion Hypothesis). However, other studies find above-chance performance on comprehension of passive sentences (e.g., Burchert & De Bleser, 2004), complicating this explanation. Berndt had originally argued that such difficulties in comprehending semantically reversible passive sentences occurred in all patients with Broca's aphasia but later recognized that only some patients experienced this problem (Berndt *et al.*, 1996).

Although individuals with Broca's aphasia generally perform poorly on a variety of syntactic comprehension tasks, they perform relatively well on grammaticality-judgment tasks. These findings have been replicated multiple times using different sentence constructions (e.g., dative sentences, passive sentences, and sentences containing relative clauses) and different languages (e.g., English, Italian, and Chinese) (Lu *et al.*, 2000; Wulfeck, Bates, & Capasso, 1991). Nevertheless, even when making judgments about grammaticality, patients with Broca's aphasia show better accuracy for some structures over others. Examples of this are found in a study by Wulfeck *et al.* (1991), who reported that English-speaking individuals with Broca's aphasia identified errors in subject-verb agreement less accurately than errors in word order. Thus, while grammaticality-judgment tasks tell us that patients with Broca's aphasia have the competence to appreciate sentence structure, it is not at the level of healthy controls. We assume that the performance requirements even for grammaticality-judgment tasks, not to mention comprehension ones, may demand cognitive resources beyond those that patients with Broca's aphasia can deploy in timely fashion.

The observation of differential performance for patients with Broca's aphasia across various comprehension tasks have generated several proposals regarding the underlying mechanisms of their deficits. One proposal is that while sentence production requires syntactic knowledge for every utterance, even the simplest one, comprehension processes only require syntactic knowledge for structurally complex sentences, as listeners can use other cues and non-syntactic strategies for understanding simpler sentences (Novick, Trueswell, & Thompson-Schill, 2005). In another proposal, two possibilities are considered to explain the specific comprehension patterns associated with agrammatic aphasia (Linebarger, Schwartz, & Saffran, 1983). The first possibility is that patients actually do construct syntactic representations during comprehension, but are incapable of subsequently deriving sentence meaning from those representations. The second possibility is that patients have a general cognitive resource limitation that restricts them from being capable of both effectively parsing a sentence and composing a semantic interpretation, resulting in different behavioral patterns for different tasks. Another suggestion is that phonological short-term memory is impaired when Broca's area is damaged, resulting in sentence comprehension deficits because of a deficit in articulatory rehearsal (Rogalsky & Hickok, 2011)—a skill necessary for successful sentence comprehension (Caplan, Alpert, Waters, & Olivieri, 2000). Despite the multiple proposals of the underlying mechanisms involved in processing deficits for individuals with aphasia, no consensus has been reached on the precise causes.

Given that the most common manifestation of comprehension impairment in Broca's aphasia consists of a deficit in processing grammatical structure, one would have no reason to assume that lexical processing in individuals with Broca's aphasia would be impaired. Indeed, for an extended period, several researchers claimed that single-word comprehension in individuals with Broca's aphasia or agrammatism is unimpaired (e.g., Goodglass, Kaplan, & Barresi, 2001; Schwartz *et al.*, 1980). However, more recent studies have shown that, although lexical processing is less impaired than sentence processing, the comprehension of single words is not completely spared either (e.g., Moineau, Dronkers, & Bates, 2005).

Several studies have shown that individuals with Broca's aphasia perform worse on semantic priming tasks (e.g., recognizing the word *dog* faster after seeing the word *cat* than after seeing the word *ring*) than healthy participants (e.g., Prather, Shapiro, Zurif, & Swinney, 1991; Utman, Blumstein, & Sullivan, 2001). This is interpreted as evidence that their lexical-semantic representations are at least partly impaired. Along these lines, it has been suggested that the problems seen for these patients in sentence contexts can be attributed to impaired selection of the contextually appropriate meaning and integration of lexical information (Swaab, Brown, & Hagoort, 1998).

As mentioned above, unlike the research on language comprehension in healthy elderly, relatively little research has focused on what roles cognitive skills play in the comprehension problems of Broca's aphasia. For example, individuals with agrammatic aphasia showed no evidence for a deficit in working memory capacity (Friedmann & Gvion, 2003). This is a surprising finding since Broca's area is often linked to domain-general working memory, which, moreover, has been shown to play a role in sentence comprehension (e.g., Santi & Grodzinsky, 2007). However, reasoning the other way around, cognitive functions could remain intact as grammatical deficits might not affect cognition. This is supported by a case study of an individual with severe agrammatism whose causal reasoning and Theory of Mind skills were intact, leading to the conclusion that grammatical disfluency and those cognitive processes are independent from each other (Varley & Siegal, 2000).

In sum, while earlier studies on Broca's aphasia considered comprehension abilities to be intact, over the years multiple studies have documented problems in both sentence comprehension and single-word comprehension. More specifically, researchers have proposed that an impairment of the lexical-semantic system may affect both levels of comprehension, although the underlying mechanisms are still under debate (Choy & Thompson, 2010). Despite the fact that views have changed over the years regarding which linguistic levels evidence impairment in Broca's aphasia, the research on single-word comprehension is relatively scarce. Rather, the focus in Broca's aphasia remains comprehension (and production) at the sentence level, with agrammatism the hallmark of interest for psycholinguists and neurolinguists.

## *Wernicke's aphasia*

In contrast to Broca's aphasia, Wernicke's aphasia has classically been described as a fluent aphasia, which refers to fluent language production. Wenicke's aphasia is associated with damage to posterior areas of the left hemisphere, with patients manifesting substantial problems in language comprehension. This comprehension deficit is more often attributed to impairment at the single-word level than at the sentence level. In fact, in the revision of a standard assessment tool for aphasia, the Boston Diagnostic Aphasia Examination (BDAE: Goodglass, Kaplan, & Barresi, 2001), having single-word comprehension problems is one of the three crucial diagnostic criteria for Wernicke's aphasia, in addition to fluent speech and impaired repetition abilities. For example, individuals with Wernicke's aphasia show significantly more impairment for picture-word matching than patients with Broca's aphasia, patients with right-hemisphere brain damage, patients with anomic aphasia (a relatively mild aphasia in which word-retrieval problems predominate), and healthy controls (Moineau *et al.*, 2005). These results are consistent with the results of other standardized aphasia tests and the general clinical understanding of the importance of comprehension problems as a key syndrome of Wernicke's aphasia.

The underlying mechanism causing the lexical comprehension problems in Wernicke's aphasia is often thought to be a deficit with phonological analysis, in which the decoding from acoustics to phonology is disrupted, but the analysis of prosody and other contextual cues is not (e.g., Luria, 1976). As a result, comprehension scores in formal testing situations (in which prosody and contextual cues are often minimized) may be different from those seen in everyday communication, where patients are able to exploit contextual cues to facilitate their comprehension (Robson, Sage, & Lambon Ralph, 2012).

An alternative proposal is that individuals with Wernicke's aphasia are impaired in their semantic processing (e.g., Hickok, 2000). This view is mostly based on neuroanatomical studies of healthy individuals that show that brain areas involved in semantic processing (i.e., posterior middle superior temporal areas—the latter defined as "Wernicke's area") are the same sites that are most often lesioned in individuals with symptoms of Wernicke's aphasia (e.g., Vigneau *et al.*, 2006). Tasks that have been used to assess semantic processing include picture-word matching (with written or orally presented words), object categorization, word association, and semantic priming tasks. A plausible account is that a combination of these two deficits, acoustic-phonological decoding and semantic processing, underlies the comprehension impairment in individuals with Wernicke's aphasia (Robson *et al.*, 2012).

Unlike with studies of Broca's aphasia, sentence comprehension deficits are so obvious in Wernicke's aphasia that they are diagnosed with clinical tasks (matching pictures to words or sentences, answering yes-no questions like *Is a good pair of rubber boots good for keeping water out?*) (Goodglass, Kaplan, & Barresi, 2001). Thus, sentence-level problems have rarely been the focus of experimental research in this population. Friederici and colleagues have focused on the influence of grammatical morphology in sentence comprehension and proposed that grammatical morphology, such as case and gender morphemes, provides cues to sentence meaning. Their studies showed that individuals with Wernicke's aphasia perform normally during on-line comprehension tasks (Friederici, 1985), but poorly during off-line comprehension, such as grammaticality-judgment tasks (Bates, Friederici, & Wulfeck, 1987), similar to what we described for healthy adults in the previous section. However, instead of attributing this to a working memory deficit, Bates *et al.* (1987) argued it is a linguistic deficit in that grammatical morphology, used in sentence comprehension, is at least as vulnerable in Wernicke's aphasia as it is in Broca's aphasia. Dronkers and Larsen (2001) argued that some individuals with Wernicke's aphasia may demonstrate problems with syntactic comprehension, but only in those cases where the lesion extends into more anterior areas towards the frontal lobe instead of being restricted to the mid-portion of the superior temporal gyrus. However, subsequent research showed that individuals with posterior temporal lesions, often seen in Wernicke's aphasia, were the most impaired in grammaticality judgment, though interestingly showing equal degrees of impairment across the different syntactic structures tested (Wilson, Saygin, Sereno, & Iacoboni, 2004). The authors concluded that how well a patient with aphasia performs on grammaticality judgment has nothing to do with the specific sentence structure.

In sum, comprehension problems are quite prominent in Wernicke's aphasia but it is unclear the extent to which these derive from lexical comprehension difficulties caused by either impaired phonological analysis or semantic processing deficits or both. It is also not clear whether syntactic and morphosyntactic comprehension processes themselves are impaired, since grammaticality judgment is challenging for patients with Wernicke's aphasia, but on-line comprehension has been shown to be spared in at least some patients.

## *Summary on comprehension in aphasia*

This section described two subtypes of aphasia, Broca's and Wernicke's aphasia, in which comprehension has been studied in markedly different ways and in which markedly different deficit patterns in comprehension have been observed. While Wernicke's aphasia has always been well known for its obvious comprehension problems, Broca's aphasia has only more recently been shown to involve specific problems in comprehension, mainly at the sentence level, though lexical processing may be impaired in Broca's aphasia as well. Wernicke's aphasia, by contrast, shows great deficits in single-word comprehension, a crucial diagnostic criterion of this subtype, and obvious sentence-level deficits (though these are rarely the focus of experimental studies). The language performance of these two different aphasia types demonstrates the dissociation between levels of language comprehension.

## Comprehension in dementia

Dementia is an umbrella term for a variety of progressive syndromes that include cognitive impairment among the outcomes of brain disease. The most frequently occurring type of dementia is Alzheimer's disease, but two other syndromes of interest to readers of this chapter are two of the three variants of primary progressive aphasia non-fluent and semantic primary progressive aphasia. Their names alone suggest that these syndromes fall somewhere between aphasia and dementia. The phenomenon of primary progressive aphasia is itself a relatively recent addition to the neurologist's diagnostic toolkit (Mesulam, 1982). The subtypes are considered to be a type of aphasia as the problems associated with them—at least at onset—are primarily linguistic; they exhibit the signature of dementia, however, in that there is progressive decline which towards the end also includes non-linguistic cognitive decline.[2] Thus, we highlight these dementia syndromes in this chapter as they are the ones in which impairments for sentence comprehension and single-word comprehension have been reported.

### *Alzheimer's disease*

Alzheimer's disease is typified by misfolded amyloid beta protein outside of neurons (amyloid plaques) and misfolded tau protein within neurons (neurofibrillary tangles), which are suspected to cause cell death, resulting in brain degeneration. In the early stage of the disease, the brain areas most affected are the temporo-parietal, lateral prefrontal, and medial temporal cortices. The most striking cognitive impairment in Alzheimer's disease is memory loss, but language problems are also a common feature. More specifically, one of the early language problems of Alzheimer's disease is word-finding substitutions, most of which are semantically related to the target (Mathews, Obler, & Albert, 1994). As the disease progresses, additional language problems surface. Alois Alzheimer (1911) described the language problems of his patient Johann F. in a case report (as quoted in Möller & Graeber, 1998: 112) giving us the following information about the patient's comprehension difficulties: "Very dull, slightly euphoric, slow in comprehension […], rare answers, frequent repetition of the question. […] Does not realise contradictions in speech."

The literature on comprehension problems in Alzheimer's disease contains varying findings. Some studies suggest that individuals with Alzheimer's disease have problems at both the single-word level and the sentence level. Others only focus on comprehension at the single-word level. Moreover, the degree to which the comprehension problems occur varies across studies; most likely, this results—at least in part—from testing patients in different stages of decline. In general, sentence comprehension is only mildly impaired in the early stages of Alzheimer's disease, whereas in middle-stages both comprehension and language production come to resemble that of Wernicke's aphasia (Bickel, Pantel, Eysenbach, & Schröder, 2000; Obler & Albert, 1984). By late stages of the disease, a few formulaic utterances may be comprehended, but comprehension is poor enough that patients can no longer follow directions (Lamar, Obler, Knoefel, & Albert, 1995).

Regarding sentence comprehension, there is quite a bit of controversy as to whether or not this level of language processing is affected. When testing 11 patients with Alzheimer's disease on auditory comprehension of language, Kontiola, Laaksonen, Sulkava, and Erkinjuntti (1990) found impairments in understanding complex grammatical structures in all of the patients. Difficulty with sentences has also been shown on reading tasks, such as the Test for Reception of Grammar, an 80-item sentence-picture matching task testing different lexical, morphosyntactic, and syntactic constructions (Croot, Patterson, & Hodges, 1999).

However, other studies have found that sentence comprehension is preserved in Alzheimer's disease. Relative preservation of syntactic operations was found in a case study of a patient with Alzheimer's disease, despite the fact that she was severely impaired on semantic knowledge (she could no longer match even highly familiar words with their real-life or depicted referents) (Schwartz, Marin, & Saffran, 1979). Similarly in other studies, individuals with mild or moderate Alzheimer's disease did not perform significantly differently from healthy controls on comprehending verbally presented simple and complex sentences (Grossman *et al.*, 1996a) and were able to understand syntactically complex sentences in a sentence-picture matching test (Rochon, Waters, & Caplan, 1994; Waters, Caplan, & Rochon, 1995). In a sentence acceptability judgment task, patients with Alzheimer's disease were not disproportionately affected when identifying the referent for a reflexive pronoun (which relies on syntactic knowledge) using three different sentence types: simple sentences with only one full lexical noun phrase, complex sentences with two full lexical noun phrases (only one of which was in the appropriate syntactic position to be the antecedent of the reflexive), and a third sentence type controlling for the effect of sentence length (Waters & Caplan, 1997). This led the authors to conclude that individuals with Alzheimer's disease, even though they perform more poorly overall than control subjects, are not directly affected by working memory limitations, as the increase of working memory demands for complex sentences did not disproportionally affect their performance.

There is more agreement among researchers regarding the presence of an impairment of single-word comprehension compared to sentence comprehension in Alzheimer's disease (Waters, Rochon, & Caplan, 1998). Comprehension of words with both superordinate and specific semantics is impaired in Alzheimer's disease, except when words are related to emotion (Martin & Fedio, 1983). In a spoken-word/picture matching task, individuals with Alzheimer's disease showed impaired single-word comprehension, making mostly semantic errors such as choosing a semantically related distracter (for instance, *goat* when the target was *sheep*) (Diesfeldt, 1989; see also Masterson *et al.*, 2007). Patients with Alzheimer's disease have also shown impairment of specific semantic categories. For example, their processing of the animate category *animals* was significantly worse than the inanimate category *furniture* on a written semantic similarity-judgment task (Vonk, Jonkers, De Santi, & Obler, 2012). This is consistent with other studies on semantic categories that show that individuals with Alzheimer's disease are generally more impaired for comprehension of words denoting living objects than non-living objects (e.g., Almor *et al.*, 2009).

The semantically related error patterns in sentence comprehension and the specific semantic category deficit in single-word comprehension in patients with Alzheimer's disease suggest that the underlying mechanism responsible for comprehension problems in Alzheimer's disease lies at the semantic level; some type of breakdown of semantic boundaries appears to take place. Martin and Fedio (1983) proposed that the semantic deficit affects the differentiation between items within the same category but that semantic knowledge at a broader level is relatively preserved. Another explanation is that the language problems in Alzheimer's disease are related to an impairment of episodic memory, that is, memory for specific events in the past (Holland, Boiler, & Bourgeois, 1986). The authors proposed that repetition skills stay relatively intact due to preserved short-term memory, auditory comprehension, and syntax, while more general functions, such as episodic memory, deteriorate, which results in the specific language problems seen in Alzheimer's disease. However, the sentence comprehension deficits found by some researchers are hard to explain as resulting merely from an overall impairment in episodic memory. Therefore, the account of a semantic deficit in patients with Alzheimer's disease is more widely accepted.

## *Primary progressive aphasia*

As mentioned above, primary progressive aphasia is classified as a dementia syndrome because decline is progressive, but it carries the term *aphasia* in its name because the main impairment is of language and speech. Primary progressive aphasia is linked to atrophy in the left frontal and temporal brain areas, regions that include both Broca's and Wernicke's areas, plus regions extending beyond them. The primary progressive aphasia classification can be subdivided into three distinct variants: non-fluent, semantic, and logopenic (Gorno-Tempini *et al.*, 2011); here we discuss non-fluent and semantic primary progressive aphasia, for which comprehension problems are regularly reported.

The neuroanatomical damage in non-fluent primary progressive aphasia is specifically located in left inferior frontal and adjacent anterior-superior temporal regions, similar to the brain damage seen in Broca's aphasia (Charles *et al.*, 2014). Patients with non-fluent primary progressive aphasia have severe deficits in sentence comprehension, but do not appear to be impaired for single-word comprehension (Mesulam, 2001). The difficulty understanding grammatical aspects of sentences is a fundamental feature of the non-fluent primary progressive aphasia (Charles *et al.*, 2014). The deficit has been shown in several tasks, such as answering questions about complex sentences containing either subject- or object-relative clauses (e.g., Peelle *et al.*, 2008) and completing a sentence picture-matching task involving complex syntactic structures such as subordination and center-embedding (e.g., Hodges & Patterson, 1996). These impairments are found for both online and offline sentence comprehension tasks (Grossman, Rhee, & Moore, 2005). The sentence comprehension deficit appears to be greater for specific types of sentences; patients with non-fluent primary progressive aphasia—compared to patients with Alzheimer's disease—had greater difficulty with sentences containing subordinate clauses than syntactically simple sentences in an auditory comprehension task (Grossman *et al.*, 1996b). Moreover, the patients with non-fluent primary progressive aphasia had more difficulty with object-relative subordinate clauses than subject-relative clauses (Grossman *et al.*, 1996b). Charles *et al.* (2014) showed that although all variants of primary progressive aphasia demonstrate difficulty with center-embedded subordinate clauses, only individuals with non-fluent primary progressive aphasia exhibit difficulty with cleft sentence structures.

Grossman *et al.* (2005) have developed two theories of the underlying mechanisms responsible for this specific deficit in sentence comprehension as compared to single-word comprehension. In one, the deficit is attributed to a limited and/or slow working memory system. According to this view, during sentence comprehension, crucial information degrades in working memory before it can be retrieved because of abnormally slow processing of syntactic information. This explanation is supported by findings that individuals with non-fluent primary progressive aphasia have a greatly impaired auditory-verbal working memory (Grossman *et al.*, 1996b). In addition, Grossman *et al.* (2013) suggested that part of the damaged neuroanatomical area characteristic in non-fluent primary progressive aphasia is involved in working memory—analogous to what Santi and Grodzinsky (2007) proposed—which would explain why working memory abilities are poor in non-fluent primary progressive aphasia. However, as we discussed previously in the section on Broca's aphasia—whose patients show similarities in the affected brain areas to those in non-fluent primary progressive aphasia—this argument does not work when investigating the role of working memory in agrammatic aphasia (e.g., Friedmann & Gvion, 2003).

The alternate account by Grossman *et al.* (2005) to explain the difficulties patients with non-fluent primary progressive aphasia experience in sentence comprehension is that they have impaired grammatical knowledge and processing skills. This point of view is supported by the difference in performance across various sentence types (e.g., subject-relative versus object-relative clauses). In sentences of equal length and equivalent lexical content, better performance is seen for simple transitive sentences than for grammatically complex sentences with a center-embedded subordinate clause or a sentence-final subordinate clause. As Grossman *et al.* (2005) point out, both accounts explain the data, so the truth about the underlying mechanism of these deficits in non-fluent primary progressive aphasia is not yet known.

Semantic primary progressive aphasia is associated with brain atrophy in the anterior temporal region and ventral and lateral portions of the left temporal lobe (Gorno-Tempini *et al.*, 2011; Grossman et al., 2005), that is, regions that overlap with the mid-temporal regions reported to be most often lesioned in patients with Wernicke's aphasia (Vigneau *et al.*, 2006). Behavioral studies show that the comprehension problems in individuals with semantic primary progressive aphasia are mostly centered on single words (Mesulam, 2001). In fact, this impairment is considered to be one of the core features of semantic primary progressive aphasia. This impairment for single-word comprehension surfaces even at the earliest stages of the disease and is thought to be due to loss of semantic memory.[3] The semantic deficit in semantic primary progressive aphasia is present in most semantic categories of words (e.g., animals, tools) and is especially noticeable in the comprehension of low-frequency words (Gorno-Tempini *et al.*, 2011).

Patients with this type of dementia typically exhibit no substantial impairment of sentence comprehension (Rochon, Kavè, Cupit, Jokel, & Winocur, 2004). However, some studies show subtle signs of difficulty by individuals with semantic primary progressive aphasia in comprehending sentences. For example, poor single-word comprehension of individuals with semantic primary progressive aphasia also compromises their sentence comprehension (Grossman *et al.*, 2004). In another study, patients with semantic primary progressive aphasia had poor comprehension of center-embedded sentences (Charles *et al.*, 2014). On the other hand, several case studies showed that although single-word processing is degraded in semantic primary progressive aphasia, the comprehension of even complex sentences (e.g., object-relative structures) remained relatively intact (e.g., Breedin & Saffran, 1999; Rochon *et al.*, 2004).

In general, the language comprehension problems in semantic primary progressive aphasia have been attributed to impaired semantic feature knowledge (i.e., knowledge about certain aspects of concepts) (Grossman *et al.*, 2005). Possible problems in sentence comprehension are considered to be a result of the same deficit. However, except for the few studies mentioned above, sentence processing has not been studied extensively in semantic primary progressive aphasia and more research is needed to provide better insight into possibly distinguishing different comprehension impairments at the sentence or single-word level in individuals with semantic primary progressive aphasia.

## Summary on comprehension in dementia

Language comprehension problems vary across different variants of dementia. In this section we focused on the three types of dementia with prominent comprehension problems, namely Alzheimer's disease and two variants of primary progressive aphasia non-fluent and semantic. Comprehension problems in Alzheimer's disease are widely acknowledged to exist at the single-word level. However, whether or not individuals with Alzheimer's disease suffer from sentence comprehension difficulties independent of their single-word problems is still under debate. The two variants of primary progressive aphasia discussed in this section show opposite comprehension patterns, analogous to those found between the aphasia subtypes described above. Individuals with non-fluent primary progressive aphasia mainly have problems at the sentence level, but not at the single-word level, while individuals with semantic primary progressive aphasia show a single-word comprehension deficit with no substantial deficit of syntactic processing (although some studies also show problems at the sentence level).

An important consideration to take into account when investigating language processing in dementia is that the severity of any of these dementias increases over time, due to the progressive nature of the disease. In most cases, associated language problems follow the same path, and comprehension in general decreases in proportion to the course of the disease. Therefore, language problems that did not seem to exist at first diagnosis could surface at a later stage of the disease.

## Summary

Across the populations we have reported on in this chapter, comprehension of "more complex" syntactic structures is more likely to show decline than that of "less complex" structures, but whether this is due to problems with the linguistic representations or with declines in the cognitive skills that support complex structure is still unspecified. What we do know at this point, however, is that comprehension declines in somewhat specific ways in healthy aging and in patients with different types of brain damage. When frontal regions of the brain are impaired, comprehending reversible passives is difficult if no context is provided to aid in the correct syntactic interpretation of sentences; when posterior regions are impaired, single-word comprehension—especially of low-frequency words—declines, rendering both single word and sentence comprehension problematic.

Nevertheless, on-line comprehension is often better spared in patients with aphasia and early stages of Alzheimer's disease than is off-line comprehension, similar to what is seen in older adults who are considered healthy. If off-line comprehension is differentially impaired, then the relatively common finding of a link between working memory deficits and poor comprehension likely contributes to the interpretation problems. Other executive control mechanisms, too, are probably called upon for comprehension, but they have been less studied in the literature to date.

Clearly further research is necessary to clarify how word and sentence comprehension can be affected by changes in brain structures associated with advancing age, aphasia and dementia. Ideally participants—both monolingual and multilingual—from a broad range of adult ages, across broad educational and socio-economic status ranges, would be tested on a broad range of syntactic structures. As technology permits isolation of specific neuronal networks, moreover, our understanding of how word comprehension operates and how it integrates with syntactic parsing to permit sentence comprehension will improve.

## NOTES

1 Tip-of-the-tongue states are a type of word retrieval in which the speaker feels very close to retrieving a word without being able to do so. Sometimes, partial phonological information (such as the first sound or number of syllables) and grammatical information (such as the grammatical gender) may be known, but full phonology cannot be retrieved. These states sometimes resolve within a few seconds and sometimes take much longer.
2 In classic forms of aphasia, there is usually a sudden onset of symptoms from a stroke or accident (though tumors can cause a slowly progressive decline). Cognitive deficits are rare and non-progressive.
3 Semantic memory is defined here according to the definition by Tulving (1972: 401–402) as "a system for receiving, retaining, and transmitting information about meaning of words, concepts, and classification of concepts."

## REFERENCES


Adrover-Roig, D., Sesé, A., Barceló, F., & Palmer, A. (2012). A latent variable approach to executive control in healthy ageing. *Brain and Cognition*, 78(3), 284–299. doi: 10.1016/j.bandc.2012.01.005

Allen, P. A., Madden, D. J., Weber, T. A., & Groth, K. E. (1993). Influence of age and processing stage on visual word recognition. *Psychology and Aging*, 8(2), 274–282. doi: 10.1037/0882-7974.8.2.274

Almor, A., Aronoff, J. M., MacDonald, M. C., Gonnerman, L. M., Kempler, D., Hintiryan, H.,…Andersen, E. S. (2009). A common mechanism in verb and noun naming deficits in Alzheimer's patients. *Brain and Language*, 111(1), 8–19. doi: 10.1016/j.bandl.2009.07.009

Alzheimer, A. (1911). Über eigenartige Krankheitsfälle des späteren Alters. *Zeitschrift für die Gesamte Neurologie und Psychiatrie*, 4(1), 356–385. doi: 10.1007/BF02866241

Baddeley, A. D., Baddeley, H., Chincotta, D., Luzzi, S., & Meikle, C. (2005). Working memory and aging. In J. Duncan, L. H. Phillips, & P. McLeod (Eds.), *Measuring the mind: Speed, control, and age* (pp. 193–316). Oxford, UK: University Press. doi: 10.1093/acprof:oso/9780198566427.003.0012

Balota, D. A., & Duchek, J. M. (1988). Age-related differences in lexical access, spreading activation, and simple pronunciation. *Psychology and Aging*, 3(1), 84–93. doi: 10.1037/0882-7974.3.1.84

Barresi, B. A., Nicholas, M., Connor, L., Obler, L. K., & Albert, M. L. (2000). Semantic degradation and lexical access in age-related naming failures. *Aging, Neuropsychology, and Cognition*, 7(3), 169–178. doi: 10.1076/1382-5585(200009)7:3;1-Q;FT169

Bastiaanse, R., & Van Zonneveld, R. (2006). Comprehension of passives in Broca's aphasia. *Brain and Language*, 96(2), 135–142. doi: 10.1016/j.bandl.2005.06.012

Bates, E., Friederici, A., & Wulfeck, B. (1987). Comprehension in aphasia: A cross-linguistic study. *Brain and Language*, 32(1), 19–67. doi: 10.1016/0093-934X(87)90116-7

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., & Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *Journal of Speech, Language, and Hearing Research*, 54(1), 243–262. doi: 10.1044/1092-4388 (2010/09-0233)

Benichov, J., Cox, L. C., Tun, P. A., & Wingfield, A. (2012). Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability, and cognitive function. *Ear and Hearing*, 33(2), 250–256. doi: 10.1097/AUD.0b013e31822f680f

Berndt, R. S., Mitchum, C. C., & Haendiges, A. N. (1996). Comprehension of reversible sentences in "agrammatism": A meta-analysis. *Cognition*, 58(3), 289–308. doi: 10.1016/0010-0277(95)00682-6

Bickel, C., Pantel, J., Eysenbach, K., & Schröder, J. (2000). Syntactic comprehension deficits in Alzheimer's disease. *Brain and Language*, 71(3), 432–448. doi: 10.1006/brln.1999.2277

Breedin, S. D., & Saffran, E. M. (1999). Sentence processing in the face of semantic loss: A case study. *Journal of Experimental Psychology: General*, 128(4), 547–562. doi: 10.1037/0096-3445.128.4.547

Brickman, A. M., Zimmerman, M. E., Paul, R. H., Grieve, S. M., Tate, D. F., Cohen, R. A.,…Gordon, E. (2006). Regional white matter and neuropsychological functioning across the adult lifespan. *Biological Psychiatry*, 60(5), 444–453. doi: 10.1016/j.biopsych.2006.01.011

Broca, P. (1861). Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau. *Bulletin de la Société Anthropologique*, 2, 235–238.

Burchert, F., & De Bleser, R. (2004). Passives in agrammatic sentence comprehension: A German study. *Aphasiology*, 18(1), 29–45. doi: 10.1080/02687030344000409

Burke, D. M., MacKay, D. G., & James, L. E. (2000). Theoretical approaches to language and aging. In T. J. Perfect & E. A. Maylor (Eds), *Models of cognitive aging* (pp. 204–237). New York, NY: Oxford University Press.

Burke, D. M., White, H., & Diaz, D. L. (1987). Semantic priming in young and older adults: Evidence for age constancy in automatic and attentional processes. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 79–88. doi: 10.1037/0096-1523.13.1.79

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging*, 17(1), 85–100. doi: 10.1037/0882-7974.17.1.85

Cahana-Amitay, D., & Albert, M. L. (2015). *Redefining recovery from aphasia.* New York, NY: Oxford University Press. doi: 10.1093/med/9780199811939.003.0009

Cahana-Amitay, D., Spiro III, A., Sayers, J. T., Oveis, A. C., Higby, E., Ojo, E. A.,… Obler, L. K. (2016). How older adults use cognition in sentence-final word recognition. *Aging, Neuropsychology, and Cognition*, 23(4), 418–444. doi: 10.1080/13825585.2015.1111291

Caplan, D., Alpert, N., Waters, G., & Olivieri, A. (2000). Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping*, 9(2), 65–71. doi: 10.1002/(SICI)1097-0193(200002)9:2<65::AID-HBM1>3.0.CO;2-4

Caplan, D., DeDe, G., Waters, G., Michaud, J., & Tripodis, Y. (2011). Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging*, 26(2), 439–450.

Caplan, D., & Waters, G. (2003). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory*, 13(3-4), 403–413. doi: 10.1037/a0021837

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22(1), 77–94. doi: 10.1017/S0140525X99001788

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and*

*Language*, 3(4), 572–582. doi: 10.1016/0093-934X(76)90048-1

Charles, D., Olm, C., Powers, J., Ash, S., Irwin, D. J., McMillan, C. T.,…Grossman, M. (2014). Grammatical comprehension deficits in non-fluent/agrammatic primary progressive aphasia. *Journal of Neurology, Neurosurgery, and Psychiatry*, 85(3), 249–256. doi: 10.1136/jnnp-2013-305749

Choy, J., & Thompson, C. (2010). *Effects of lexical processing deficits on agrammatic sentence comprehension: An eyetracking study*. Paper presented at the Clinical Aphasiology Conference, Isle of Palms, SC.

Croot, K., Patterson, K., & Hodges, J. R. (1999). Familial progressive aphasia: Insights into the nature and deterioration of single word processing. *Cognitive Neuropsychology*, 16(8), 705–747. doi: 10.1080/026432999380627

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, 3(4), 422–433. doi: 10.3758/BF03214546

DeDe, G., Caplan, D., Kemtes, K., & Waters, G. (2004). The relationship between age, verbal working memory, and language comprehension. *Psychology and Aging*, 19(4), 601–616. doi: 10.1037/0882-7974.19.4.601

Diesfeldt, H. (1989). Semantic impairment in senile dementia of the Alzheimer type. *Aphasiology*, 3(1), 41–54. doi: 10.1080/02687038908248975

Dronkers, N. F., & Larsen, J. (2001). Neuroanatomy of the classical syndromes of aphasia. In R. S. Berndt (Ed.), *Handbook of neuropsychology: Language and aphasia* (2nd ed.) (pp. 19–30). Amsterdam, Netherlands: Elsevier.

Elliott, L. L., Hammer, M. A., & Evan, K. E. (1987). Perception of gated, highly familiar spoken monosyllabic nouns by children, teenagers, and older adults. *Perception & Psychophysics*, 42(2), 150–157. doi: 10.3758/BF03210503

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, 115(3), 149–161. doi: 10.1016/j.bandl.2010.07.006

Federmeier, K. D., Van Petten, C., Schwartz, T. J., & Kutas, M. (2003). Sounds, words, sentences: Age-related changes across levels of language processing. *Psychology and Aging*, 18(4), 858–872. doi: 10.1037/0882-7974.18.4.858

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology*, 26(7), 874–890. doi: 10.1080/13803390490510680

Foster, J., Black, S., Buck, B. H., & Bronskill, M. J. (1997). Ageing and executive functions: A neuroimaging perspective. In P. Rabbit (Ed.), *Methodology of frontal and executive function* (pp. 177–190). Hove, UK: Psychology Press.

Friederici, A. D. (1985). Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition*, 19(2), 133–166. doi: 10.1016/0010-0277(85)90016-2

Friedmann, N., & Gvion, A. (2003). Sentence comprehension and working memory limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain and Language*, 86(1), 23–39. doi: 10.1016/S0093-934X(02)00530-8

Giaquinto, S., Ranghi, F., & Butler, S. (2007). Stability of word comprehension with age: An electrophysiological study. *Mechanisms of Ageing and Development*, 128(11), 628–636. doi: 10.1016/j.mad.2007.09.003

Gold, B. T., Andersen, A. H., Jicha, G. A., & Smith, C. D. (2009). Aging influences the neural correlates of lexical decision but not automatic semantic priming. *Cerebral Cortex*, 19(11), 2671–2679. doi: 10.1093/cercor/bhp018

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58(3), 787–814. doi: 10.1016/j.jml.2007.07.001

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The assessment of aphasia and related disorders*. Philadelphia, PA: Lippincott Williams & Wilkins.

Goral, M., Clark-Cotton, M., Spiro III, A., Obler, L. K., Verkuilen, J., & Albert, M. L. (2011). The contribution of set switching and working memory to sentence processing in older adults. *Experimental Aging Research*, 37(5), 516–538. doi: 10.1080/0361073X.2011.619858

Goral, M., & Obler, L. K. (2003). Root-morpheme processing during word recognition in Hebrew speakers across the adult life span. In J. Shimron (Ed.), *Language processing and acquisition in languages of Semitic, root-based, morphology* (pp. 223–242). Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/lald.28.12gor

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F.,…Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006–1014. doi: 10.1212/WNL.0b013e31821103e6

Grieder, M., Crinelli, R. M., Koenig, T., Wahlund, L., Dierks, T., & Wirth, M. (2012). Electrophysiological and behavioral correlates of stable automatic semantic retrieval in aging. *Neuropsychologia*, 50(1), 160–171. doi: 10.1016/j.neuropsychologia.2011.11.014

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language*, 27(1), 135–159. doi: 10.1016/0093-934X(86)90009-X

Grodzinsky, Y. (1995). Trace deletion, $\theta$-roles, and cognitive strategies. *Brain and Language*, 51(3), 469–497. doi: 10.1006/brln.1995.1072

Grossman, M., Cooke, A., DeVita, C., Alsop, D., Detre, J., Chen, W., & Gee, J. (2002). Age-related changes in working memory during sentence comprehension: An fMRI study. *NeuroImage*, 15(2), 302–317. doi: 10.1006/nimg.2001.0971

Grossman, M., Cooke, A., McMillan, C., Moore, P., Gee, J., & Work, M. (2004). Sentence comprehension in progressive aphasia and frontotemporal dementia: An fMRI study. *Brain and Language*, 91(1), 134–135. doi: 10.1016/j.bandl.2004.06.070

Grossman, M., D'Esposito, M., Hughes, E., Onishi, K., Biassou, N., White-Devine, T., & Robinson, K. M. (1996a). Language comprehension profiles in Alzheimer's disease, multi-infarct dementia, and frontotemporal degeneration. *Neurology*, 47(1), 183–189. doi: 10.1212/WNL.47.1.183

Grossman, M., Mickanin, J., Onishi, K., Hughes, E., D'Esposito, M., Ding, X.,…Reivich, M. (1996b). Progressive nonfluent aphasia: Language, cognitive, and PET measures contrasted with probable Alzheimer's disease. *Journal of Cognitive Neuroscience*, 8(2), 135–154. doi: 10.1162/jocn.1996.8.2.135

Grossman, M., Powers, J., Ash, S., McMillan, C., Burkholder, L., Irwin, D., & Trojanowski, J. Q. (2013). Disruption of large-scale neural networks in non-fluent/agrammatic variant primary progressive aphasia associated with frontotemporal degeneration pathology. *Brain and Language*, 127(2), 106–120. doi: 10.1016/j.bandl.2012.10.005

Grossman, M., Rhee, J., & Moore, P. (2005). Sentence processing in frontotemporal dementia. *Cortex*, 41(6), 764–777. doi: 10.1016/S0010-9452(08)70295-8

Harley, T. A., Jessiman, L. J., & MacAndrew, S. B. (2011). Decline and fall: A biological, developmental, and psycholinguistic account of deliberative language processes and ageing. *Aphasiology*, 25(2), 123–153. doi: 10.1080/02687031003798262

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22), 193–225. doi: 10.1016/S0079-7421(08)60041-9

Hickok, G. (2000). Speech perception, conduction aphasia, and the functional neuroanatomy of language. In Y. Grodzinsky, L. P. Shapiro, & D. Swinney (Eds.), *Language and the brain: Representation and processing* (pp. 87–104). San Diego, CA: Academic Press. doi: 10.1016/B978-012304260-6/50006-2

Higby, E., Cahana-Amitay, D., Vogel, A., Spiro III, A., Albert, M., & Obler, L. K. (submitted). The role of executive functions in naming among older adults.

Hodges, J. R., & Patterson, K. (1996). Nonfluent progressive aphasia and semantic dementia: A comparative neuropsychological study. *Journal of the International Neuropsychological Society*, 2(6), 511–524. doi: 10.1017/S1355617700001685

Holland, A. L., Boiler, F., & Bourgeois, M. (1986). Repetition in Alzheimer's disease: A longitudinal study. *Journal of Neurolinguistics*, 2(1), 163–177. doi: 10.1016/S0911-6044(86)80010-0

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(6), 1378–1391. doi: 10.1037/0278-7393.26.6.1378

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging*, 8(2), 283–293. doi: 10.1037/0882-7974.8.2.283

Kane, M. J., Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Connelly, S. L. (1994). Inhibitory attentional mechanisms and aging. *Psychology and Aging*, 9(1), 103–112. doi: 10.1037/0882-7974.9.1.103

Kavé, G., & Levy, Y. (2005). The processing of morphology in old age: Evidence from Hebrew. *Journal of Speech, Language, and Hearing Research*, 48(6), 1442–1451. doi: 10.1044/1092-4388(2005/100)

Kemmer, L., Coulson, S., De Ochoa, E., & Kutas, M. (2004). Syntactic processing with aging: An event-related potential study. *Psychophysiology*, 41(3), 372–384.

Kemper, S., Crow, A., & Kemtes, K. (2004). Eye-fixation patterns of high- and low-span young and older adults: Down the garden path and back again. *Psychology and Aging*, 19(1), 157–170. doi: 10.1111/1469-8986.2004.00180.x

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging*, 16(2), 312–322. doi: 10.1037/0882-7974.19.1.157

Kidd, G. R., & Humes, L. E. (2012). Effects of age and hearing loss on the recognition of interrupted words in isolation and in sentences. *The Journal of the Acoustical Society of America*, 131(2), 1434–1448. doi: 10.1121/1.3675975

Kontiola, P., Laaksonen, R., Sulkava, R., & Erkinjuntti, T. (1990). Pattern of language impairment is different in Alzheimer's disease and multi-infarct dementia. *Brain and Language*, 38(3), 364–383. doi: 10.1016/0093-934X(90)90121-V

Kramer, A. F., Humphrey, D. G., Larish, J. F., & Logan, G. D. (1994). Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. *Psychology and Aging*, 9(4), 491–512. doi: 10.1037/0882-7974.9.4.491

Kwong See, S. T., & Ryan, E. B. (1995). Cognitive mediation of adult age differences in language performance. *Psychology and Aging*, 10(3), 458–468. doi: 10.1037/0882-7974.10.3.458

Lamar, M. A. C., Obler, L. K., Knoefel, J. E., & Albert, M. L. (1995). Communication patterns in end-stage Alzheimer's disease: Pragmatic analyses. In R. L. Bloom, L. K. Obler, S. De Santi. & J. S. Ehrlich (Eds.), *Discourse analysis and applications: Studies in adult clinical populations* (pp. 217–236). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lash, A., Rogers, C. S., Zoller, A., & Wingfield, A. (2013). Expectation and entropy in spoken word recognition: Effects of age and hearing acuity. *Experimental Aging Research*, 39(3), 235–253. doi: 10.1080/0361073X.2013.779175

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging*, 8(1), 34–43. doi: 10.1037/0882-7974.8.1.34

Lee, C., & Federmeier, K. D. (2012). Ambiguity's aftermath: How age differences in resolving lexical ambiguity affect subsequent comprehension. *Neuropsychologia*, 50(5), 869–879. doi: 10.1016/j.neuropsychologia.2012.01.027

Lin, F. R., Yaffe, K., Xia, J., Xue, Q., Harris, T. B., Purchase-Helzner, E.,…Simonsick, E. M. (2013). Hearing loss and cognitive decline in older adults. *JAMA Internal Medicine*, 173(4), 293–299. doi: 10.1001/jamainternmed.2013.1868

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, 13(3), 361–392. doi: 10.1016/0010-0277(83)90015-X

Lu, C., Bates, E., Li, P., Tzeng, O., Hung, D., Tsai, C.,…Chung, Y. (2000). Judgements of grammaticality in aphasia: The special case of Chinese. *Aphasiology*, 14(10), 1021–1054. doi: 10.1080/02687030050156593

Luria, A. (1976). *Disturbances of understanding of verbal communication in patients with sensory aphasia*. The Hague: Mouton & Co.

Luzzatti, C., Toraldo, A., Guasti, M. T., Ghirardi, G., Lorenzi, L., & Guarnaschelli, C. (2001). Comprehension of reversible active and passive sentences in agrammatism. *Aphasiology*, 15(5), 419–441. doi: 10.1080/02687040143000005

MacKay, D. G., Abrams, L., & Pedroza, M. J. (1999). Aging on the input versus output side: Theoretical implications of age-linked asymmetries between detecting versus retrieving orthographic information. *Psychology and Aging*, 14(1), 3–17. doi: 10.1037/0882-7974.14.1.3

Madden, D. J., Pierce, T. W., & Allen, P. A. (1993). Age-related slowing and the time course of semantic priming in visual word identification. *Psychology and Aging*, 8(4), 490–507. doi: 10.1037/0882-7974.8.4.490

Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language*, 19(1), 124–141. doi: 10.1016/0093-934X(83)90059-7

Masterson, J., Druks, J., Kopelman, M., Clare, L., Garley, C., & Hayes, M. (2007). Selective naming (and comprehension) deficits in Alzheimer's disease? *Cortex*, 43(7), 921–934. doi: 10.1016/S0010-9452(08)70691-9

Mathews, P. J., Obler, L. K., & Albert, M. L. (1994). Wernicke and Alzheimer on the language disturbances of dementia and aphasia. *Brain and Language*, 46(3), 439–462. doi: 10.1006/brln.1994.1024

Mesulam, M. M. (1982). Slowly progressive aphasia without generalized dementia. *Annals of Neurology*, 11(6), 592–598. doi: 10.1002/ana.410110607

Mesulam, M. M. (2001). Primary progressive aphasia. *Annals of Neurology*, 49(4), 425–432. doi: 10.1002/ana.91

Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, 25(1), 31–43. doi: 10.1016/j.jneuroling.2011.08.001

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. doi: 10.1006/cogp.1999.0734

Moineau, S., Dronkers, N. F., & Bates, E. (2005). Exploring the processing continuum of single-word comprehension in aphasia. *Journal of Speech, Language, and Hearing Research*, 48(4), 884–896. doi: 10.1044/1092-4388(2005/061)

Möller, H., & Graeber, M. (1998). The case described by Alois Alzheimer in 1911. *European Archives of Psychiatry and Clinical Neuroscience*, 248(3), 111–122. doi: 10.1007/s004060050027

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 5(3), 263–281. doi: 10.3758/CABN.5.3.263

Obler, L. K., & Albert, M. (1984). Language in aging. In M. L. Albert (Ed.), *Clinical neurology of aging*, (pp. 245–253). New York, NY: Oxford University Press.

Obler, L. K., Nicholas, M., Albert, M. L., & Woodward, S. (1985). On comprehension across the adult lifespan. *Cortex*, 21(2), 273–280. doi: 10.1016/S0010-9452(85)80032-0

Park, D. C., & Hedden, T. (2001). *Working memory and aging*. New York, NY: Psychology.

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition*, 130(2), 157–173. doi: 10.1016/j.cognition.2013.10.005

Peelle, J. E., Troiani, V., Gee, J., Moore, P., McMillan, C., Vesely, L., & Grossman, M. (2008). Sentence comprehension and voxel-based morphometry in progressive nonfluent aphasia, semantic dementia, and nonaphasic frontotemporal dementia. *Journal of Neurolinguistics*, 21(5), 418–432. doi: 10.1016/j.jneuroling.2008.01.004

Pichora-Fuller, M. K. (2003). Processing speed and timing in aging adults: Psychoacoustics, speech perception, and comprehension. *International Journal of Audiology*, 42, S59–S67. doi: 10.3109/14992020309074625

Pick, A. (1902). Über Agrammatismus als Folge Herderkrankung. *Zeitschrift für Heilkunde*, 23, 82–90.

Prather, P., Shapiro, L., Zurif, E., & Swinney, D. (1991). Real-time examinations of lexical processing in aphasics. *Journal of Psycholinguistic Research*, 20(3), 271–281. doi: 10.1007/BF01067219

Robson, H., Sage, K., & Lambon Ralph, M. A. (2012). Wernicke's aphasia reflects a combination of acoustic-phonological and semantic control deficits: A case-series comparison of Wernicke's aphasia, semantic dementia and semantic aphasia. *Neuropsychologia*, 50(2), 266–275. doi: 10.1016/j.neuropsychologia.2011.11.021

Rochon, E., Kavè, G., Cupit, J., Jokel, R., & Winocur, G. (2004). Sentence comprehension in semantic dementia: A longitudinal case study. *Cognitive Neuropsychology*, 21(2–4), 317–330. doi: 10.1080/02643290342000357

Rochon, E., Waters, G. S., & Caplan, D. (1994). Sentence comprehension in patients with Alzheimer′ s disease. *Brain and Language*, 46(2), 329–349. doi: 10.1006/brln.1994.1018

Rogalsky, C., & Hickok, G. (2011). The role of Broca's area in sentence comprehension. *Journal of Cognitive Neuroscience*, 23(7), 1664–1680. doi: 10.1162/jocn.2010.21530

Salthouse, T. A., Fristoe, N., McGuthry, K. E., & Hambrick, D. Z. (1998). Relation of task switching to speed, age, and fluid intelligence. *Psychology and Aging*, 13(3), 445–461. doi: 10.1037/0882-7974.13.3.445

Santi, A., & Grodzinsky, Y. (2007). Working memory and syntax interact in Broca's area. *NeuroImage*, 37(1), 8–17. doi: 10.1016/j.neuroimage.2007.04.047

Schwartz, M. F., Marin, O. S., & Saffran, E. M. (1979). Dissociations of language function in dementia: A case study. *Brain and Language*, 7(3), 277–306. doi: 10.1016/0093-934X(79)90024-5

Schwartz, M. F., Saffran, E. M., & Marin, O. S. (1980). The word order problem in agrammatism: I. Comprehension. *Brain and Language*, 10(2), 249–262. doi: 10.1016/0093-934X(80)90055-3

Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-related cochlear synaptopathy: An early-onset contributor to auditory functional decline. *The Journal of Neuroscience*, 33(34), 13686–13694. doi: 10.1523/JNEUROSCI.1783-13.2013

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging*, 14(3), 458–472. doi: 10.1037/0882-7974.14.3.458

Stenfelt, S., & Rönnberg, J. (2009). The Signal-Cognition interface: Interactions between degraded auditory signals and cognitive processes. *Scandinavian Journal of Psychology*, 50(5), 385–393. doi: 10.1111/j.1467-9450.2009.00748.x

Stern, C., Prather, P., Swinney, D., & Zurif, E. (1991). The time course of automatic lexical access and aging. *Brain and Language*, 40(3), 359–372. doi: 10.1016/0093-934X(91)90135-N

Stine-Morrow, E. A., Milinder, L., Pullara, O., & Herman, B. (2001). Patterns of resource allocation are reliable among younger and older readers. *Psychology and Aging*, 16(1), 69–84. doi: 10.1037/0882-7974.16.1.69

Stine-Morrow, E. A., Shake, M. C., Miles, J. R., Lee, K., Gao, X., & McConkie, G. (2010). Pay now or pay later: Aging and the role of boundary salience in self-regulation of conceptual integration in sentence processing. *Psychology and Aging*, 25(1), 168–176. doi: 10.1037/a0018127

Stites, M. C., Federmeier, K. D., & Stine-Morrow, E. A. (2013). Cross-age comparisons reveal multiple strategies for lexical ambiguity resolution during natural reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(6), 1823–1841. doi: 10.1037/a0032860

Swaab, T. Y., Brown, C., & Hagoort, P. (1998). Understanding ambiguous words in sentence contexts: Electrophysiological evidence for delayed contextual selection in Broca's aphasia. *Neuropsychologia*, 36(8), 737–761. doi: 10.1016/S0028-3932(97)00174-7

Taylor, J. K., & Burke, D. M. (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. *Psychology and Aging*, 17(4), 662–676. doi: 10.1037/0882-7974.17.4.662

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–402). New York, NY: Academic Press.

Tun, P. A., Benichov, J., & Wingfield, A. (2010). Response latencies in auditory sentence comprehension: Effects of linguistic versus perceptual challenge. *Psychology and Aging*, 25(3), 730–735. doi: 10.1037/a0019300

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: The modulation of the frontotemporal language system in the context of age-related atrophy. *Cerebral Cortex*, 20(2), 352–364. doi: 10.1093/cercor/bhp105

Utman, J. A., Blumstein, S. E., & Sullivan, K. (2001). Mapping from sound to meaning: Reduced lexical activation in Broca's aphasics. *Brain and Language*, 79(3), 444–472. doi: 10.1006/brln.2001.2500

Van der Linden, M., Hupet, M., Feyereisen, P., Schelstraete, M., Bestgen, Y., Bruyer, R.,…Seron, X. (1999). Cognitive mediators of age-related differences in language comprehension and verbal memory performance. *Aging, Neuropsychology, and Cognition*, 6(1), 32–55. doi: 10.1076/anec.6.1.32.791

Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and 'theory of mind' in an agrammatic aphasic patient. *Current Biology*, 10(12), 723–726. doi: 10.1016/S0960-9822(00)00538-8

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging*, 18(2), 332–339. doi: 10.1037/0882-7974.18.2.332

Vigneau, M., Beaucousin, V., Herve, P., Duffau, H., Crivello, F., Houde, O.,… Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, 30(4), 1414–1432. doi: 10.1016/j.neuroimage.2005.11.002

Vonk, J., Jonkers, R., De Santi, S., & Obler, L. (2012). Object and action processing in Alzheimer's disease: The embodied view of cognition. *Procedia-Social and Behavioral Sciences*, 61, 218–219. doi: 10.1016/j.sbspro.2012.10.156

Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, 103(4), 761–772. doi: 10.1037/0033-295X.103.4.761

Waters, G. S., & Caplan, D. (1997). Working memory and on-line sentence comprehension in patients with Alzheimer's disease. *Journal of Psycholinguistic Research*, 26(4), 377–400. doi: 10.1023/A:1025073902269

Waters, G. S., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging*, 16(1), 128–144. doi: 10.1037/0882-7974.16.1.128

Waters, G. S., & Caplan, D. (2004). Verbal working memory and on-line syntactic processing: Evidence from self-paced listening. *Quarterly Journal of Experimental Psychology Section A*, 57(1), 129–163. doi: 10.1080/02724980343000170

Waters, G. S., Caplan, D., & Rochon, E. (1995). Processing capacity and sentence comprehension in patients with Alzheimer's disease. *Cognitive Neuropsychology*, 12(1), 1–30. doi: 10.1080/02643299508251990

Waters, G. S., Rochon, E., & Caplan, D. (1998). Task demands and sentence comprehension in patients with dementia of the Alzheimer's type. *Brain and Language*, 62(3), 361–397. doi: 10.1006/brln.1997.1880

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7(7), 701–702. doi: 10.1038/nn1263

Wingfield, A., & Lindfield, K. C. (1995). Multiple memory systems in the processing of speech: Evidence from aging. *Experimental Aging Research*, 21(2), 101–121. doi: 10.1080/03610739508254272

Wlotko, E. W., Lee, C., & Federmeier, K. D. (2010). Language of the aging brain: Event-related potential studies of comprehension in older adults. *Language and Linguistics Compass*, 4(8), 623–638. doi: 10.1111/j.1749-818X.2010.00224.x

Wulfeck, B., Bates, E., & Capasso, R. (1991). A cross-linguistic study of grammaticality judgments in Broca's aphasia. *Brain and Language*, 41(2), 311–336. doi: 10.1016/0093-934X(91)90158-W

Yoon, J., Campanelli, L., Goral, M., Marton, K., Eichorn, N., & Obler, L. K. (2015). The effect of plausibility on sentence comprehension among older adults and its relation to cognitive functions. *Experimental Aging Research*, 41(3), 272–302. doi: 10.1080/0361073X.2015.1021646

# 19 Neurolinguistic Studies of Sentence Comprehension

MICHAEL A. SKEIDE AND
ANGELA D. FRIEDERICI

Max Planck Institute for Human Cognitive and Brain Sciences

## Introduction

Language is a crucial evolutionary achievement of the Homo sapiens sapiens. Sophisticated auditory pattern learning skills have been discovered in non-human primates (Fitch & Hauser, 2004) and songbirds (Abe & Watanabe, 2011) and are discussed as precursors of the ability to acquire genuine linguistic structures. However, in order to comprehend language, it is essential to be able to discriminate, combine, and integrate a very complex set of linguistic elements, including acoustic, phonological, morphological, and lexical elements with their semantic, syntactic and prosodic information according to complex syntactic rules. This faculty is a uniquely human trait.

One of the decisive prerequisites for the emergence of language skills within the past 100,000 years of human phylogeny was the highly developed and thus language-ready Homo sapiens sapiens brain (Berwick, Friederici, Chomsky, & Bolhuis, 2013). The first systematic accounts of the neuroanatomical and neurophysiological foundations of language were provided in the nineteenth century when the notion of language-selective cortical areas and long-distance language networks was established based on ex vivo findings in aphasic patients (Broca, 1863; Wernicke, 1874; Lichtheim, 1885). With the availability of new in vivo techniques, particularly event-related potentials (ERPs) magnetic resonance imaging (MRI), our understanding of the neural basis of language has dramatically increased in the last 20 years. In this chapter we discuss the available neurolinguistic literature with the goal to provide a coherent picture of the diverse results.

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

## *Links and disparities between neurolinguistics and psycholinguistics*

Neurolinguistics as the study of the structural architecture and functional mechanisms of the brain underlying the processing of language, and psycholinguistics as the study of the mental representations and cognitive operations facilitating language processing are neighboring experimental linguistic disciplines.

Classical neurolinguistics has its roots in the nineteenth-century aphasiology, the study of deficits in language production and comprehension as a result of brain lesions. To date, the language-impaired brain is still a major neurolinguistic research topic (see Vonk, Higby, & Obler, this volume). Younger basic research focuses either on language processing in the mature brain of young adults or in the developing, that is, maturing or ageing brain (Skeide *et al.*, 2016a).

Modern neurolinguistic models are strongly theoretically informed by psycholinguistic models of sentence comprehension and incorporate concepts of both serial syntax-first and interactive constraint-satisfaction models (Marslen-Wilson & Tyler, 1980; Frisch, Hahne, & Friederici, 2004). Furthermore, neurolinguistic experiments frequently adapt behavioral experimental designs of psycholinguistic origin in order to relate controlled behavioral observations to brain measures.

## *Neurolinguistic methods*

Event-related brain potentials (ERPs) acquired with electroencephalography (EEG) systems are the most frequently used measure in neurolinguistics. ERPs quantify electrical activity of the cortex in response to a particular stimulus event with high temporal resolution in the order of milliseconds. Averaged electro-cortical activity appears as waveforms, in which so-called ERP components, with an either positive or negative inflection (polarity) relative to baseline, with a certain temporal latency after stimulus onset and with a characteristic but poorly resolved spatial distribution (topography) over the scalp can be identified. Based on these features, several components, discussed below, have been associated with particular stages of the language comprehension process. Magnetencephalography (MEG), a related neurophysiological method, records magnetic fields induced by electro-cortical activity. MEG provides information about the amplitude, latency and topography of language-related magnetic components with a temporal resolution comparable to ERPs but with an improved spatial resolution.

Magnetic resonance imaging (MRI) is another technique that is widely used for neurolinguistic experiments. It has replaced positron emission tomography (PET) as the state-of-the-art method for spatial reconstruction of the language network in the order of submillimeter. However, the temporal resolution of MRI is limited as it measures the hemodynamics changes (i.e., changes in blood flow, blood volume, and blood oxygenation) induced by brain activity, which takes place in the order of seconds. Functional MRI reveals precise information about the magnitude and the location of neural activity changes in response to external stimulation or intrinsic fluctuations at rest. These neural activity changes are reflected in blood-oxygen-level dependent (BOLD) signal changes based on the effect of neurovascular coupling. Structural MRI provides detailed morphometric and geometric features of the neural gray and white matter like its volume, density, thickness, and surface area. Diffusion-weighted MRI, especially diffusion tensor imaging, is used to reconstruct the trajectory and quantify tissue probabilities of white matter pathways interconnecting brain areas.

Near infrared spectroscopy allows for a more flexible recording of the BOLD response than MRI since the registration system is mounted directly on the participant's head. This advantage made it an important method for language acquisition research in infants and young children. However, the spatial resolution of near infrared spectroscopy is much lower than that of MRI while its temporal resolution is similarly poor.

Each of the mentioned non-invasive methods provides either fine-grained temporal or spatial information but not both. Currently, the best method to combine high temporal and spatial resolution within a single approach is to acquire EEG and MRI data simultaneously, which, however, requires an experimental paradigm suitable for both the high and the low temporal resolution method. Invasive techniques, particularly intracranial electrophysiology, overcome this issue but are exclusively feasible in clinical settings. Another limitation of the mentioned non-invasive methods is that they hardly allow the inference of causality. Nevertheless, causal relations between brain activity and behavior can be established non-invasively with neurostimulation methods such as transcranial magnetic stimulation and transcranial direct current stimulation. However, these techniques are limited to a few millimeters in terms of the spatial definition of their cortical target regions.

## Neurolinguistic studies of sentence comprehension

In the following section, we will give an overview of the organization and functionality of the brain structures that form the neural network underlying sentence comprehension according to the present state of research. Subsequently we will outline the two currently most influential neurolinguistic models of sentence comprehension. Finally, in the ensuing sections, which are the main focus of the chapter, we will report the main insights on the neural processing of syntax and semantics. Note that a detailed review of the literature on segmental (phonological) and suprasegmental (prosodic) acoustic processing of speech is beyond the scope of this chapter (for a review on phonology, see Hickok & Poeppel, 2007; Giraud & Poeppel, 2012; see also Pisoni, this volume); for a review on prosody, see Friederici, 2002, 2011; see also Pratt, this volume.

### *The sentence processing network*

Two gray matter regions underlying sentence processing, the left-lateralized inferior frontal cortex (Broca's area) and the posterior part of the superior temporal cortex (Wernicke's area), were already broadly defined by the pioneering nineteenth-century patient studies in postmortem brains (Broca, 1863; Wernicke, 1874). However, it is now assumed that the entire inferior frontal cortex—Brodmann Areas (BA) 44, 45, and 47, in addition to the frontal operculum (FOP)—is of major relevance for sentence comprehension (Hagoort, 2005; Friederici, 2012) and not only for production as Broca suggested. Furthermore, contributions of the temporal cortex to sentence comprehension were not only detected in the posterior part of the superior temporal gyrus (pSTG) identified by Wernicke, but in the entire STG including its mid (mSTG) and anterior (aSTG) portions as well as in the superior temporal sulcus (STS) and also in the middle temporal gyrus (MTG) (Friederici, 2012) (see Figure 16.1).

The IFG, FOP, STG, STS, and MTG form the core computational units of sentence processing across varying experimental designs and task demands (Caplan, Chen, & Waters, 2008). Nevertheless, all these larger cortical regions support multiple cognitive functions. However, there is evidence that within the left IFG, specific language-selective areas can be spatially disentangled from surrounding domain-general areas serving working memory, cognitive control and music processing (Makuuchi, Bahlmann, Anwander, & Friederici, 2009; Fedorenko, Behr, & Kanwisher, 2011; Fedorenko, Duncan, & Kanwisher, 2012). Notably, the tight functional relation between the frontal and temporal language regions is corroborated by their common genetic signatures (Johnson *et al.*, 2009) as well as their common distribution of different types of neurotransmitter systems (Zilles, Bacha-Trams, Palomero-Gallagher, Amunts, & Friederici, 2014).

The core sentence processing network has several cortical interfaces with sensory regions and, moreover, with memory-related regions in the inferior frontal sulcus (IFS) and inferior parietal cortex (IPC) (Meyer, Obleser, Anwander, & Friederici, 2012; Makuuchi & Friederici, 2013). Additionally, language is linked to several subcortical interfaces including the cerebellum, the basal ganglia and the thalamus that are most frequently interpreted as responsible for the temporal sequencing of speech (Kotz & Schwartze, 2010), but also for language-related cognitive control processes, particularly the left caudate nucleus (Friederici, 2006).

Within the framework of nineteenth-century aphasiology, Lichtheim already introduced the notion that the frontal and temporal language areas must be interconnected in order to provide their function as a network (Lichtheim, 1885). Indeed, recent fMRI studies demonstrated that the left BA 44 and the left FOP are effectively functionally connected with the left STG/STS (Lohmann *et al.*, 2010) and also with memory and vision interfaces (IFS, IPC and fusiform gyrus) during sentence comprehension (Makuuchi & Friederici, 2013). The basis for the signal exchange between the cortical gray matter regions underlying sentence processing are structural connections via dorsal and ventral white matter fiber tracts. Dorsally, BA 44 of the IFG is connected with the STG/STS and the MTG via the superior longitudinal fasciculus (SLF) including the arcuate fasciculus (AF) (Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006; Saur *et al.*, 2008). Ventrally, the inferior fronto-occipital fasciculus (IFOF) connects BA 45/47 with the STG/STS and the uncinate fasciculus (UF) connects the FOP with the aSTG/STS (Friederici, 2011) (Figure 16.1). The SLF/AF has reached a uniquely high degree of

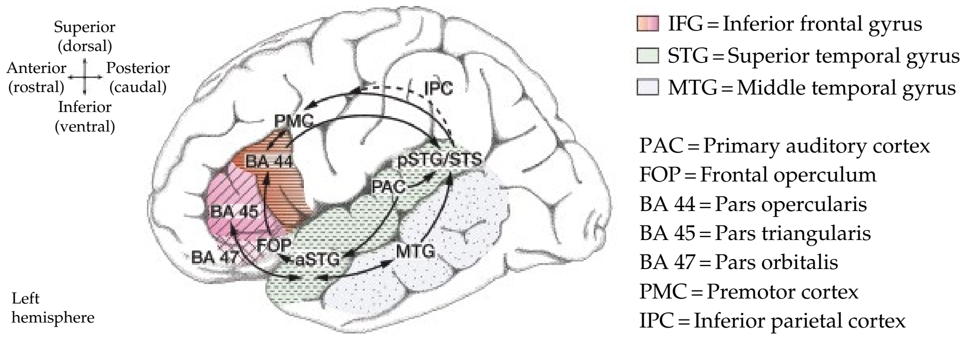



**Figure 19.1** The cortical sentence processing network (schematic view of the left hemisphere). The major areas involved in sentence processing are color coded. In the frontal cortex, five language-related regions are labeled: three Brodmann areas (BA 47, 45, 44), the premotor cortex (PMC) and the ventrally located frontal operculum (FOP). In the temporal and parietal cortex the following regions are labeled: the primary auditory cortex (PAC), the anterior (a) and posterior (p) portions of the superior temporal gyrus (STG) and sulcus (STS), the middle temporal gyrus (MTG) and the inferior parietal cortex (IPC). The solid black lines schematically indicate the direct pathways between these regions. The dashed black line indicates an indirect connection between the pSTG/STS and the PMC mediated by the IPC. The arrows indicate the assumed major direction of the information flow between these regions. During auditory sentence comprehension, information flow starts from PAC and proceeds from there to the anterior STG and via ventral connections along the uncinate fasciculus (UF) to the frontal cortex. Back-projections from BA 45 to anterior STG and MTG via ventral connections along the inferior fronto-occipital fasciculus (IFOF) are assumed to support top–down processes in the semantic domain, and the dorsal back-projection from BA 44 to posterior STG/STS along the superior longitudinal fasciculus (SLF) including the arcuate fasciculus (AF) are assumed to subserve top–down processes relevant for the assignment of grammatical relations. The dorsal pathway from PAC via pSTG/STS to the PMC is assumed to support auditory-to-motor mapping. Furthermore, within the temporal cortex, anterior and posterior regions are connected via the inferior and middle longitudinal fasciculi, branches of which may allow information flow from and to the mid-MTG.

differentiation in humans compared to their closest primate relatives (Rilling *et al.*, 2008). It is, moreover, not only phylogenetically specific but also ontogenetically specific since it only provides its full function for complex syntax once it is fully matured in young adults (Skeide, Brauer, & Friederici, 2015).

In addition to these interlobar long-distance connections between remote regions involved in sentence processing, numerous intralobar short-distance connections within neighboring cortical areas were identified in anatomical studies but their in vivo reconstruction is still in its very beginnings. Recently, short-distance connections were tracked within the left inferior frontal cortex (BA 44 and IFS) (Makuuchi *et al.*, 2009) and the left superior temporal cortex (mSTG and aSTG; mSTG and pSTG) (Upadhyay *et al.*, 2008). Language-related cortico-subcortical circuitry is even more understudied, but the first successful reconstructions of the connections between the inferior frontal cortex, the basal ganglia and the thalamus, respectively, have been recently reported (Jeon, Anwander, & Friederici, 2014). A major goal for future studies will be to provide a comprehensive picture of the connectivity and the microcircuitry of the neural networks involved in sentence comprehension.

## *Neurolinguistic models of sentence comprehension*

Two comprehensive neurolinguistic models of sentence comprehension covering the entire processing cascade from perception to interpretation were proposed recently: Friederici's model (Friederici, 2002, 2011, 2012) and Hagoort's Memory, Unification, Control (MUC) model (Hagoort, 2005, 2013). Both models assume that acoustic, syntactic and semantic processing levels are fundamental elements of the human language faculty (Figure 16.1). There is also a general agreement that the incoming speech signal initially undergoes spectrotemporal analyses in the auditory cortex and its vicinity (mSTG/STS). However, the neural implementation of subsequent phonological, syntactic and semantic processes is the aspect where the two models differ from each other. Moreover, unlike the MUC model, Friederici's model comprises a suprasegmental acoustic processing level. We first outline Friederici's model before turning to the MUC model.

Friederici's model comprises three initial processing stages of auditory sentence comprehension in the temporal cortex. These stages are dominated by unconscious bottom up mechanisms that occur early (onset within 20–120 ms), run fully automatically and proceed rapidly (within 30–60 ms). First, the incoming speech signal undergoes acoustic-phonological analyses in the left mSTG/STS within a time window of 20–50 ms until the phonological word forms are detected. Second, the phonological word form information is transferred from the left mSTG/STS to the left aSTG/STS along ventral short-distance fiber tracts, where it undergoes parallel lexical-semantic (50–80 ms) and morphosyntactic (40–90 ms) categorization. Third, once the lexical categories are identified, the corresponding lexicon entry can be accessed and retrieved (110–170 ms). This process is not confined to the left aSTG/STS but also includes the left mSTG. Once the syntactic categories are identified, phrase structures can be reconstructed (120–150 ms) in the left aSTG/STS and the left FOP connected by the UF. Suprasegmental prosodic information is processed in the right superior temporal cortex in a similar time window after delivery of the phonological word form information via transcallosal fibers. Higher-level sentence comprehension is characterized by the additional involvement of the inferior frontal cortex. It can be subdivided into two final stages reflecting consciously controllable top down processes that occur typically within 200–600 ms but also later, that not necessarily run automatically and that usually proceed relatively slowly (150–200 ms or beyond). First, semantic relations between lexical units and syntactic relations between phrases have to be determined based on verb-argument-related predictions. Lexical information is delivered from the left aSTG/STS and the left mSTG along the IFOF to the anterior portion (pars triangularis, BA 45) and the inferior portion (pars orbitalis, BA 47) of the left IFG where the contextual compatibility between lexical items is evaluated (200–400 ms). In parallel, phrase structure information is transmitted from the left aSTG/STS and the left FOP to the posterior portion (pars opercularis) of the left IFG (BA 44) along the UF and ventral short-distance fiber tracts. BA 44 and the bordering posterior part of BA 45 crucially support the (re-)ordering of phrases in a sentence at 300–500 ms. Complex suprasegmental acoustic information, comprising intonation and accentuation, is processed within a later time window of 400–600 ms in the right superior temporal cortex. Second, in order to enable sentence comprehension, syntactic and semantic information are integrated in the posterior temporal cortex at around 600 ms. Therefore, syntactic information is sent from the left BA44/45 to the left pSTG/STS along the SLF/AF and semantic information is sent from the left BA45/47 to the left pSTG/STS along the IFOF. However, the role of the SLF/AF and IFOF in the model is not limited to the integration of syntax and semantics. The SLF/AF is assumed to be involved in the hierarchization of phrases and in verb-argument related predictions (within 400–600 ms) for which an information exchange between the left BA44/45 and the left pSTG/STS is necessary. The IFOF is considered to mediate the strategically controlled access to lexical-semantic features in the MTG under regulation of BA 45 and 47 before integration takes place.

Hagoort assumes that auditory language comprehension is driven by multiple processing cycles that are repeated until an utterance can be interpreted. The starting point of the first stage of the first cycle is the arrival of the speech signal in the left primary auditory cortex. Subsequently, along a dorsal-to-ventral gradient in the left posterior temporal cortex, information about acoustic-phonological properties is retrieved from an area covering the left pSTG/STS, then information about syntactic properties is retrieved from a region spanning the left pSTS to the pMTS and finally information about conceptual semantic properties of the signal is retrieved from an area extending from the left pMTG to the pITG. Depending on the type of linguistic information, another dorsal-to-ventral gradient activates in the left inferior parietal cortex with phonological representations retrieved from the left supramarginal gyrus (SMG), syntactic representations retrieved from an area between the left SMG and the left angular gyrus (AG) and semantic representations retrieved from the AG. These retrieval mechanisms, the core of the so-called memory component, operate sequentially at different, but partially overlapping time scales until they are completed at around 250 ms. During the second stage of the first cycle, phonological, syntactic, and semantic information is relayed from the left posterior temporal and inferior parietal cortex to the left IFG over a feed-forward pathway including the SLF and the AF. This pathway is mediated by fast AMPA and $GABA_A$ transmission streams in which the signals rapidly decay. The IFG in turn crucially supports the maintenance of the information due to its ability for self-sustained firing. Furthermore, it is involved in the rule-constrained combination of linguistic elements from all three domains along a posterior-dorsal to anterior-ventral gradient with the posterior dorsal BA44 and the neighboring ventral BA6 unifying phonological word forms into intonational phrases, the anterior ventral BA44 and posterior dorsal BA45 linking together syntactic phrases and the anterior ventral BA45 and BA47 integrating lexical items into the discourse context. This so-called unification component selects the optimal candidates out of multiple possible environments and optional alternative links. As soon as the selection procedure is over, the information is sent back from the left IFG to the left posterior temporal cortex along a slow feedback pathway comprising the UF, the IFOF and the inferior longitudinal fasciculus (ILF) until the first cycle is completed at around 400 ms. This pathway strengthens the role of the IFG for the maintenance of linguistic information since it is mediated by an NMDA transmission system in which the signals decay relatively slowly. Usually, full comprehension requires additional recurrent feedforward and feedback signaling cycles between the posterior temporal and the inferior frontal cortex. The higher the competition among suitable candidate combinations, that is, the higher the unification load, the longer it takes until a final interpretation of an utterance can be achieved. Finally, successful comprehension is guided by a so-called control system represented in the anterior cingulate cortex, the dorsolateral prefrontal cortex and the inferior parietal cortex that subserves attentional and pragmatic operations such as attention allocation to a target language during language switching or turn taking in a conversation.

## Syntax

### *Low-level syntactic feature detection*

MEG work strongly suggests that the human brain responds to morphosyntactic category errors as early as between 40–90 ms after speech input reflected by an early left anterior negativity (ELAN) (Herrmann, Maess, Hahne, Schröger, & Friederici, 2011). Additionally, it was recently shown in an ERP study that this information is processed implicitly, in the absence of awareness (Batterink & Neville, 2013). Accordingly, auditory language comprehension is initially driven by unconscious bottom up feature detection mechanisms triggered automatically by low-level syntactic properties such as phrase types. There is converging evidence from MEG and fMRI studies that the anterior portion of the left superior temporal cortex forms the neural basis of these recognition mechanisms (Friederici, Wang, Herrmann, Maess, & Oertel, 2000; Friederici, Rüschemeyer, Hahne, & Fiebach, 2003; Snijders *et al.*, 2009; Herrmann *et al.*, 2011; DeWitt & Rauschecker, 2012; Brennan *et al.*, 2012).

Although first traces of the ELAN can be detected in very early time windows, it is well-documented that it unfolds in time until 180–200 ms after stimulus onset (Friederici, Pfeifer, & Hahne, 1993; Friederici *et al.*, 2000; Herrmann *et al.*, 2011). Given the relatively long temporal extension of the component, it is unlikely that its functional role is limited to the identification of syntactic categories. Instead, it was argued that the ELAN is also related to the reconstruction of the internal structure of a phrase (Friederici, 2011, 2012). These initial local phrase structure building mechanisms seem to recruit the left anterior temporal cortex which is assumed to work in cooperation with the FOP (Friederici *et al*., 2003, 2006).

The question how low-level syntactic information is propagated from the left anterior superior temporal cortex and the FOP to the left inferior frontal cortex for high-level processing cannot be answered definitively given the existing evidence. Nevertheless, although there is currently no study available that directly investigated this problem, most reviews suggest that the information flows along a ventral tract, namely either the UF or the extreme capsule fiber system (ECFS) (Hickok & Poeppel, 2007; Weiller, Musso, Rijntjes, & Saur, 2009; Friederici, 2012). Hagoort, however, assumes that low-level syntactic information is transferred along a dorsal pathway comprising the SLF/AF (Hagoort, 2013).

## *High-level syntactic computations*

There is broad agreement in the literature that the left temporal cortex alone cannot provide the full range of functions required for syntax processing but that the left inferior frontal cortex is necessarily involved at higher processing levels where the structural relations between phrases must be analyzed. However, some researchers hold the position that it is not the left inferior frontal cortex but only the anterior temporal cortex that is specifically involved in syntactic computation (Rogalsky & Hickok, 2011).

Within the inferior frontal cortex, BA44 and BA45 are the subregions that were most frequently associated with syntax (Caplan *et al*., 2008; Makuuchi *et al*., 2009, 2013; Santi & Grodzinsky, 2010; Newman *et al*., 2010; Tyler *et al*., 2011; Meyer *et al*., 2012; Kinno, Ohta, Muragaki, Maruyama, & Sakai, 2014) although a few studies also reported syntax-related activity in neighboring areas, namely BA47 (Pallier, Devauchelle, & Dehaene, 2011; Kinno *et al*., 2014) and BA6 (Kinno *et al*., 2014). The particular functions ascribed to BA44 and/or BA45 reflect either a general involvement, termed syntactic processing, syntactic parsing, or syntactic structuring, or a more specific involvement, termed combination of syntactic elements, syntactic movement, evaluation of hierarchical relations between phrases, or reordering of phrases. Despite ongoing controversies regarding the exact implementation of syntax processing in the left inferior frontal cortex, the clear majority of studies on the topic shares the general assumption that BA44 and/or BA45 are decisively involved in the assignment of syntactic relations between morphosyntactically pre-categorized elements of a sentence. In free word order languages, such as German and Japanese, early morphosyntactic and later syntactic processing can be sharply dissociated from each other into the ELAN component and a so-called left anterior negativity (LAN) component peaking between 300–500 ms. In fixed word order languages, such as English and Dutch, however, in which word order and not morphosyntactic marking is the main cue for the assignment of syntactic relations in a sentence, the LAN is found less frequently (Friederici, 2002, 2011; see Friederici & Weissenborn, 2007 for an overview).

A coherent picture of the factors potentially determining the precise anatomical location of high-level syntactic computational processes within the inferior frontal cortex has not emerged yet. Studies using German sentences consistently revealed confined activation in BA44 (Makuuchi *et al.*, 2009, 2013; Brauer, Anwander, & Friederici, 2011; Meyer *et al.*, 2012; Skeide *et al.*, 2016b; Skeide *et al.*, 2014). Hence, it could be hypothesized that free word order languages in general recruit BA44 rather than BA45. This view, however, was challenged by noun phrase scrambling studies in languages with a more fixed word order such as Japanese or English reporting activation both in BA44 and BA45 (Caplan *et al.*, 2008; Kinno, Kawamura, Shioda, & Sakai, 2008; Santi & Grodzinsky, 2010; Newman, Ikuta, & Burns Jr., 2010; Tyler *et al.*, 2011, Kinno *et al.*, 2014). Disentangling the specific contributions of BA44 and BA45 based on distinct syntactic features to be processed appears to be a more fruitful approach. Friederici (2012) hypothesized that the (re-)ordering of morphosyntactically marked phrases in a sentence mainly involves BA44 whereas the movement of elements from subordinate sentence parts mainly involves BA45. Future studies could test this hypothesis by applying multivariate analyses, since these will have more sensitivity than mass-univariate analyses to how specific syntactic features drive the spatial distribution of neural activity. Multivariate analyses might also facilitate the isolation of core activation sources within largely overlapping BOLD signal distributions.

Most authors agree that sentence comprehension depends not only on ventral connections between the temporal and the inferior frontal cortex but also on a dorsal connection along the SLF/AF. Nevertheless, there are currently two open issues debated in the literature, namely first, what the anatomically valid subsegments of the SLF/AF are and second, what the exact cortical termination areas of the SLF/AF within the frontal, temporal, and parietal lobe are. Despite the existing controversies that cannot be fully discussed here, diffusion tensor imaging studies consistently reconstructed a coherent tract with a roughly similar trajectory described in classical postmortem dissection studies connecting the posterior inferior frontal and the mid and superior temporal cortices (Burdach, 1822; Dejerine, 1885; Catani, Jones, & ffytche, 2005; Makris *et al.*, 2005; Friederici *et al.*, 2006; Rilling *et al.*, 2008; Wilson *et al.*, 2011; Brown *et al.*, 2014; Fernández-Miranda *et al.*, 2015).

The exact division of labor between the dorsal and ventral pathway with respect to syntax processing is controversially discussed in the literature. There are advocates of the view that both pathways equally contribute to the transmission of syntactic information (Rolheiser, Stamatakis, & Tyler, 2011; Papoutsi, Stamatakis, Griffiths, Marslen-Wilson, & Tyler, 2011; Griffiths, Marslen-Wilson, Stamatakis, & Tyler, 2012) but also advocates of the view that each pathway supports different syntactic aspects (Wilson *et al.*, 2011; Friederici, 2012; Hagoort, 2013). Currently, there is stronger evidence for specific roles of each pathway. The studies suggesting similar roles of the dorsal and the ventral pathway in sentence comprehension either used syntax manipulations that were not controlled for possibly interfering semantic processes (Papoutsi *et al.*, 2011; Griffiths *et al.*, 2012) or reported effects that were much more pronounced in the dorsal tract compared to the ventral tract (Rolheiser *et al.*, 2011). These limitations were overcome in a semantic-free artificial grammar learning experiment (Friederici *et al.*, 2006) and in an experiment on primary progressive aphasia patients (see Vonk, Higby, & Obler, Chapter 18, this volume) with uniquely confined lesions either to the SLF/AF or the IFOF (Wilson *et al*., 2011). Both studies provide converging evidence for a specific role of the dorsal pathway in comprehending syntactically complex sentences. This link between complex syntax and the SLF/AF was corroborated in a recent developmental study (Skeide *et al*., 2016b). Nevertheless, it is generally hard to determine at present what the precise functional role of the dorsal pathway within language network is since the information flow along a certain fiber tract cannot be directly measured but only indirectly inferred by correlation to functionally relevant measures or reconstruction in a model. Two recent studies applying dynamic causal modelling (DCM) suggest that, during high‐level sentence comprehension, information flows unidirectionally from the left inferior frontal to the left posterior temporal cortex via a dorsal pathway (den Ouden *et al*., 2012; Makuuchi *et al*., 2013). This observation fits the assumption that the SLF/AF might support the transmission of syntactic information from BA44/45 to the posterior temporal cortex where it can be finally integrated with semantic information as reflected in the P600 known from the ERP literature (Friederici, 2012). More sophisticated modeling of linguistic representations and their possible physiological implementations is necessary for a more direct testing of this hypothesis in future studies.

## Semantics

### *Lexical categorization and lexical‐semantic access*

Similar to syntactic feature detection, lexical‐semantic categorization is a rapid process (Marslen‐Wilson, 1973; Rayner & Clifton, 2009). A recent MEG experiment demonstrated that the recognition of a word form's lexical status (words versus pseudowords) elicits cortical responses already at around 50–80 ms in bilateral temporal sources (MacGregor, Pulvermüller, van Casteren, & Shtyrov, 2012). Hence, not only early morphosyntactic but also early lexical‐semantic processing, like the assignment of the lexical status, is driven by unconscious bottom up feature detection mechanisms triggered automatically by low‐level domain‐specific properties.

The actual access to lexical‐semantic representations occurs in a later time window of around 110–170 ms (MacGregor *et al*., 2012) in several distinct brain regions varying as a function of the sensory input modality (Fiebach & Friederici, 2003). Fast lexical‐semantic access to spoken words is associated with the left STG, STS and MTG in the vicinity of the primary auditory cortex (Indefrey, P., Hagoort, P., Herzog, H., Seitz, R.J., & Brown, 2001; Fiebach & Friederici, 2003; Dronkers, Wilkins, Van Valin Jr., Redfern, & Jaeger, 2004; Gold *et al*., 2006; Gagnepain *et al*., 2008). Accordingly, it has been suggested that easy and quickly accessible lexical items, that is, particularly concrete words that are learned early during language acquisition, are represented in the brain in a sensory manner (Fiebach & Friederici, 2003). Some authors have reported right‐hemispheric contributions to lexical-semantic access, especially in the right temporal cortex (MacGregor *et al.*, 2012) but more experimental work is necessary to corroborate the specificity of these observations.

Accessing lexical-semantic representations is not always an automatic process but can also be under strategic control guided by predictions about lexical properties. Consciously controlled top down mechanisms are consistently related to BA45 and BA47 of the left inferior frontal gyrus. These areas were shown to be selectively modulated by word frequency with stronger hemodynamic activity for low frequency words compared to high frequency words (Chee, Hon, Caplan, Lee, & Goh, 2002; Fiebach, Friderici, Müller, & von Cramon, 2002) as well as by age of word acquisition and abstractness with stronger hemodynamic activity for late versus early learned words and abstract versus concrete words (Fiebach *et al.*, 2002; Fiebach & Friederici, 2003). There is extensive evidence for the hypothesis that BA45/47 mediates the top–down retrieval of lexical-semantic representations stored in the MTG (Lau, Phillips, & Poeppel, 2008). Lexical-semantic access might recruit more parts of the cortex linked to semantic representations, as has been discussed in several reviews (Patterson, Nestor, & Rogers, 2007; Binder, Desai, Graves, & Conant, 2009) and observed in recent experiments on the cortical distribution of perceptual object and action categories (Huth, Nishimoto, Vu, & Gallant, 2012). However, it is broadly assumed that the language-specific lexicon and general conceptual representations can be separated from another at the neural level although they are thought to form a continuum of semantic knowledge or semantic memory (Patterson *et al.*, 2007). The most comprehensive meta-analytic comparison of lexical semantics (i.e., representations acquired through language) against perceptual semantics (i.e., representations of concrete objects derived from sensory-motor experience), revealed that the left BA47 and the anterior STS are specifically related to lexical semantics (Binder *et al.*, 2009).

Interestingly, the left anterior temporal cortex, particularly the left aSTS seems to be selectively and specifically related to both syntax and semantics at low-level processing stages. Several authors emphasized its involvement in combinatorial mechanisms across these linguistic modalities (Hickok & Poeppel, 2007; Saur *et al.*, 2008). It was argued, that the anterior portion of this region is sensitive to syntax whereas a more posterior part is sensitive to both syntax and semantics (Humphries, Binder, Medler, & Liebenthal, 2006). A recent meta-analysis corroborated the modality specificity of the anterior temporal cortex indicating that linguistic stimuli recruit its lateral part whereas visual stimuli rather recruit its ventral part (Visser, Jefferies, & Lambon Ralph, 2010).

## *High-level semantic processing*

Successful sentence comprehension requires a speaker not only to access and retrieve lexical-semantic information of a given word but also to analyze semantic relations between lexical items in certain sentence contexts. This top–down language function is well established in the ERP literature where it is assumed to be reflected in the N400 component (Kutas & Hillyard, 1980, 1984; Lau *et al.*, 2008; MacGregor *et al.*, 2012).

High-level semantic processing is most frequently ascribed to the anterior portion of the left inferior frontal gyrus including BA45 and BA47 (Friederici, 2012). These areas must receive information from preceding semantic processing stages. With respect to auditory sentence comprehension there is cumulating evidence, most directly from patient studies, that a ventral pathway along the IFOF (or ECFS) supports the transfer of lexical-semantic information from the anterior temporal cortex, where phonological word forms are mapped to meaning (Tsapkini, Frangakis, & Hillis, 2011), to BA45 and BA47 (Weiller *et al.*, 2009; Weiller, Bormann, Saur, Musso, & Rijntjes 2011).

Semantic context effects in the left BA45 and BA47 were demonstrated for a broad range of semantic aspects at the sentence level, including ambiguity (Rodd, Davis, & Johnsrude, 2005), predictability (Obleser, Wise, Dresner, & Scott, 2007) and conceptual relatedness (Newman *et al.*, 2010). Activation in the left BA47 in response to contextual semantic manipulations was repeatedly found to go along and covary with activation in the inferior parietal cortex, particularly the AG and its vicinity (Obleser *et al.*, 2007; Uddin *et al.*, 2010; Xiang, Fonteijn, Norris, & Hagoort, 2010). Accordingly several authors emphasized the role of the AG in the retrieval of semantic knowledge (Binder *et al.*, 2009; Hagoort *et al.*, 2013) and/or the contextual integration of complex semantic information (Lau *et al.*, 2008; Binder *et al.*, 2009; Friederici, 2012). However, the question of if and how BA45/47 and the AG can be functionally differentiated in terms of specific contributions to high-level semantic processing cannot be answered given the existing evidence. Weiller *et al.* (2011) suggested that BA45/47 and the AG might form a fronto-parietal network engaged not only in verbal working memory but also in the top down control of semantic processing in the temporal cortex. Even if this assumption is valid, further work is needed to clarify the open issue of the division of labor between BA45/47 and the AG. Moreover, future studies will have to provide conclusive evidence that language-specific selectivity for semantics can be disentangled from domain general conceptual activity within the AG given that it is a highly multimodal cortical region (Binder *et al.*, 2009).

In the ERP literature, the P600 marks a final step of sentence comprehension (Friederici, 2011). Attempts to localize this component revealed sources in the posterior MTG (Kwon *et al.*, 2005) and the posterior STG at the border to the inferior parietal cortex (Service, Helenius, Maury, & Salmelin, 2007). Accordingly, if the hypothesis that the P600 reflects the integration of semantic information into syntactic constructions is true (Gunter, Friederici, & Schriefers, 2000; Kuperberg *et al.*, 2003), it is unlikely that these integration processes take place in the IFG or at least it is unlikely that the syntax-semantics interface exclusively involves the IFG without any temporo-parietal contributions. The flow of semantic information between BA45/47 and the posterior superior temporal and/or the inferior parietal cortex for integration and interpretation could recruit a direct connection along the ECFS as proposed in two recent reviews (Weiller *et al.*, 2011; Friederici, 2012). This plausible position must await further support from more direct experimental results.

## Summary

Auditory sentence comprehension can be divided into two classes of neural mechanisms. First, unconsciously proceeding bottom up mechanisms that are triggered externally and automatically in domain-selective regions within the temporal cortex by salient basal domain-specific features enable a fast and partly parallel categorization of the speech input. Second, consciously controllable mechanisms that are generated internally in domain-selective regions within the inferior frontal cortex supported by multimodal temporo-parietal and subcortical structures allow higher-order bottom up and predictive top down assignment of complex relations between the elements detected in a sentence and finally their integration into a conceptual whole. Both mechanisms are implemented in dorsal, ventral and cortico-subcortical networks of fiber tracts ensuring a rapid flow of information between the involved brain areas. The two neural mechanisms of auditory sentence comprehension reflect two aspects of the evolutionary advantage of human language. Language is not only an efficient medium for rapid information exchange but also a unique tool for sophisticated conceptual representation of the world and therefore it maximizes the chance of survival in a dynamic and complex environment.


## REFERENCES


Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, 14(8), 1067–1074. doi: 10.1038/nn.2869

Batterink, L., & Neville, H. (2013). The human brain processes syntax in the absence of conscious awareness. *The Journal of Neuroscience*, 33(19), 8528–8533. doi: 10.1523/JNEUROSCI.0618-13.2013

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain and the nature of language. *Trends in Cognitive Sciences*, 17(2), 89–98. doi: 10.1016/j.tics.2012.12.002

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796. doi: 10.1093/cercor/bhp055

Brauer, J., Anwander, A., & Friederici, A. D. (2011). Neuroanatomical prerequisites for language functions in the maturing brain. *Cerebral Cortex*, 21(2), 459–466. doi: 10.1093/cercor/bhq108

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., & Pylkkänen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain & Language*, 120(2), 163–173. doi: 10.1016/j.bandl.2010.04.002

Broca, P. (1863). Localisations des fonctions cérébrales. Siège de la faculté du langage articulé. *Bulletin de la Société d' Anthropologie de Paris*, 4, 200–204.

Brown, E. C., Jeong, J. W., Muzik, O., Rothermel, R., Matsuzaki, N., Juhász, C., … Asano, E. (2014). Evaluating the arcuate fasciculus with combined diffusion-weighted MRI tractography & electrocorticography. *Human Brain*

*Mapping*, 35(5), 2333–2347. doi: 10.1002/hbm.22331

Burdach, C. F. (1822). *Vom Baue und Leben des Gehirns und Rückenmarks*. Leipzig: Dyk.

Caplan, D., Chen, E., & Waters, G. (2008). Task-dependent and task-independent neurovascular responses to syntactic processing. *Cortex*, 44(3), 257–275. doi: 10.1016/j.cortex.2006.06.005

Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, 57(1), 8–16. doi: 10.1002/ana.20319

Chee, M. W. L., Hon, N. H. H., Caplan, D., Lee, H. L., & Goh, J. (2002). Frequency of concrete words modulates prefrontal activation during semantic judgments. *Neuroimage*, 16(1), 259–268. doi: 10.1006/nimg.2002.1061

Dejerine, J. (1895). *Anatomie des centres nerveux*. Paris: Rueff et Cie.

Den Ouden, D. B., Saur, D., Mader, W., Schelter, B., Lukic, S., Wali, E., … Thompson, C. K. (2012). Network modulation during complex syntactic processing. *Neuroimage*, 59(1), 815–823. doi: 10.1006/nimg.2002.1061

DeWitt, I., & Rauschecker, J. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences of the USA*, 109(8), 2709–2710. doi: 10.1073/pnas.1113427109

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D. Jr, Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, 92(1–2), 145–177. doi: 10.1016/j.cognition.2003.11.002

Fedorenko, E., Behr, M. K., & Kanwisher N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences of the USA*, 108(39), 16428–16433. doi: 10.1073/pnas.1112937108

Fedorenko, E., Duncan, J., & Kanwisher N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, 22(21), 2059–2062. doi: 10.1016/j.cub.2012.09.011

Fernández-Miranda, J. C., Wang, Y., Pathak, S., Stefaneau, L., Verstynen, T., & Yeh, F. C. (2015). Asymmetry, connectivity, and segmentation of the arcuate fascicle in the human brain. *Brain structure & function*, 220(3), 1665–1680. doi: 10.1007/s00429-014-0751-7

Fiebach, C. J., Friederici, A. D., Müller, K., & von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, 14(1), 11–23. doi: 10.1162/089892902317205285

Fiebach, C. J., & Friederici, A. D. (2003). Processing concrete words: fMRI evidence against a specific right-hemisphere involvement. *Neuropsychologia*, 42(1), 62–70. doi: 10.1016/S0028-3932(03)00145-3

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303(5656), 377–380. doi: 10.1126/science.1089401

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1(3), 183–192. doi: 10.1016/0926-6410(93)90026-2

Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping*, 11(1), 1–11. doi: 10.1002/1097-0193(200009)11:1<1::AID-HBM10>3.0.CO;2-B

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78–84. doi: 10.1016/S1364-6613(00)01839-8

Friederici, A. D., Rüschemeyer, S.-A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic & semantic processes. *Cerebral Cortex*, 13(2), 170–177. doi: 10.1093/cercor/13.2.170

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the USA*, 103(7), 2458–2463. doi: 10.1073/pnas.0509389103

Friederici, A. D. (2006). What's in control of language? *Nature Neuroscience*, 9(8), 991–992. doi: 10.1038/nn0806-991

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, 91(4),1357–1392. doi: 10.1152/physrev.00006.2011

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262–268. doi: 10.1016/j.tics.2012.04.001

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax-semantic interface. *Brain Research*, 1146, 50–58. doi: 10.1016/j.brainres.2006.08.038

Frisch, S., Hahne, A., & Friederici, A. D. (2004). Word category and verb–argument structure information in the dynamics of parsing. *Cognition*, 91(3), 191–219. doi: 10.1016/j.cognition.2003.09.009

Gagnepain, P., Chételat, G., Landeau, B., Dayan, J., Eustache, F., & Lebreton, K. (2008). Spoken word memory traces within the human auditory cortex revealed by repetition priming and functional magnetic resonance imaging. *The Journal of Neuroscience*, 28(20), 5281–5289. doi: 10.1523/JNEUROSCI.0565-08.2008

Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Andersen, A. H. (2006). Dissociation of automatic and strategic lexical-semantics: functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *The Journal of Neuroscience*, 26(24), 6523–6532. doi: 10.1523/JNEUROSCI.0808-06.2006

Giraud, A. L., & Poeppel D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–507. doi: 10.1038/nn.3063

Griffiths, J. D., Marslen-Wilson, W. D., Stamatakis, E. A., & Tyler, L. K. (2012). Functional organization of the neural language system: dorsal and ventral pathways are critical for syntax. *Cerebral Cortex*, 23(1), 139–147. doi: 10.1093/cercor/bhr386

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender & semantic expectancy: ERPs reveal early autonomy & late interaction. *Journal of Cognitive Neuroscience*, 12(4), 556–568.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, 9(9), 416–423. doi: 10.1016/j.tics.2005.07.004

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, 4(416), 1–13. doi: 10.3389/fpsyg.2013.00416

Herrmann, B., Maess, B., Hahne, A., Schröger, E., & Friederici, A. D. (2011). Syntactic and auditory spatial processing in the human temporal cortex: An MEG study. *Neuroimage*, 57(2), 624–633. doi: 10.1016/j.neuroimage.2011.04.034

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402. doi: 10.1038/nrn2113

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6), 1210–1224. doi: 10.1016/j.neuron.2012.10.014

Humphries, C., Binder, J.R., Medler, D.A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, 18(4), 665–679. doi: 10.1162/jocn.2006.18.4.665

Indefrey, P., Hagoort, P., Herzog, H., Seitz, R. J., & Brown, C. M. (2001). Syntactic processing in left prefrontal cortex is independent of lexical meaning. *Neuroimage*, 14(3), 546–555. doi: 10.1006/nimg.2001.0867

Jeon, H.-A., Anwander, A., & Friederici, A. D. (2014). Functional network mirrored in the prefrontal cortex, caudate nucleus and thalamus: high resolution functional imaging and structural connectivity. *The Journal of Neuroscience*, 34(28), 9202–9212. doi: 10.1523/JNEUROSCI.0228-14.2014

Johnson, M. B., Kawasawa, Y. I., Mason, C. E., Krsnik, Z., Coppola, G., Bogdanovic, D., … Sestan, N. (2009). Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron*, 62(4), 494–509. doi: 10.1016/j.neuron.2009.03.027

Kinno, R., Kawamura, M., Shioda, S., & Sakai, K. L. (2008). Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task. *Human Brain Mapping*, 29(9), 1015–1027. doi: 10.1002/hbm.20441

Kinno, R., Ohta S, Muragaki Y, Maruyama T, Sakai K. L. (2014). Differential reorganization of three syntax-related networks induced by a left frontal glioma. *Brain*, 137(4), 1193–1212. doi: 10.1093/brain/awu013

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences*, 14(9), 392–399. doi: 10.1016/j.tics.2010.06.005

Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, 15(2), 272–293. doi: 10.1162/089892903321208204

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–208.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307, 161–163.

Kwon, H., Kuriki, S., Kim, J. M., Lee, Y. H., Kim, K., & Nam, K. (2005). MEG study on neural activities associated with syntactic and semantic violations in spoken Korean sentences. *Neuroscience Research*, 51(4), 349–357. doi: 10.1016/j.neures.2004.12.017

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de) constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933. doi: 10.1038/nrn2532

Lichtheim, L. (1885). Über Aphasie. *Deutsches Archiv für klinische Medicin*, 36, 204–268.

Lohmann, G., Hoehl, S., Brauer, J., Danielmeier, C., Bornkessel-Schlesewsky, I., Bahlmann, J., … Friederici, A. D. (2010). Setting the frame: The human brain activates a basic low-frequency network for language processing. *Cerebral Cortex*, 20(6), 1286–1292. doi: 10.1093/cercor/bhp190

MacGregor, L. J., Pulvermüller, F., van Casteren, M., & Shtyrov, Y. (2012). Ultra-rapid access to words in the brain. *Nature Communications*, 3(711). doi: 10.1038/ncomms1715

Makris, N., Kennedy, D. N., McInerney, S., Sorensen, A. G., Wang, R., Caviness Jr., V. S., Pandya, D. N. (2005). Segmentation of subcomponents within the superior longitudinal fascicle in humans: A quantitative, in vivo, DT-MRI study. *Cerebral Cortex*, 15(6), 854–869. doi: 10.1093/cercor/bhh186

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences of the USA*, 106(20), 8362–8367. doi: 10.1073/pnas.0810928106

Makuuchi, M. & Friederici, A. D. (2013). Hierarchical functional connectivity between the core language system & the working memory system. *Cortex*, 49(9), 2416–2423. doi: 10.1016/j.cortex.2013.01.007

Marslen-Wilson, W. D. (1973). Linguistic structure & speech shadowing at very short latencies. *Nature*, 244, 522–523.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8(1), 1–71.

Meyer, L., Obleser, J., Anwander, A., & Friederici, A.D. (2012). Linking ordering in Broca's area to storage in left temporo-parietal regions: The case of sentence processing. *Neuroimage*, 62(3), 1987–1998. doi: 10.1016/j.neuroimage.2012.05.052

Newman, S. D., Ikuta, T., & Burns Jr., T. (2010). The effect of semantic relatedness on syntactic analysis: An fMRI study. *Brain & Language*, 113(2), 51–58. doi: 10.1016/j.bandl.2010.02.001

Obleser, J., Wise, R. J., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *The Journal of Neuroscience*, 27(9), 2283–2289. doi: 10.1523/JNEUROSCI.4663-06.2007

Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences of the USA*, 108(6), 2522–2527. doi: 10.1073/pnas.1018711108

Papoutsi, M., Stamatakis, E.A., Griffiths, J., Marslen-Wilson, W.D., & Tyler, L.K. (2011). Is left fronto-temporal connectivity essential for syntax? Effective connectivity, tractography and performance in left-hemisphere damaged patients. *Neuroimage*, 58(2), 656–664. doi: 10.1016/j.neuroimage.2011.06.036

Patterson, K., Nestor, P.J., & Rogers, T.T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976–987. doi: 10.1038/nrn2277

Rayner, K., & Clifton, C.J. (2009). Language processing in reading & speech perception is fast & incremental: implications for event related potential research. *Biological Psychiatry*, 80(1), 4–9. doi: 10.1016/j.biopsycho.2008.05.002

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., Behrens, T. E. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, 11, 426–428. doi: 10.1038/nn2072

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, 15(8), 1261–1269. doi: 10.1093/cercor/bhi009

Rogalsky, D., & Hickok, G. (2011). The role of Broca's area in sentence comprehension. *Journal of Cognitive Neuroscience*, 23(7), 1664–1680. doi: 10.1162/jocn.2010.21530

Rolheiser, T., Stamatakis, E. A., & Tyler, L. K. (2011). Dynamic processing in the human language system: synergy between the arcuate fascicle and extreme capsule. *The Journal of Neuroscience*, 31(47), 16949–16957. doi: 10.1523/JNEUROSCI.2725-11.2011

Santi, A., & Grodzinsky, Y. (2010). FMRI adaptation dissociates syntactic complexity dimensions. *Neuroimage*, 51(4), 1285–1293. doi: 10.1016/j.neuroimage.2010.03.034

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., … Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the USA*, 105(46), 18035–18040. doi: 10.1073/pnas.0805234105

Service, E., Helenius, P., Maury, S., & Salmelin, R. (2007). Localization of syntactic and semantic brain responses using magnetoencephalography. *Journal of Cognitive Neuroscience*, 19(7), 1193–1205. doi: 10.1162/jocn.2007.19.7.1193

Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: an FMRI study using word-category ambiguity. *Cerebral Cortex*, 19(7), 1493–1503. doi: 10.1093/cercor/bhn187

Skeide, M. A., Brauer, J., & Friederici, A. D. (2014). Syntax gradually segregates from semantics in the developing brain. *Neuroimage*, 100, 106–111.

Skeide, M. A., & Friederici, A. D. (2016a). The ontogeny of the cortical language network. *Nature Reviews Neuroscience*, 17(5), 323–332.

Skeide, M. A., Brauer, J., & Friederici, A. D. (2016b). Brain functional and structural predictors of language performance. *Cerebral Cortex*, 26(5), 2127–2139. doi: 10.1093/cercor/bhv042

Tsapkini, K., Frangakis, C. E., & Hillis, A. E. (2011). The function of the left anterior temporal pole: evidence from acute stroke & infarct volume. *Brain*, 134(10), 3094–3105. doi: 10.1093/brain/awr050

Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., … Stamatakis, E. A. (2011). Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. *Brain*, 134(2), 415–431. doi: 10.1093/brain/awq369

Uddin, L. Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D.A., Greicius, M. D., & Menon, V. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional & structural connectivity. *Cerebral Cortex*, 20(11), 2636–2646. doi: 10.1093/cercor/bhq011

Upadhyay, J., Silver, A., Knaus, T. A., Lindgren, K. A., Ducros, M., Kim, D. S., & Tager-Flusberg, H. (2008). Effective and structural connectivity in the human auditory cortex. *The Journal of Neuroscience*, 28(13), 3341–3349. doi: 10.1523/JNEUROSCI.4434-07.2008

Visser, M., Jefferies, E., & Lambon Ralph M. E. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, 22(6), 1083–1094. doi: 10.1162/jocn.2009.21309

Weiller, C., Musso, M., Rijntjes, M., & Saur, D. (2009). Please don't underestimate the ventral pathway in language. *Trends in Cognitive Sciences*, 13(9), 369–370. doi: 10.1016/j.tics.2009.06.007

Weiller, C., Bormann, T., Saur, D., Musso, M., & Rijntjes, M. (2011). How the ventral pathway got lost: and what its recovery might mean. *Brain and Language*, 118(1-2), 29–39. doi: 10.1016/j.bandl.2011.01.005

Wernicke, C. (1874). *Der aphasische Symptomenkomplex. Eine psychologische Studie auf anatomischer Basis*. Breslau: Cohn & Weigert.

Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., … Gorno-Tempini, M. L. (2011). Syntactic processing depends on dorsal language tracts. *Neuron*, 72(2), 397–403. doi: 10.1016/j.neuron.2011.09.014

Xiang, H. D., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2010). Topographical functional connectivity pattern in the perisylvian language networks. *Cerebral Cortex*, 20(3), 549–560. doi: 10.1093/cercor/bhp119

Zilles, K., Bacha-Trams, M., Palomero-Gallagher, N., Amunts, K., & Friederici, A. D. (2014). Common molecular basis of the sentence comprehension network revealed by neurotransmitter receptor fingerprints. *Cortex* 63, 79–89. doi: 10.1016/j.cortex.2014.07.007

# Part III Acquisition

# 20 Overview

VIRGINIA VALIAN
Hunter College and the Graduate Center, City University of New York

The key questions in language acquisition concern what linguistic concepts, if any, might be built into the child's language-learning mechanism, what the role of linguistic experience is, and what procedures the child uses to develop her language(s). How, in short, does the child use the language she hears—which comes to her ears in the form of an undifferentiated stream—to arrive at the abstract knowledge of language that adults and older children have? Language acquisition involves mastery of sounds and phonemes (or their handshape equivalents), morphemes, words, syntax, semantics, and pragmatics. Each of those aspects of language requires abstract mental representations. The consonants, vowels, and tones of language are abstract; the rhythm, intonation, and stress are abstract; words are abstract; syntax is abstract; and semantics is abstract. The miracle of language acquisition and use is that learners turn mere sounds into the abstract syntactic structures used to recover meaning.

At birth, the neonate has already analyzed the basic prosody of her language and has analyzed specific vowel sounds. Within the first few days of life the child distinguishes the rhythm, intonation, and stress of the language heard in the womb from languages with a different prosodic pattern—but not from languages with a similar prosodic pattern. For example, four-day-old infants exposed to French *in utero* could distinguish French from Russian, languages with different prosodic patterns, while infants exposed to neither language *in utero* did not distinguish them (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, & Amiel-Tison, 1988). Thus, experience *in utero* allows neonates to extract the pattern of their language. That infants do not distinguish the pattern of their language from another with a similar pattern is evident from a study with two-month-old English-hearing infants. They can distinguish English from Japanese, but not from Dutch. That is because the prosody of English at the word level is much different from the prosody of Japanese but similar to the prosody of Dutch (Christophe & Morton, 1998). Infants

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

who heard English in the womb also do not distinguish French from Japanese, despite the prosodic differences between the two languages: as English-hearers, they have extracted a prosodic pattern only for languages like English.

Newborn infants show that they have not only acquired prosodic patterns in the womb, but particular features of their language as well. They respond differentially to vowel sounds from their language and similar vowel sounds from another language. Neonates who heard either English or Swedish in the womb differentiated between English and Swedish vowels within a few days after birth (Moon, Lagercrantz, & Kuhl, 2013). Further, the number of hours post-birth was not related to how sharp their differentiation was. It was the *in utero* learning that mattered. The infants heard variants of the vowels and not just a single example. They appeared to treat the variants from their own language as examples of the same vowel, but to treat the variants from the other language as different vowels. Not only were children learning in the womb, but they were creating classes of vowel sounds that they treated as equivalent. The creation of equivalence classes—categories in which all members behave similarly in one or more ways—is crucial for language acquisition.

Studies like these demonstrate that the learning that occurs in the womb is spontaneous and already abstract. Learning takes place even though it occurs in social isolation, with no meanings attached to the sounds, and with no reinforcement given for the development of categories. Most of language acquisition, including acquisition of the sound structure of one's language, of course takes place outside the womb. The purpose of these examples is to show that the acquisition of linguistic patterns occurs even earlier than one might have thought.

The protean nature of language acquisition is amply demonstrated in the chapters in this section. Language acquisition occurs at all major life stages—in infancy, childhood, and adulthood. People are typically exposed to their first—and in some cases, only—language in infancy and early childhood, but "late" acquisition is common for deaf individuals in hearing-only households. Most existing research is on monolingual English hearer-speakers, reflecting the fact that there are more researchers who are native English speakers than there are researchers speaking any other language or set of languages, but, fortunately, this has begun to change. Deen provides examples of morphological development in a number of different languages. According to Ethnologue (Lewis, Simons, & Fennig, 2016, https://www.ethnologue.com/guides/how-many-languages), 7097 languages were documented as of 2016, so a concentration on a mere handful of the world's languages would be misguided and benighted. Since the field is increasingly benefiting from studies of acquisition of a wide range of languages, both spoken and signed, we can expect a corresponding broad range of insights.

What counts as a language is important at two levels in language acquisition. One level is in distinguishing one language from another (and its related manifestation of noticing that two languages are mutually intelligible). American and British English seem intuitively to be dialects of the same language, even though there are some lexical differences (*elevator* versus *lift*) and some minor syntactic differences (British English allows "I might do" as an answer to "Are you going to the movies tonight?" and some dialects allow "I gave it her" in addition to "I gave her it"). But many Chinese people speak two "dialects" that are not mutually intelligible, such as Mandarin and the language of their local community. If mutual unintelligibility is the criterion for distinguishing languages, those individuals are actually bilinguals.

The other level at which what counts as a language matters is in distinguishing language from communication systems. Kegl makes that distinction in order to separate some gestural communication systems that deaf children with their hearing relatives might develop and use at home, from natural language. She suggests that the type and amount of input deaf children receive are related to the type of system that the child develops.

Only intensive cross-linguistic research can provide answers about what is universal in language acquisition and what is language specific. The wealth of data presented in the chapters in this section demonstrate that intensive study, rather than a cursory look, is necessary, because a great deal of data are required to confirm or disconfirm hypotheses. Many of the chapters describe results that are inconsistent or difficult to explain fully via any existing models. Best, for example, compares three models of how children learn that some sound contrasts are meaningful in their language, while others are not. The models are not necessarily mutually exclusive, and none of them fully account for children's behavior. Explanatory theories require a great deal of information. With small amounts of data, it is possible to prematurely accept incorrect theories.

As with the concentration on English, research on monolinguals is more common than research on bilinguals, even though some researchers suggest that bilingualism is at least as common as and perhaps more common than monolingualism in the world (Grosjean, 2010). More than 50% of citizens in the European Union (EU) can carry on a conversation in more than one language, and in some EU countries, more than 90% of the inhabitants speak more than one language (Luxembourg, Latvia, Slovenia, Lithuania, Malta, the Netherlands, and Sweden; European Commission (2012), http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf). In the United States, according to census data from 2009-2013, 21% of the population over age 5 speaks a language other than English at home (United States Census Bureau, 2015, https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html). Of those, 58% say they speak English less than very well. Knowing more than one language is common. For that reason alone, we need to understand how it occurs. Again, without intensive study of bilinguals as well as monolinguals, we will be unable to determine what is universal about language acquisition and what is specific to learning a single language.

There are many ways of being bilingual. Some children grow up being exposed, roughly equally, to more than one language. But that pattern is only one of many patterns. Some children instead spend their first years as a monolingual, speaking their single language at home; they are exposed to a second language only when they start attending school. If the language at school is the language of the community, or the majority language, the child may get increasingly less exposure to her first language, and end up knowing it less well than monolinguals of that language do. So-called *heritage* speakers are often in a position where there are two home languages, one of which is the majority language. The majority language can become more and more dominant, and the learner may either never acquire or lose information that full native speakers have. In French, Spanish, and many other languages, for example, nouns have a gender; adjectives and articles must agree in gender with the nouns they are in construction with. Heritage language learners make more errors and have less facility with gender agreement than native speakers do in some tasks (e.g., Montrul, Foote, & Perpiñán, 2008; Montrul, Davidson, De La Fuente, & Foote, 2014; Polinsky, 2008). Yet other children become exposed to another language later in life, sometimes through immigration, and retain their first language.

Papers in this section address whether acquisition differs, depending on whether it is a first or second language, or on whether it is one of two languages being simultaneously acquired, or on whether it is acquired late. A second language can be acquired either simultaneously with a first language, or in later childhood or adulthood. How similar are simultaneous and successive types of acquisition? What effect does already knowing a language, for example, have on acquiring a second (or third) language? Several authors in this section address bilingual, second, and late language acquisition: Kegl, Klein and Martohardjono, Meisel, and Byers-Heinlein and Lew-Williams. One issue with acquiring more than one language, or acquiring a language late, is the quantity and quality of language that the learner is exposed to—the input. Children who are exposed to two languages from birth, for example, effectively receive half as much language input as children who are exposed to a single language. For children who are exposed to language late, as is the case for many deaf children, the late and partial exposure may lead to non-optimal acquisition. Although researchers do not find strong evidence for a critical period, except perhaps for acquiring a native-like accent, it may be necessary to be exposed to *some* natural language early in life. Kegl addresses these and other issues.

Klein and Martohardjono distinguish between bilingual acquisition and second language acquisition. If a child is exposed to two languages before the age of three, she is bilingual. If one language is not present until after age three, the child is acquiring that language as a second language. If an individual is not exposed to the second language until after puberty, that person is considered an adult second language learner, rather than a child second language learner, but Klein and Martohardjono note that different researchers have different time periods.

Acquisition post-birth almost always occurs in a social context. Roseberry Lytle and Kuhl hypothesize that natural language learning requires social interaction between the child and the people around her. Chinese sounds, for example, are learned by nine-month-old English-speaking infants when they interact with a live speaker, but not when they see a video of a speaker or hear a recording of a speaker. The same superiority of social interaction holds when word learning and syntax learning are examined. A video chat is as conducive to learning as a live chat with 24 to 30 month olds, showing that the speaker does not have to be physically present, but does have to be responsive. Roseberry Lytle and Kuhl suggest that social interaction acts at many levels, by directing the child's attention to speech, increasing the amount of information the child receives, and developing and maintaining the child's motivation to participate meaningfully in the interaction. Mutual eye gaze is one social cue children use in word learning. Byers-Heinlein and Lew-Williams report that bilinguals more effectively use eye gaze than monolinguals do in detecting where a toy has been hidden. Bilinguals may be even more sensitive to social cues than monolinguals. Clear evidence of the value of social interaction comes from studies of deaf children of hearing relatives who enter a school for the deaf, as Kegl describes. By having language partners who also use only a visual-manual system, individuals develop a much more extensive communication system.

As Roseberry Lytle and Kuhl point out, the results with social interaction might seem surprising, since infants do learn patterns from strings of syllables that are presented to them in the laboratory, with no social interaction. And, as we have already noted, the fetus learns the prosody of ambient language and creates equivalence classes of sounds that are specific to the ambient language. Roseberry Lytle and Kuhl point out that babies exist in a very rich linguistic world. The kind of learning that is required when a baby is exposed to a full language from multiple speakers may be very different from the kind of learning that is required in a laboratory setting or in the womb.

How the infant processes sounds brings up the question of whether the child brings domain-specific or domain-general abilities to the process of language acquisition. Best addresses the difficulty of answering this question in practice, even though the two are easy to separate in principle. If the child has domain-specific abilities, for example, she may take speech sounds that are on a continuum and impose a categorical structure on them. The difference between *ba* and *pa* seems categorical in perception, even though they are actually on a continuum. We accept a range of sounds as *ba* and then suddenly shift to perceiving a range of sounds as *pa*. A nice demonstration of this can be found at The Virtual Linguistics Campus (2015, Jun 25).

As voice onset time (the time it takes the vocal cords to start vibrating after air flow is released following an initial blockage due to pronunciation of the start of the consonant) decreases we continue to hear *ba* as *ba*, but somewhere between 30 ms and 0 ms we hear the sound as *pa*. Even one-month-old infants show this phenomenon (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). As Best points out, such phenomena were interpreted as demonstrating that "speech is special" and unique to humans. But it turns out that humans treat non-speech sounds as well as speech sounds categorically. It also turns out that chinchillas—after much training with the end points of the continuum—can also make categorical distinctions of speech sounds (Kuhl & Miller, 1975). Perhaps, then, children are using general perceptual mechanisms when hearing speech, mechanisms that are shared with other animals and that are utilized for a range of stimuli. Or, categorical perception may, in its origin, be specific to speech and recruited to handle other stimuli. Chinchillas may be exploiting a different mechanism than humans do, as the extensive training suggests.

Along similar lines, rats can use changes in item frequency in a sequence as an ordering cue (after training) in a way similar to infants (Toro, Nespor, & Gervain, 2016), suggesting another domain-general ability. They can also use pitch alternations to distinguish different sequences (de la Mora, Nespor, & Toro, 2013). When frequency is pitted against prosody, human infants (Gervain & Werker, 2013), but not rats (Toro *et al*., 2016), use prosody as a cue, suggesting that the ability to use some cues for higher-order processing, such as drawing inferences about syntactic structure, may be unique to humans.

Best considers at length how the infant's experience affects her ability—and loss of ability—to make certain speech contrasts, and how the child uses speech perception to begin recognizing words. Best compares a number of models for each process. The different models are unusually specific and detailed compared to models for the development of syntax, semantics, and other aspects of grammar. In the case of reorganization of the sound system, one possibility is that the child tunes her system to the sounds that are used in her language, thus reducing—but not forever losing—the ability to distinguish sounds used only in other languages. Another possibility is that during an early critical period the child's brain becomes committed to the sounds of the system she is acquiring. And a third possibility, which Best favors, is similar to the idea that the child reorganizes her sound system, but also predicts that the child will be able to make non-native contrasts if they can be assimilated to the contrasts in her native language.

Best also compares theories that address how the child solves the problem that people pronounce the same word in different ways. Female speakers sound different from male speakers, old people sound different from young people, and so on. How do children come to distinguish sound differences that do not signal word-level differences from sound differences that do signal word-level differences? The word *tree*, for example, can be pronounced in different ways. A canonical American and British pronunciation can be found here: http://dictionary.cambridge.org/us/pronunciation/english/tree. Most English speakers may not even detect a difference between them. But they do hear the difference between *tree* and *three*, even though some speakers, especially non-natives, pronounce them the same.

However complicated early speech perception is for monolinguals, it is even more complicated for bilinguals. Yet, as Byers-Heinlein and Lew-Williams describe, bilingual babies appear to make most of the same distinctions that monolinguals do. At the same time, bilingual babies may have a different learning trajectory for exactly the case we've been describing, where sound differences may or may not signal word differences. Spanish-Catalan babies hear vowels in Catalan that mark word differences, and distinguish them at age 4 months and at age 12 months, but not always at age 8 months. Monolingual Catalan babies distinguish the vowels throughout that period. One conjecture is that bilingual babies go through a period when they realize that, across their two languages, a vowel difference need not signal a word difference. The words *pilota* and *pelota* both mean "ball." That similarity in word meaning coupled with a difference in vowel character may lead infants to temporarily ignore such vowel differences.

The way that children learn labels for words also differs in some respects between mono- and bilinguals. Although both groups may take as their first hypothesis about a new noun that it refers to a whole object rather than a part of an object, Byers-Heinlein and Lew-Williams note that the groups differ with respect to mutual exclusivity. Monolinguals assume, when they hear a new noun, that the noun is not a synonym for a noun they already know, but refers to a different object. For bilinguals, however, who have the experience of learning different words that mean the same thing across their two languages, mutual exclusivity does not hold. Bilinguals also differ from monolinguals in knowing fewer words in each of their languages than a comparable monolingual does. That is presumably one consequence of having less input in either language than a monolingual peer does. If the total number of different words the child knows across her two languages is tabulated, her total vocabulary is comparable to a monolingual's. That too suggests that the child's vocabulary is tightly linked to the linguistic input she receives. For vocabulary, it could not be otherwise.

The only way a child can learn a word is by hearing it. Thus, vocabulary size in one language is not correlated with vocabulary size in another language. Word learning depends on exposure. The richer the input at 18 months, the greater the vocabulary and processing efficiency at 24 months. As Byers-Heinlein and Lew-Williams note, processing efficiency, like vocabulary, is not correlated across a child's two languages. The child may be much more efficient in processing one of her two languages.

Once the child has been exposed to a word, whether it is a noun or a verb, what strategies does she use to learn what the word means? Levine, Strother-Garcia, Hirsh-Pasek, and Golinkoff suggest that only a hybrid model can explain how the child acquires word meanings. That hybrid makes three main assumptions. The first is that the child will use a variety of cues in learning a word. We have already seen that social interaction, and the myriad cues it provides, aids word learning. Perceptual cues are another aid, as are linguistic cues. The second assumption is that the child may use cues to different degrees as she develops. Perceptual cues may be paramount early in acquisition, but less important later in acquisition. The third assumption is that the child has internal biases that she brings to the word-learning situation, such as the bias that a noun refers to a whole object.

Word learning is not an all-or-nothing phenomenon. Levine, Strother-Garcia, Hirsh-Pasek, and Golinkoff note that it takes several different experiences with a word, in different contexts, before a child can determine its full meaning. Word learning, unlike other forms of language acquisition, continues across a lifetime. English-speaking adults know many thousands of words, while one year olds know only a few. Nouns tend to be produced more frequently than verbs cross-linguistically, though children in some languages produce verbs more often than children in other languages do. The "noun bias," Levine, Strother-Garcia, Hirsh-Pasek, and Golinkoff suggest, is due to the greater concreteness of nouns compared to verbs. Nouns that refer to objects that have a consistent shape and are easy to perceive, nouns that are concrete and imageable, are nouns that are easier to acquire.

Although, as we have said, vocabulary is dependent on exposure, that statement is only true if we are talking about word roots. We can distinguish between word roots and the morphemes, or small units of meaning, that can be combined with words. There are two productive processes that allow a child to produce new words even if she has never heard them. A child can be productive by adding *–s* to create plural nouns or to create third person singular present tense verbs. Such processes are morphological: they deal with the structure of words. Deen describes the two types of morphology: inflectional, as in the example just presented, and derivational. Derivational morphology takes prefixes and suffixes to create a new word. When a child understands how morphological processes work, and has the relevant morphemes, she can create new words based on roots and affixes she already knows. A word like *antidisestablishmentarianism* has *establish* as its root, with the affixes *anti-*, *dis-*, *ment-*, *arian-*, and *ism* added to the front and back ends.

Inflectional morphology is at the border between morphology and syntax. The distinctions that are made in inflectional morphology, such as tense, person, grammatical gender, and so on, are distinctions that are relevant to syntax. When we speak of subject-verb agreement, for example, we are relating two word forms. In English, with the verb *be*, subject-verb agreement is more visible than it is with standard main verbs. Only the form *I* can be used with *am*; in that sense the subject and the verb agree. Similarly, as already discussed, in languages like Spanish, the article and the noun agree in gender. In English, inflectional morphology is rather limited compared to other languages.

Regardless of language, Deen notes that children acquire inflectional morphology before age four. In principle, the moment the child understands that *–ed* is how English represents the past tense and can be added to any verb (except irregulars), she should uniformly use the past tense when the occasion demands. Yet two year olds are inconsistent in their use of past tense in English. When they use it, they use it correctly, but they do not always use it when they should. The reasons for omission are not clear, particularly since, as Deen notes, children in languages with rich morphology seem to master morphology very early, although omission is common in the acquisition of every language. Although errors of omission are common, errors of commission are rare. For example, children very seldom use third person singular present tense *–s* with first person verbs. In agglutinative languages like Turkish and Swahili, where each affix encodes a different grammatical property (like tense, person, and gender), children put the affixes in the correct order. Children's errors of overregularization, such as saying *foots* instead of *feet*, or *runned* instead of *ran*, are particularly good evidence that the child has an internal rule. Those are forms that the child has never heard, so the child's production shows that she is overapplying a rule.

The acquisition of morphosyntax by children who are bilinguals is described by Meisel. A major question is whether the child's two languages develop independently of each other or whether one of them influences the other. Meisel provides evidence that, even at the beginning of acquisition, the child does not have a single merged system but two systems. If, for example, the two year old's two languages differ in the word order they use for subjects, verbs, and objects, the child properly obeys the word order of each. Similarly, in cases where one language requires the verb to be in the second position but the second one does not, children do not appear confused. There is, however, some cross-linguistic interaction, though it does not occur for all children. Meisel suggests that children's use of properties from one language in a second language reflect processing rather than grammatical differences. The interaction effects appear to be quantitative. For example, English-Italian bilingual children may use subjects more in Italian than is standard for monolinguals (Serratrice, Sorace, & Paoli, 2004), though it is by no means the rule (Valian, 2016).

The acquisition of morphosyntax in a second language, as described by Klein and Martohardjono, shows both similarities to and differences from acquisition of morphosyntax in a child's first language. Among the differences is the fact that errors of commission are more common in children's acquisition of a second language. One example is the use of *be* where it does not belong, as in "he is go." Child learners of a second language, such as learners of German who had Chinese as the native language, appear to have no difficulty acquiring tense markers even though their native language does not express tense overtly. In other instances, there may be evidence of transfer from the child's first language to their second language. Unlike child learners, adult learners of a second language may persist in their morphosyntactic errors, especially in production.

Klein and Martohardjono review different models of those persistent errors. One set of models attributes the adult's errors to a lack of knowledge. In one variant, the problem is at the level of morphosyntax, in a second it is at the level of phonology or prosody, in a third it is at the level of the lexicon. In all the cases, aspects of the learner's first language are being transferred to the second language. Depending on the model, the learner is or is not hypothesized to be able to recover from the erroneous transfer from the first language. In models that propose that the learner continues to have full access to universal language principles, as well as transferring properties from the first language to the second language, the learner can "recover" from the errors.

Another set of models attributes the adult's errors to processing or parsing difficulties, rather than to lack of linguistic knowledge. For example, the fact that learners' performance is equivalent to native speakers' in grammaticality judgment tasks, but worse in on-line tasks, suggests that processing difficulties may be responsible when poor performance is observed. The second language learner may use different parsing strategies than the first language learner uses, or may use the same strategies but have difficulty employing them due to processing difficulties.

The theme of needing to distinguish between errors that reflect differences in grammar from errors that reflect differences in processing recurs throughout this section. Just as it is relevant in morphosyntax, it is relevant in syntax. Hyams and Orfitelli consider several syntactic error phenomena, some from two year olds and some from older children. In each case they review the possibility that the children's errors are due to deficits in grammatical knowledge (competence deficits) or to processing (performance deficits). For the cases they review, they conclude that the children have deficient grammars, even if they also have processing problems. As one example, two year olds in languages like English and German produce what are sometimes called root infinitives. In English, children produce verbs without inflections; in German, they produce the infinitival form instead of a tensed form. These errors are most common in languages, like English and German, that require that sentences have subjects. Hyams and Orfitelli describe competence- and performance-deficit accounts, as well as a hybrid account. As another example, Hyams and Orfitelli describe explanations of children's failure, in languages like English and German, to produce subjects for sentences as often as they should. Here, too, there are competence- and performance-deficit models, and, here, too, Hyams and Orfitelli conclude that children have a competence deficit. Four and five year olds seem to misunderstand sentences like "Bert hugged Ernie before playing the piano," in some cases taking the player of the piano to be anyone at all, even someone not mentioned in the sentence. Here, there are several explanations, all of them relating the child's non-adult interpretation of such sentences to immature structural analyses or to extra-syntactic factors.

Learners must not only acquire knowledge of their languages, but they have to put that knowledge to use. That in turn requires skills in planning, integrating, and remembering. McKee, McDaniel, and Garrett suggest that limited production or comprehension on the child's part cannot be taken to imply imperfect knowledge on the child's part, a point that Klein and Martohardjono also make when considering children's acquisition of a second language. For adults, speaking and listening are such highly practiced skills that they seem effortless. As speech errors attest, even this highly practiced behavior occasionally goes awry, and when it does, it goes awry in principled ways. Consider speech errors in which the speaker substitutes an intruded word instead of the target, as when a child says *cookie* instead of *candy*. When intrusions occur, both children and adults substitute a word of the same syntactic category. Children also make errors that adults do not, such as by omitting words like Determiners (e.g., *a* and *the*) and inflectional suffixes (e.g., third person singular –*s*). Comprehension experiments suggest that the child's grammar represents such elements but the production system has difficulty retrieving them and fitting them into a prosodic pattern. McKee, McDaniel, and Garrett propose that the architecture of the child's and adult's production system is the same. The difference is that the child is unable to exploit all the resources of that system.

No part of acquisition occurs in isolation from any other part. Once one contemplates the sheer range of knowledge and abilities that the child or adult learner must bring to bear in typical language acquisition, one is stunned that learners succeed so well so quickly. How in the world do they do it? The chapters in this section explore several answers to these questions, bringing to bear a wealth of data. The reader will be awed by children's accomplishments and by researchers' ingenuity in investigating those accomplishments.

# REFERENCES


Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, 1(2), 215–219. doi: 10.1111/1467-7687.00033

de la Mora, D., Nespor, M., & Toro, J. M. (2013). Do humans and non-human animals share the grouping principles of the Iambic–Trochaic Law? *Attention, Perception and Psychophysics*, 75, 92–100. doi: 10.3758/s13414-012-0371-3

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171(3968), 303–306. doi: 10.1126/science.171.3968.303

European Commission (2012). *Europeans and their languages: Special Eurobarometer 386 report*. Directorate-General for Communication. Retrieved on 26 October 2016 from http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf.

Gervain, J., & Werker, J. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, 4, 1490. doi: 10.1038/ncomms2430

Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, MA: Harvard University Press. doi: 10.4159/9780674056459

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, 190(4209), 69–72. doi: 10.1126/science.1166301

Lewis, M. P., Simons, G. F., &. Fennig, C. D. (Eds.). (2016). Ethnologue: Languages of the world, 19th ed. Dallas, TX: SIL International. Retrieved on 26 October 2016 from https://www.ethnologue.com/guides/how-many-languages.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143–178. doi: 10.1016/0010-0277(88)90035-2

Moon, C., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: a two-country study. *Acta Paediatrica*, 102(2), 156–160.doi: 10.1111/apa.12098

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58, 3–53. doi: 10.1111/j.1467-9922.2008.00449.x

Montrul, S., Davidson, J., De La Fuente, I., & Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 17(01), 118–138. doi: 10.1017/S1366728913000114

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal*, 6(1), 40–71.

Serratrice, L., Sorace, A., & Paoli, S. (2004). Cross-linguistic influence at the syntax–pragmatics interface: Subjects and objects in English–Italian bilingual and monolingual acquisition. *Bilingualism: Language and cognition*, 7(03), 183–205. doi: 10.1017/S1366728904001610

Toro, J. M., Nespor, M., & Gervain, J. (2016). Frequency-based organization of speech sequences in a nonhuman animal. *Cognition*, 146, 1–7. doi: 10.1016/j.cognition.2015.09.006

United States Census Bureau. (2015). Detailed languages spoken at home and ability to speak English for the population 5 years and over: 2009-2013. Retrieved on 26 October 2016 from http://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html.

Valian, V. (2016). Null subjects. In J. Lidz, W. Snyder, & J. Pater (Eds.), *Oxford handbook of developmental linguistics* (Ch 17). Oxford, UK: Oxford University Press.

[The Virtual Linguistic Campus] (2015, June 25). Linguistic Micro-Lectures: Categorical Perception [Video File]. Retrieved from https://youtu.be/4V5pQyKsgg4

# 21 Speech Perception in Infants: Propagating the Effects of Language Experience

CATHERINE T. BEST

Western Sydney University and Haskins Laboratories

## Introduction

During their first year and a half, children make remarkable progress toward becoming native perceivers and beginning speakers of the specific languages their caregivers direct to them. In the first postnatal year, infants show progressive perceptual attunement to multiple aspects of native speech. This attunement establishes the foundation that allows normal children to achieve a watershed in lexical development by18–20 months on average. By this time most children have moved well beyond their first 25 or so spoken words, which they had acquired relatively slowly as holistic forms, and they have built up an expressive vocabulary of 50–100+ words that displays various signs of (childlike) phonological rules and regularizations. This accomplishment reflects their discovery of certain basic phonological principles of native words (see Peperkamp, 2003), specifically, that consonant and vowel distinctions can serve as the basis for lexical contrast; that the same consonant or vowel can recur across different words, for example, they start to recognize and play with rhyming and alliteration; and that the pronunciation of these segmental subcomponents in different contexts is governed by phonological rules. These insights into the internal structure of native words and the phonological rules that apply across words support more rapid vocabulary growth. They also herald other linguistic accomplishments. At virtually the same time, children also begin to produce the combinatorial morphological and syntactic features of their language that will extend their early single-word utterances into phrases of increasing length, which they will use to build ever more sophisticated propositional utterances and narratives over their preschool and early school years.

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

The perceptual advances in year one that presage those linguistic advances of year two and beyond are initially grounded in inborn perceptual sensitivities and attentional biases. Infants begin postnatal life equipped with a small but exquisitely adapted repertoire of cries, comfort-state sounds, facial expressions, and other behaviors that are produced more often toward people than toward objects or events. These tendencies draw young infants into frequent, speech-infused interactions with caregivers and family members. Within such contexts, their perception of native speech is accentuated both by perceptual preferences for human speech and faces over other sounds and sights, and by wide-ranging sensitivities to many critical properties of speech, including consonant and vowel phonetic distinctions, as well as rhythmic and intonation patterns, that is, the prosodic properties of connected speech. Importantly, in addition, infants evince remarkable adaptive capabilities, which together with exposure to the languages spoken in their homes, rapidly "tune up" their detection of phonetic distinctions and phonotactic sequences that are frequent in native speech. All this provides a strong foundation for acquiring productive and receptive vocabulary and linguistic characteristics of the native language early on.

An extensive body of empirical findings on speech perception across the first year indicates that by six months, infants are already tuning in perceptually to the prosodic properties of their to-be-native language, as well as to some of its specific vowels. Next, between 6–12 months, evidence of attunement to the consonant contrasts and to the frequency of occurrence of phonotactic patterns (permissible sequences of consonants and/or vowel) in that language begins to emerge. So does the ability to perceptually segment from connected speech those word-forms the child either has already heard frequently in their environment (*mommy, daddy*, their own name) or has just been familiarized with in the laboratory. Also during the second half-year and beyond, infants show progressive improvement in their ability to accommodate to variations in talkers, emotions, age and gender, and foreign or non-native regional accents when discriminating phonetic contrasts or segmenting words from continuous speech.

This chapter will review key findings from that literature, focusing in particular on the developmental changes in speech perception that result from language experience across the first year and up through the middle of the second year. We conclude by considering the implications of infants' perceptual attunement to native speech, and their increasing skill in handling systematic phonetic variations in the target stimuli, for their progress in lexical, phonological, morphological, and syntactic development during the second year and beyond.

We begin by placing the infant speech perception work in its historical and theoretical context, starting with an overview of the trajectory of three main theoretical issues that have driven empirical research on the topic since it began nearly five decades ago (see also Gervain & Mehler, 2010). Next, we will summarize and compare the key claims of several influential theoretical models of experiential effects on infant speech perception. We will go on to review the key research findings, and end by discussing their broader implications and raising several important remaining open questions.

## Brief history of fundamental theoretical issues in infant speech perception research

Experimental studies of infant speech perception began to appear in the literature in the early 1970s (e.g., Eimas, Siqueland, Jusczyk, & Vigorito, 1971). Research on the topic since then has moved through three global phases that have differed in their main theoretical foci, although these theoretical issues have overlapped and continued across phases (see also Jusczyk & Luce, 2002). The first phase (1970s-early 1990s) focused primarily on the question of whether infants possess a *biological specialization* for speech perception, as evidenced particularly but not only by their discrimination of consonant and vowel contrasts. On one side of this still-unresolved debate stands the premise that humans alone are biologically specialized for perceiving and producing speech (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985, 1989; see also Studdert-Kennedy & Goldstein, 2003, re: the *particulate principle* as the specialized mechanism), as an integral part of a broader human specialization for language (e.g., Chomsky, 1975) that is instantiated in a devoted neural system in the human brain distributed within the left hemisphere (e.g., Lenneberg, 1967). This language system is posited to employ specialized speech-specific mechanisms, such as innate neural links between speech motor commands and the mechanisms for perceiving the corresponding speech contrasts. These human-specialized mechanisms are claimed to differ qualitatively from the domain-general cognitive/perceptual processes employed in perception of non-linguistic stimuli, which we share with other animals.

During this first research phase, the proponents of biological specialization sought to identify innate, universal, speech-specific perceptual abilities in young infants. Candidate behaviors included:

a. *categorical perception* of synthesized continua or naturally produced (though sometimes manipulated) tokens of minimal segmental distinctions, that is, pairs of consonants or vowels differentiated by a single phonetic feature (e.g., Eimas, 1975; Eimas *et al*., 1971; Morse, 1972);
b. adult-like sensitivity to multiple and overlapping signal properties for those distinctions, such as *trading relations* between critical spectral and temporal properties of consonant distinctions (e.g., Eimas, 1985; Miller & Eimas, 1983; see foundational work with adults and older children: Best, Morrongiello, & Robson, 1981; Morrongiello, Robson, Best, & Clifton, 1984), *perceptual compensation for coarticulation* between sequences of consonants (e.g., of stop consonants preceded by/r/versus/l/) (Fowler, Best, & McRoberts, 1990; see also Fowler, 1984; Mann, 1980, 1986), and *perceptual equivalence* (constancy) across differences in phonetic context or talkers (Jusczyk & Derrah, 1987; Kuhl, 1983);
c. *cross-modal integration* of heard and seen speech by young infants, including looking preferences for the one of two talking faces that phonetically matches a

simultaneously-heard vowel (Kuhl & Meltzoff, 1982) even if that vowel is non-native (not previously experienced) (Walton & Bower, 1993), and the *McGurk effect*, in which a phonetic conflict between synchronized heard and seen consonants is perceived as a third consonant (e.g., Burnham & Dodson, 2004; Rosenblum, Schmuckler, & Johnson, 1997; see originating study: McGurk & MacDonald, 1976);

d. a *left hemisphere advantage* in infant speech perception (e.g., Best, Hoffman, & Glanville, 1982; Glanville, Best, & Levenson, 1977; MacKain, Studdert-Kennedy, Spieker, & Stern, 1983; Molfese, Freeman, & Palermo,1975; Molfese & Molfese, 1985; see also Best, 1988), including a left hemisphere bias in preferential looking to the talking face that matches a simultaneously heard disyllable (MacKain, Studdert-Kennedy, Spieker, & Stern, 1983).

On the other side of this debate, counter-evidence has been sought to support the view that infant speech perception phenomena can be explained by domain-general cognitive/perceptual mechanisms that are neither unique to speech stimuli (e.g., categorical perception of nonspeech stimuli contrasting in *tone-onset-time:* Jusczyk, Pisoni, Walley, & Murray, 1980; in *rise-time*: Jusczyk, Rosner, Cutting, Foard, & Smith, 1977; and in *rapid spectrum change*: Jusczyk, Pisoni, Reed, Fernald, & Myers, 1983), nor exclusive to humans (i.e., evidence that other primates or even birds show *categorical perception* and *perceptual compensation for coarticulation* and *perceptual constancy* for speech contrasts: e.g., Kuhl, 1981; Kuhl & Miller, 1975; Kuhl & Padden, 1983; Kluender, Diehl, & Killeen, 1987; Kluender & Lotto, 1994; Lotto, Kluender, & Holt, 1997).

However, deciding unequivocally between human speech-specialized versus domain-general and cross-species accounts is virtually impossible (see, e.g., Diehl & Kluender, 1989, in context of the commentaries by Fowler, 1989; Remez, 1989; Studdert-Kennedy, 1989). For example, parallel findings on speech and nonspeech stimuli can be interpreted either as evidence that speech perception depends on domain-general processes, or that specialized speech processes were triggered ("tricked") by the uniquely speech-critical properties of the artificially created non-speech stimuli (i.e., which represent no naturally occurring sound-producing events). Alternatively, it may be that the perceptual phenomenon under study is not speech-specialized but this does not logically preclude the existence of other potentially speech-specialized perceptual processes. Conversely, a *lack* of parallel between perception of speech and the correspondingly devised nonspeech stimuli could either be taken to reflect that perception of speech *is* specialized, or that those particular nonspeech stimuli lacked the critical acoustic properties needed to elicit the domain-general perceptual pattern under investigation. In short, the same data can be claimed to support or at least to admit room for either side of the debate, and there is ample opportunity for hypothesizing other untested processes or stimulus properties that might provide the lynchpin between domain-general auditory mechanisms and biologically specialized speech perception mechanisms. The debate is unlikely to be settled by any set of experiments, as the assumptions of both sides lie in metaphysics rather than observables (see Kuhn, 1962; Popper, 1957).

In any case, the biological specialization debate has since faded as the primary rationale for most studies on infant speech perception. It has been subsumed by the newer focus of the second, overlapping research phase (roughly the 1980s through the 2000s, but still ongoing). Regardless of whether infants' initial responses to speech arise from specialized or general perceptual abilities, subsequent research has made it increasingly clear that language experience imposes dramatic effects on speech perception before the end of the first year (Best, 1984, 1993, 1994; Kuhl, 1993; Kuhl & Iverson, 1995; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1984). Still, the biological specialization debate paved the way for this shift in research emphasis, which was initially motivated by the Critical Period hypothesis (CPH; e.g., Lenneberg, 1967). The CPH assumes that language experience must "tune up" the specialized language system during a critical period of early development, to optimize its functions for the requirements of the specific native language the child is learning. This, in turn, implies there should be developmental shifts in the left hemisphere advantage for native versus non-native speech (see Best & Avery, 1999). More recently, other neural plasticity mechanisms, and indeed other developmental processes, have been proposed by which early experience might influence perception much more strongly than does later experience (e.g., Aslin & Pisoni, 1980; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola, & Nelson, 2008). For example, cross-modal audiovisual speech sensitivities clearly change across the first year as a result of language-specific experience (Pons, Lewkowicz, Soto-Faraco, & Sébastian-Gallés, 2009; Best, Kroos, & Irwin, 2010, 2011; Best, Kroos, Gates, & Irwin, 2014).

The third phase of research (1990s into the present) has extended investigations upward from infancy to include the first half of the second year (here, "toddlers") and beyond. This third research phase reflects further overlap with, and refocusing of, the two prior main theoretical issues. Specifically, the current research phase embodies a transition from the emphasis of the first two phases on infants' perception of *phonetic contrasts* to an emphasis on infants' and toddlers' discovery and recognition of *word-forms* as larger, lexically-relevant units that lead the way to the child's further insights that words are composed of recombinant sub-elements (consonants, vowels, syllables, stress patterns, inflectional morphemes, etc.), and can themselves be combined into larger units of meaning according to syntactic and discourse rules (phrases, sentences, stories, etc.).

The remainder of the chapter will focus on theoretical models and empirical findings from the second and third phases of infant speech research, that is, effects of specific language experience on perception of consonant and vowel contrasts in the first year, and their relevance to emerging recognition of known/familiarized word-forms from the end of the first year through the middle of the second year. Several models have been advanced to attempt to account for developmental changes in infants' perception of non-native and native phonetic contrasts (the second research phase). Some of these models have been extended more recently to address development of word-form recognition and learning (third phase). We will now describe the key points of the main theoretical models, to set the stage for summarizing key research findings on infant's speech perception and toddlers' word-form learning and recognition.

## Theoretical models of the effects of language experience

We summarize the three core models of experiential effects on infant speech perception, along with a more recent model of universal perceptual biases in infant perception of vowel contrasts. The subsequent section will then summarize three key models of the development of spoken word recognition. Note that the differences among the models in each section primarily reflect differences in perspective and emphasis, rather than mutually exclusive accounts of the effects of language experience on the development of speech perception and word recognition.

## Models of perceptual attunement to native speech contrasts in infancy

### *Perceptual Reorganization*

Werker and colleagues reported the first unequivocal evidence of developmental changes in infants' perception of non-native speech contrasts in the first year, by including both an age comparison and cross-linguistic native and non-native speech contrasts in their design. They found that Canadian English-learning 6–7 month olds discriminated non-native (Hindi, Nthlakampx [Thompson Salish]) minimal-pair consonant distinctions as well as a native (Canadian English) one. In contrast, the 10–12 month olds only discriminated the native English contrast, failing to exceed chance performance on the non-native contrasts, similar to adult English listeners. Importantly, infants learning Hindi or Nthlakampx discriminated their native contrast at both ages (Werker *et al.*, 1981; Werker & Tees, 1984; see Best, 2002 for discussion of the historical context and importance of these findings). Werker and colleagues argued against the premise that the decline in non-native discrimination by 10-12 months reflects an absolute loss of ability to perceive speech contrasts that are lacking in the native language, that is, they rejected a strict Critical Period account. They reasoned, rather, that as a result of native language experience a *perceptual reorganization* occurs in speech perception by the third quarter of the first year, which along with other cognitive changes around that age focuses infants' attention toward the phonemes and contrasts that are used in their language environment, and away from contrasts not used in that language (see Werker, 1989, 1991). Perceptual reorganization, unlike permanent loss of sensorineural capacity to discriminate a speech contrast due to lack of early exposure (Critical Period), allows for at least some perceptual readjustment to phonetic distinctions (e.g., in a later-learned second language, or L2) that are not used in the native or first language (L1). The perceptual reorganization account is certainly compatible with both everyday and empirical observations that children well past their first birthday are very good at learning new languages including their speech contrasts, despite the perceptual change in discrimination around 10–12 months. It is also consistent with evidence that even adult second language (L2) learners do show perceptual readjustment to at least some degree for certain types of non-native L2 contrasts.

## *Native Language Magnet (NLM, and updated versions NLM-e [NLM-expanded] and NLNC [Native Language Neural Commitment])*

In the NLM model in its various forms, Kuhl and colleagues instead have adhered to the Critical Period premise that exposure to native speech in the first year narrows the brain's initial state of universal sensorineural responsiveness to the acoustic features of speech in any language, toward just those features used in the language environment. The core premise is that the brain's auditory-phonetic neural circuitry becomes "neurally committed" to processing just those distinctions that are statistically supported in the speech input to the infant (e.g., Kuhl, 1993, 2004; Kuhl & Iverson, 1995; Kuhl *et al.*, 2008). This early exposure thus "warps" perception, such that prototype representations develop for each native phoneme category, which function as perceptual magnets, increasing the perceived similarity to phonetic variants within a category by reducing perceptual sensitivity to those within-category differences, while conversely enhancing sensitivity to phonetic differences across the boundaries between contrasting categories (Kuhl, 1991; Kuhl *et al.*, 1992). This "magnet effect" is the source of the facilitated discrimination of native speech contrasts and reduced non-native speech discrimination seen in the second half of year one (Kuhl *et al.*, 2006).

## *Perceptual Assimilation Model (PAM)*

The Perceptual Assimilation Model is more compatible with perceptual reorganization principles than with the notion of a critical period. However, PAM diverges from both models above in positing a linguistically-relevant basis by which L1 and L2 learners, as well as mature L1 users, perceive non-native consonants and vowels (and lexical tones), specifically, it uniquely posits that listeners *perceptually assimilate* non-native/L2 phones to L1 phonological categories based on perceived phonetic similarities, both within and between L1 categories, of the native phonological system (Best, 1984, 1993, 1994, 1995; also Best & Tyler, 2007). A key contribution of PAM, supported by empirical reports, is its prediction, based on those principles, that although native language experience yields poor categorization and discrimination of some non-native contrasts, not all non-native contrasts will become difficult to perceive, as the perceptual reorganization and critical period principles would imply. This reflects variations in the phonetic fit of specific non-native phones to the perceiver's L1 phonological system, whereby the perceived goodness of fit determines whether the non-native phone is perceptually assimilated to a given native phoneme (Categorized to a given native phonological category), is partially assimilated to two or more native phonemes (Uncategorized speech sound), or fails altogether in perceptual assimilation to native phonology (Non-Assimilated, heard as a non-speech sound). These assimilation differences lead to systematic variations in discrimination of non-native *contrasts*, which range from poor categorization and discrimination of contrasting non-native segments when both are assimilated with equal goodness of fit to the same Single Category

(SC) in the native phonological system, to excellent near-native discrimination when the contrasting non-native segments are assimilated to Two Categories (TC) in the native phonological system or as an Uncategorized-Categorized distinction (UC), or as distinct Non-Assimilable non-speech sounds (NA) (for full details, see e.g., Best, 1995; Best & Tyler, 2007). Thus, some non-native contrasts remain easily distinguishable even by adults. The PAM perspective on infant speech perception (Best, 1984, 1993, 1994) is that the native phonological system does not emerge during the first year. Infants are instead becoming attuned to the surface phonetic properties and distributions in the ambient speech of their language environment, and this phonetic-only level of attunement is what is reflected in the developmental changes in infants' discrimination of non-native versus native speech distinctions. Thus, even at their first birthday and beyond, infants do not yet display the adult pattern of assimilation to native *phonological* segments and contrasts, because they do not yet grasp the phonological structure of native words (see section below on development of word-form recognition). One extension of PAM with respect to infant speech perception has been the *Articulatory Organ Hypothesis (AOH)* (e.g., Goldstein & Fowler, 2003). The AOH posits that young infants are initially sensitive to *between-organ* speech distinctions, that is, those made by different primary articulators (such as a lip versus a tongue tip constriction:/p/ versus/t/), regardless of whether they occur in the native language or not. Conversely, young infants are relatively insensitive to *within-organ* contrasts, that is, speech distinctions made by the same primary articulator (e.g., complete closure versus narrow constriction of the tongue tip: stop/t/versus fricative/s/), again regardless of whether they are native or non-native. Between-organ contrasts should remain easy for older infants, children, and adults to perceive even if they do not occur in the native language, whereas non-native within-organ contrasts should become more difficult to perceive if they are non-native, and may require some amount of language experience to become perceptible even if they are native. The AOH applies primarily to consonant distinctions (many vowels contrast by height and/or frontness differences in tongue body position, a single articulator). Some findings are compatible with the AOH (Best & McRoberts, 2003; Kuhl *et al*., 2006), but others are difficult to reconcile with it (Tyler, Best, Goldstein, & Antoniou, 2014).

What might account for the discrepant findings? We offer the following speculative analysis. AOH-compatible results of Best and McRoberts (2003) were for stop consonants differing in within-organ laryngeal distinctions versus between-organ place of articulation distinctions, whereas the AOH-incompatible results of the Tyler *et al*. study instead were for fricative place of articulation differences, within-versus between-organ by the AOH definition. In fricatives, however, because turbulent airflow is generated at a narrow constriction formed between the active articulator and a passive articulatory landmark along the upper surface of the vocal tract (place of articulation), the frication "noise" reflects the location and shape of the constriction more than it reflects the active articulator that formed the constriction. For example, English labiodental/f/and interdental/θ/fricatives have fairly similar amplitude, spectral properties and duration despite the fact that they are formed by different *active* articulators (lower lip versus tongue tip).

There is, though, another articulatory commonality between them: they use the same *passive* constriction location (upper front teeth). Conversely,/s/uses the same active articulator as/θ/(tongue tip) but the constriction appears at a different passive articulator location (alveolar ridge); as a result, these two fricatives instead have very different amplitude, spectral and duration characteristics. Therefore, it may be that for fricatives, a within- versus between-organ distinction might work with respect to the *passive* rather than the *active* articulator. The Tyler *et al*. results are in fact quite compatible with that possibility.

### *Natural Referent Vowels (NRV)*

This model posits, in some complement to the reasoning of the AOH, that infants are born with natural biases in perception of vowel distinctions, which can be shifted by relevant native language experience (Polka & Bohn, 2003, 2011). Specifically, the NRV proposes that there is a natural bias for vowels at the periphery of the vowel space, particularly those at the corners of the space, to serve as perceptual anchors, and thus it is intrinsically easier for infants to discriminate a peripheral from a non-peripheral vowel than vice versa (e.g., peripheral/i/→ non-peripheral [lower, more central] English/ɪ/is discriminated better than/ɪ/→/i/) regardless of whether the contrast occurs in their native language or not. Adults retain the perceptual asymmetry for non-native vowels, but it is inhibited by language experience if the contrast occurs in the native language. These patterns have been upheld by a number of their own infant vowel perception studies (Polka & Bohn, 1996; Polka, Bohn, & Molnar, 2005; Polka & Werker, 1994), as well as by findings from other labs (see Polka & Bohn, 2003, 2011).

## Developing word-form recognition from late infancy through toddlerhood

### *Word Recognition And Phonetic Structure Acquisition (WRAPSA)*

WRAPSA was the first model to bridge from infant speech perception to the first steps in language development: learning to recognize spoken words (e.g., Jusczyk, 1993, 1997; also Jusczyk & Luce, 2002). According to WRAPSA, young infants initially extract universal acoustic cues from speech, but with increasing experience they develop schemes for perceptually weighting the various acoustic dimensions appropriately for their language environment. This native-language weighting of acoustic properties in turn guides infants' attention to the critical properties of spoken native words. Infants store traces of all exemplars of each word they encounter, including indexical information about those exemplars (talker, gender, affective state, familiar person or stranger, etc.). Their comparison of newly incoming utterances against the stored exemplars allows them to recognize known words (perceived as similar to previously-encountered exemplar traces of a familiar word) or to treat unfamiliar/too-dissimilar patterns as possible novel words to learn. At 7.5 months infants can recognize familiarized words in subsequently presented sentences spoken by a different person if the two talkers are the same gender, but not if the two talkers differ in gender, which suggests that at this age they already store talker-gender information of encountered exemplars. Only by 10.5 months can infants generalize word recognition across talker gender (Houston & Jusczyk, 2000). Moreover, if sentence testing was delayed by a day, the younger group could only recognize the words if the sentences were spoken by the same talker (Houston & Jusczyk, 2003).

## *Processing Rich Information in Multidimensional Interactive Representations (PRIMIR)*

PRIMIR is an expansion of Werker's original *perceptual reorganization* hypothesis into a broader consideration of multiple factors. Like WRAPSA, it bridges infant speech perception to spoken word learning and recognition (Werker & Curtin, 2005; extended to bilingual infants by Curtin, Byers-Heinlein, & Werker, 2011), and is indeed compatible with WRAPSA's core principles. In both models, infants' statistical tracking (aka distributional learning) of speech features is the basis both for perceptual reorganization via native speech experience, and for word learning. PRIMIR posits that perceptual learning is modulated by attentional filters (innate speech preferences; infant developmental stage; task demands) that dynamically shift the salience of various types of information in speech, which are grouped onto three planes: General Perceptual, Word Form, and Phonemic information. The General Perceptual plane contains phonetic and indexical information about utterances; the Word Form plane contains tracked word exemplars and their associated concepts; the Phoneme plane emerges as generalizations about phonological features shared among word forms. Thus, word forms are initially phonetically based but later become phonemically based. The authors note that PRIMIR is a framework for understanding existing findings and guiding further research, not a theoretical model *per se*. PRIMIR attributes seemingly contradictory findings to task- and stage-driven shifts in attention, for example, observations that 14 month olds can distinguish minimal-pair nonce words in a pure speech discrimination task but not in a difficult word learning task (Stager & Werker, 1997), yet they *do* learn the words under reduced task demands (Fennell & Waxman, 2010; Gogate, 2010).

## *Perceptual Attunement to Word Structure (PAWS)*

Best and colleagues expanded the core principles of the Perceptual Assimilation Model to address how children come to recognize spoken words across natural phonetic variations *within* the native language (Best *et al.*, 2009, 2016; Mulak & Best, 2013; Watson, Robbins, & Best, 2014). PAWS research has focused especially on recognition of words/sentences spoken in unfamiliar non-native regional accents, which can deviate notably from the native accent (e.g., American English [native] versus Jamaican Mesolect [non-native], i.e., creole-influenced English). PAWS posits that infants in the first year detect phonetic-articulatory details in speech, but do not yet recognize their relevance to abstract phonological properties of the language until year two. This initial phonetic focus results in perceptual attunement to various systematic patterns in native speech by 6–12 months, which in turn paves the way for recognition of more abstract phonological features of native-accented words in year two, by the time they have acquired a 50+ word expressive vocabulary (~17–19 months). This shift from purely phonetic to more abstract phonological perception, according to PAWS, emerges as the child begins to sort out the linguistic versus indexical sources of natural phonetic variations within and across talkers in their speech environment. These insights allow the children to discover the complementary relationship between *phonological distinctiveness*, which signifies that a single critical phonetic feature difference between two word forms indicates a phonological contrast between two different native words (or between a word and a non-word); and *phonological constancy*, which signifies that the identity of a given spoken word is *not* changed by talker- or accent-related pronunciation differences that are indexical rather than linguistically relevant phonetic variations. Studies thus far have examined, and supported, the prediction that phonological constancy across accents emerges sometime between 15 and 19 months (Best *et al*., 2009, 2010, 2012, 2014; Best & Kitamura, 2014; Mulak *et al*., 2013).

## Synopsis: Infant speech perception and spoken word recognition

Reviewing the full array of findings on these two topics is beyond the scope of this chapter; numerous existing reviews offer more comprehensive details (e.g., Best, 1994; Gervain & Mehler, 2010; Jusczyk, 1997; Mulak & Best, 2013; Werker, 1989; Werker & Yeung, 2005; Werker, Yeung, & Yoshida, 2012; see also this volume: Levine, Strother-Garcia, Hirsh-Pasek, & Golinkoff). Here, we will instead synopsize key findings on infants' perception of speech segments and recognition of spoken words, in relation to theoretical models and principles.

### *Tuning in to native consonants, vowels, and minimal contrasts*

Infants under 6 months discriminate most vowel and consonant contrasts regardless of prior exposure. Native versus non-native vowel contrasts are perceived differently, however, by 6 months, and by 9–10 months for consonants (see reviews: Best, 1993, 1994; Gerken & Aslin, 2005; Werker, 1989, 1991; Werker *et al*., 2012). Perceptual constancy across talker gender and age for some native vowels (e.g., Kuhl, 1979, 1983), and a stronger perceptual "magnet effect" (perceived similarity) for prototypical versus non-prototypical tokens of one native vowel than one non-native vowel, have been found at 6 months (Kuhl *et al*., 1992), but no developmental comparisons were made. Other work has provided a clearer developmental picture: one non-native vowel contrast was discriminated well at 4 months, significantly less well at 6 months, and was not discriminated at 10 months (Polka & Werker, 1994). But other vowel contrasts show asymmetries in discrimination at both 6 and 10 months, regardless of whether they are native or non-native (Polka & Bohn, 1996, 2003, 2011). Thus, many vowel perception findings are mutually compatible with perceptual reorganization and assimilation (PAM) principles, some also with perceptual magnet principles (NLM). Importantly, however, several findings support universal natural referent vowels (NRV).

Many non-native contrasts are discriminated at 6–8 months but not 10–12 months, while most native consonant contrasts are discriminated well at both ages. However, discrimination of some non-native consonant contrasts remains good at 10 months and older (click and ejective place contrasts: Best & McRoberts, 2003; Best, Sithole, & McRoberts, 1988). Conversely, discrimination of some native consonants is poor at both ages (English/d/-/ð/as in *doze-those*: Polka, Colantonio, & Sundara, 2005) or declines between 6–8 and 10–12 months (coronal fricative voicing: Best & McRoberts, 2003). The types of non-native contrasts that do versus do not decline in discrimination are most clearly in line with the PAM model. Moreover, some of the native contrasts that have been found to be poorly discriminated by young infants and/or to show improvement with age and experience are compatible with AOH premises, that is, they are within-organ contrasts (Best & McRoberts, 2003; see also Kuhl *et al.*, 2006). However, a recent study compared within- versus between-organ consonant distinctions in both the native and a non-native language, and found all contrasts discriminated at both 6–8 and 10–12 months of age, a challenge to AOH but consistent with PAM (Tyler *et al.*, 2014).

Those studies used audio-only stimuli. If perception is guided by articulatory information (PAM/AOH) then infants should be sensitive to articulatory congruency between audio and video (talking face) speech. The few existing audiovisual (AV) studies are compatible with that premise: 3 to 4 month olds look preferentially at the talking face that matches a synchronized native (e.g., Kuhl & Meltzoff, 1982) or non-native audio vowel (Walton & Bower, 1993), while 4–5 month olds show a McGurk effect for phonetically incongruous AV consonants (Burnham & Dodson, 2004; Rosenblum *et al.*, 1997) and 6 month olds learn a difficult native audio contrast only if familiarized with distinct visual articulations of the two endpoint stimuli (Teinonen *et al.*, 2008). In cross-modal A→V congruency studies using a within-organ consonant contrast found in English but not Spanish (lips:/b/-/v/), 6 month olds of both languages preferred the silent talking face that matched the audio consonant they had been familiarized with, while only English infants did so at 11 months (Pons *et al.*, 2009). Conversely, similar studies using between-organ contrasts (lip versus tongue tip articulations) found A→V congruency preferences at 4 and 11 months for native English/p/-/t/. For non-native ejectives/p'/-/t'/ 4 month olds showed the same A→V congruency preference, whereas 11 month olds showed a contrary *incongruency* preference, while for non-native clicks/ʘ/-/|/ 4 month olds showed an incongruency preference, whereas 11 month olds lacked any A→V preference (Best *et al.*, 2010, 2011, 2014), which the authors interpret as largely consistent with PAM and AOH predictions, except that the AOH applies only to contrasts perceived as speech. PAM predicts that the older infants should perceive the clicks as nonspeech (Non-Assimilation: Best & Avery, 1999; Best *et al*., 1988) and thus fail to recognize they are related to visible speech, whereas younger infants still hear clicks as speech but recognize they deviate from any native consonants.

## *Discovering phonological structure in native words*

The foundations for recognizing referential words are established during the second half of year one (see, e.g., Anderson, Morgan, & White, 2003; for comprehensive reviews see: Gervain & Mehler, 2010; Jusczyk, 1993, 1997; Mulak & Best, 2013; Werker *et al*., 2012). Given that caregivers talk to infants primarily in connected phrases rather than isolated words, an important precursor to word recognition is the ability to recognize previously-encountered word-forms (the phonetic pattern of a word, even if referential meaning is not yet grasped) embedded within larger utterances, called *word(-form) segmentation*. This ability becomes evident by 7.5 months for words the child has been familiarized with, but can be found as early as 6 months for words that are already highly familiar from the home environment (*mommy, daddy*, the child's own name). Conversely, segmentation may not be reliable until 9 months if the pitch of the familiarization and test voices differs substantially, not until 10.5 months if the talker or emotional affect differs, and not until 12–13 months if talker accent differs (see review: Mulak & Best, 2013). Relatedly, young infants discriminate between sentences spoken in their native accent versus another regional accent of their language, and prefer listening to the native-accented sentences, but accent discrimination and preference decline by 9 months of age if the other accent is unfamiliar, by 6 months if the other accent is familiar. The authors interpreted the decline as evidence for *language constancy*, that is, the infants had begun to recognize that the native language remained constant despite the phonetic differences between the accents, and experience with an accent is not required but can facilitate the process (Kitamura, Best, & Panneton, 2013).

Much research on individual word recognition has addressed toddlers' (12–20 months) ability to recognize that words with referential meanings can be differentiated by a single consonant or vowel, that is, minimal pair words. Children under about 18 months often fail to treat such pairs as different in standard word-learning or word-recognition tasks, whereas those over 18 months do so reliably even with high task demands. Success is correlated with expressive vocabulary size for children around the transition (~17 months). However, even 14 month olds succeed if very familiar words are used, the referential goal is contextually enhanced, task demands are reduced (see review: Mulak & Best, 2013), and/or larger minimal-pair vowel differences are presented (Escudero, Best, Kitamura, & Mulak, 2014). These findings suggest that children under 17 months, particularly those with small expressive vocabularies, recognize words as holistic patterns, but do not yet recognize minimal phonological contrasts until ~18 months and/or the vocabulary spurt (50+ word expressive vocabulary), compatible with both PRIMIR and PAWS, particularly with the idea that *phonological distinctiveness* emerges then (Best, 2015; Best *et al*., 2009). The complementary ability identified by PAWS, *phonological constancy* in word recognition across lexically-irrelevant indexical variations such as regional accents, also emerges when the child has acquired a 50-100 word expressive vocabulary (~17–18 months): familiar words spoken in their native regional accent are recognized at both 15 and 19 months but only the older children recognize words in a very different, unfamiliar regional accent (Best *et al*., 2009; Mulak *et al*., 2013). If the phonetic variation in the task is increased (more words, multiple tokens, multiple talkers) then 15 month olds and 17 month olds with vocabularies under 50 words fail to recognize even native-accented familiar words, 17 month olds with 50+ words vocabularies succeed with native-accented words, and 19 month olds (with 100+ word vocabularies) succeed with both accents (Best *et al*., 2010, 2012). Interestingly, 15 month olds succeed with unfamiliar accents if the familiar words contain only consonants and vowels that adults correctly identify but hear as non-native-accented (Best, Gates, Kitamura, Docherty, Pinet, & Evans, 2014; Best & Kitamura, 2014). Thus, *phonological distinctiveness* and *phonological constancy* appear around the same time and under similar task and stimulus conditions and vocabulary growth levels, suggesting they mark a transition from recognizing words as holistic phonetic patterns to recognizing native-language phonological principles.

## Where do we go from here?

Research to date has focused largely on *processing*. Two less-studied but crucial issues deserve further attention. First is the *nature* of information infants perceive in speech, which must adequately support their discoveries of phonetic patterning and phonological principles so that they become not just native listeners but also native speakers. Acoustic features are often assumed to be the sole relevant perceptual information. However, PAM/AOH and PAWS posit that the underlying information most relevant to language acquisition is *articulatory* rather than acoustic *per se* (e.g., Best, 1994, 1995, 2015; see also Fowler, 1989; Studdert-Kennedy & Goldstein, 2003). Compatible with that premise, recent findings and theoretical claims suggest that speech perception and motor speech (i.e., articulatory) information are linked/unified in infants (e.g., Kuhl, Ramírez, Bosseler, Lin, & Imada, 2014; Bruderer, Danielson, Kandhadai, & Werker, 2015). Second, young language-learners must learn how the phonetic patterning and phonological structure of native words interrelate with both the higher-order linguistic principles of sentences and narratives, that is, prosody, intonation, morphology and syntax (this volume: Milin, Smolka & Feldman), and also with the social side of speech, that is, talker-specific and emotional features, regional and foreign accents, and registers or speech styles (e.g., Kuhl, 2007; Kuhl, Tsao, & Liu, 2003; this volume: Pardo; Roseberry-Lytle & Kuhl). In addition, although this review has focused on the majority of speech perception research that has focused on children learning a single language, many people learn more than one language from birth (see this volume: Kroll & Ma; Traxler, Hoverstern, & Brothers). Research is increasingly being devoted to various aspects of language learning and processing in children exposed to multiple languages, and how this situation may influence their acquisition of language skills ranging from speech perception to lexical access to morphosyntactic operations (e.g., this volume: Byers-Heinlein & Lew-Williams; Meisel; Martohardjono & Klein). Research on these issues has begun to provide important insights, but further investigation will be needed for fuller perspective into how infants learn their native language(s) from spoken input.

## REFERENCES

Anderson, J. L., Morgan, J. L. & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language & Speech*, *42*, 155–182. doi: 10.1177/00238309030460 020601

Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. Yeni Komshian, J. F. Kavanagh & C. A. Ferguson (Eds.), *Child Phonology: Perception and Production*, pp. 67–96. New York NY: Academic Press.

Best, C. T. (2015). Devil or angel in the details? Complementary principles of phonetic variation provide the key to phonological structure. In J. Romero & M. Riera (Eds.) *Sounds, representations and methodologies: Essays on the phonetics-phonology interface*. Current Issues in Linguistic Theory book series, pp. 3–31. John Benjamins: Amsterdam.

Best, C. T. (1984). Discovering messages in the medium: Speech perception and the prelinguistic infant. In H. E. Fitzgerald, B. Lester, & M. Yogman (Eds.) *Theory and research in behavioral pediatrics*, pp. 97–145. New York, NY: Plenum Press.

Best, C. T. (1988). The emergence of cerebral asymmetries in early human development: A literature review and neuroembryological model. In D. L. Molfese & S. J. Segalowitz (Eds.) *Brain lateralization in children: Developmental implications*. pp. 5–34. New York, NY: Guilford Press.

Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech contrasts: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.) *Developmental neurocognition: Speech and face processing in the first year of life*, pp. 289–304. Dordrecht, Netherlands: Kluwer Academic Publishers B. V.

Best, C. T. (1994). Learning to perceive the sound pattern of English. In C. Rovee-Collier & L. Lipsitt (Eds.) *Advances in Infancy Research* (vol. 9). pp. 217–304. Norwood NJ: Ablex Publishing Corporation.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (ed.) *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, pp. 167–200. Timonium MD: York Press.

Best, C. T. (2002). Revealing the mother tongue's nurturing effects on the infant ear. *Infant Behavior & Development*, *25*, 134–139. doi: 10.1016/S0163-6383(02)00111-X

Best, C. T. & Avery, R. A. (1999). Left hemisphere advantage for click consonants is determined by linguistic significance. *Psychological Science*, *10*, 65–69. doi: 10.1111/1467-9280.00108

Best, C. T., Gates, S., Kitamura, C., Docherty, G., Pinet, M. & Evans, B. G. (2014). Young Australian toddlers recognize Cockney-accented words involving only "category-goodness" differences in consonant pronunciations from Australian English. *International Conference on Infant Studies*, Berlin, Germany, 3–5 July.

Best, C. T., Goldstein, L., Tyler, M. D. & Nam, H. (2016). Articulating what infants attune to in native speech. *Ecological Psychology*, *28*, 216–261. doi: 10.1080/10407413.2016.1230372

Best, C. T., Hoffman, H. & Glanville, B. B. (1982). Development of infant ear asymmetries for speech and music. *Perception & Psychophysics*, *31*, 75–85. doi: 10.3758/BF03206203

Best, C. T., Kroos, C. H. & Irwin, J. (2010). Now I see what you said: Infant sensitivity to place congruency between audio-only and silent-video presentations of native and non-native consonants. *Proceedings of AVSP (AudioVisual Speech Perception)*. Hakone, Japan, Sept-Oct.

Best, C. T., Kroos, C. H. & Irwin, J. (2011). Do infants detect A→V articulator congruency for non-native click consonants? *Proceedings of AVSP Proceedings (AudioVisual Speech Perception)*. Volterra, Italy, Sept.

Best, C. T., Kroos, C. H., Gates, S. & Irwin, J. (2014). Baby steps in perceiving articulatory foundations of phonological contrasts: Infants detect audio→video congruency in native and non-native consonants. *Laboratory Phonology*, Tokyo, Japan, 25-27 July.

Best, C. T. & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language & Speech*, *46*, 183–216. doi: 10.1177/00238309030460020701

Best, C. T., Morrongiello, B. A. & Robson, R. (1981). Perceptual equivalence of two acoustic cues in speech and nonspeech perception. *Perception & Psychophysics*, *29*, 191–211. doi: 10.3758/BF03207286

Best, C. T. & Kitamura, C. (2014). The role of perceptual assimilation in early development of recognition of words spoken in native vs unfamiliar regional accents. *International Conference on Infant Studies*, Berlin, Germany, 3-5 July.

Best, C. T., Kitamura, C., Pal, M. & Dwyer, A. (2012). Young toddlers recognize non-native Jamaican-accented words differing only in "category-goodness" from native-accent pronunciations. *International Conference on Infant Studies*, Minneapolis MN, USA, June 7-9.

Best, C. T., Tyler, M. D., Kitamura, C. & Bundgaard-Nielsen, R. (2010). Vocabulary size at 17 months and the emergence of phonological constancy in word recognition across native and non-native dialects. *International Conference on Infant Studies*. Baltimore MD, March.

Best, C. T. & Tyler, M. D. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In M. Munro & O.-S. Bohn (Eds.) *Second Language Speech Learning*, pp. 13–34. Amsterdam: John Benjamins Publishing.

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B. & Quann, C.A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, *20*, 539–542. doi: 10.1111/j.1467-9280.2009.02327.x

Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences*, *112(44)*, 13531–13536. doi: 10.1073/pnas.1508631112

Burnham, D. & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology*, *45*, 204–220. doi: 10.1002/dev.20032

Chomsky, N. (1975). *Reflections on language*. Pantheon: New York NY.

Curtin, S. A., Byers-Heinlein, K. & Werker, J. F. (2011). Bilingual beginning as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*, 492–504. doi: 10.1016/j.wocn.2010.12.002

Diehl, R. L. & Kluender, K. L. (1989a). On the objects of speech perception. *Ecological Psychology*, *1*, 121–144. doi: 10.1207/s15326969eco0102_2

Diehl, R. L. & Kluender, K. L. (1989b). Reply to commentators. *Ecological Psychology*, *1*, 195–225. doi: 10.1207/s15326969eco0102_6

Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. *Perception & Psychophysics*, *18*, 341–347. doi: 10.3758/BF03211210

Eimas, P. D. (1985). The equivalence of cues in the perception of speech by infants. *Infant Behavior & Development*, *8*, 125–138. doi: 10.1016/S0163-6383(85)80001-1

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303–306. doi: 10.1126/science.171.3968.303

Escudero, P. N., Best, C. T., Kitamura, C. & Mulak, K. (2014). Magnitude of phonetic distinction predicts success at early word learning: A cross-accent study. *Frontiers in Language Science*, 01059. doi: 10.3389/fpsyg.2014.01059

Fennell, C. T. & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, *81*(5), 1376–1383. doi: 10.1111/j.1467-8624.2010.01479.x

Fowler, C. A. (1984). Segmentation of coarticulated speech in perception. *Perception & Psychophysics*, *36*, 359–368. doi: 10.3758/BF03202790

Fowler, C. A. (1989). Real objects of speech perception: A commentary on Diehl and Kluender. *Ecological Psychology*, *1*, 145–160. doi: 10.1207/s15326969eco0102_3

Fowler, C. A., Best, C. T. & McRoberts, G. W. (1990). Young infants' perception of liquid coarticulatory influences on following stop consonants. *Perception & Psychophysics*, *48*, 559–570. doi: 10.3758/BF03211602

Gerken, L.A. & Aslin, R. N. (2005). Thirty years of research on infant speech perception: The legacy of Peter W. Jusczyk. *Language Learning & Development*, *1*, 5–21. doi: 10.1207/s15473341lld0101_3

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, *61*, 191–218. doi: 10.1146/annurev.psych.093008.100408

Glanville, B. B., Best, C. T. & Levenson, R. (1977). A cardiac measure of cerebral asymmetries in infant auditory perception. *Developmental Psychology*, *13*, 54–59. doi: 10.1037/0012-1649.13.1.54

Gogate, L. J. (2010). Learning of syllable–object relations by preverbal infants: The role of temporal synchrony and syllable distinctiveness. *Journal of Experimental Child Psychology 105*, 178–197. doi: 10.1037/0012-1649.13.1.54

Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In: N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities*, pp. 159–208. Berlin: Mouton de Gruyter.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1570–1582. doi: 10.1037/0096-1523.26.5.1570

Houston, D. M., & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 1143–1154. doi: 10.1037/0096-1523.29.6.1143

Jusczyk, P. W. (1993). From general to language specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*, *21*, 3–28.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Jusczyk, P. W. & Derrah, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology*, *23*, 648–654. doi: 10.1037/0012-1649.23.5.648

Jusczyk, P. W. & Luce, P. A. (2002). Speech perception and spoken word recognition: Past and Present. *Ear & Hearing*, *23*, 2–40.

Jusczyk, P. W., Pisoni, D. B., Reed, M. A., Fernald, A. & Myers, M. (1983). Infants' discrimination of the duration of a rapid spectrum change in nonspeech signals. *Science*, *222*, 175–177. doi: 10.1126/science.6623067

Jusczyk, P.W., Pisoni, D. B., Walley, A. & Murray, J. (1980). Discrimination of relative onset time of two-component tones by infants. *Journal of the Acoustical Society of America*, *67*, 262–270. doi: 10.1126/science.6623067

Jusczyk, P.W., Rosner, B. S., Cutting. J. E., Foard, C. F. & Smith, L. B. (1977). Categorical perception of nonspeech sounds by 2-month-old infants. *Perception & Psychophysics*, *21*, 50–54. doi: 10.3758/BF03199467

Kluender, K. R., Diehl, R. L. & Killeen, P. L. (1987). Japanese quail can learn phonetic categories. *Science*, *237*, 1195–1197. doi: 10.1126/science.3629235

Kluender, K. R. & Lotto, A. R. (1994). Effects of first formant onset frequency on [-voice] judgments result from auditory processes not specific to humans. *Journal of the Acoustical Society of America*, *95*, 1044–1052. doi: 10.1121/1.408466

Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, *66*, 1668–1679. doi: 10.1121/1.383639

Kuhl, P. K. (1981). Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. *Journal of the Acoustical Society of America*, *70*, 340–349. doi: 10.1121/1.386782

Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior & Development*, *6*, 263–285. doi: 10.1016/S0163-6383(83)80036-8

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*, 93–107. doi: 10.3758/BF03212211

Kuhl, P. K. (1993). Innate predispositions and the effects of experience in speech perception: The Native Language Magnet Theory. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.) *Developmental neurocognition: Speech and face processing in the first year of life*, pp. 259–274. Dordrecht, Netherlands: Kluwer Academic Publishers B. V.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews: Neuroscience*, *5*, 831–843. doi: 10.1038/nrn1533

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, *10(1)*, 110–120. doi: 10.1111/j.1467-7687.2007.00572.x

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B*, *363*, 979–1000. doi: 10.1098/rstb.2007.2154

Kuhl, P. K. & Iverson, P. (1995). Linguistic experience and the "Perceptual Magnet Effect." In W. Strange (ed.) *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, pp. 121–154. Timonium MD: York Press.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*, 1138–1141. doi: 10.1126/science.7146899

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, *190(*4209*)*, 69–72. doi: 10.1126/science.1166301

Kuhl, P. K. & Padden D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America*, *73*, 1003–1010. doi: 10.1121/1.389148

Kuhl, P. K., Ramírez, R. R., Bosseler, A., Lin, J. F. L., & Imada, T. (2014). Infants' brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences*, *111(*31*)*, 11238–11245. doi: 10.1073/pnas.1410963111

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S. & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, F13–F21. doi: 10.1111/j.1467-7687.2006.00468.x

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100(*15*)*, 9096–9101. doi: 10.1073/pnas.1532872100

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606–608. doi: 10.1126/science.1736364

Kuhn, S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461. doi: 10.1037/h0020279

Liberman, A. M. & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36. doi: 10.1016/0010-0277(85)90021-6

Liberman, A. M. & Mattingly, I. G. (1989). A specialization for speech perception. *Science*, *243*, 489–494. doi: 10.1126/science.2643163

Lotto, A. J., Kluender, K. R., & Holt, L. L. (1997). Perceptual compensation for coarticulation by Japanese quail (*Coturnix coturnix japonica*). *Journal of the Acoustical Society of America*, *102*, 1134–1140. doi: 10.1121/1.419865

MacKain, K. S., Studdert-Kennedy, M., Spieker, S. & Stern, D. (1983). Infant intermodal speech perception is a left hemisphere function. *Science*, *219*, 1347–1349. doi: 10.1126/science.6828865

Mann, V. A. (1980). Influence of preceding liquid on stop-consonant perception. *Perception & Psychophysics*, *28*, 407–412. doi: 10.3758/BF03204884

Mann, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of "l" and "r." *Cognition*, *14*, 211–235. doi: 10.3758/BF03204884

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748. doi: 10.1038/264746a0

Miller, J. L., & Eimas, P. D. (1983). Studies on the categorization of speech by infants. *Cognition*, *13*, 135–165. doi: 10.1016/0010-0277(83)90020-3

Molfese, D. L., Freeman, R. B. & Palermo, D. S. (1975). The ontogeny of brain lateralization for speech and nonspeech stimuli. *Brain & Language*, *2*, 356–368. doi: 10.1016/S0093-934X(75)80076-9

Molfese, D. L. & Molfese, V. J. (1985). Electrophysiological indices of auditory discrimination in newborn infants: Predicting later language development? *Infant Behavior & Development*, *8*, 197–211. doi: 10.1016/S0093-934X(75)80076-9

Morrongiello, B. A., Robson, R., Best, C. T. & Clifton, R. K. (1984). Trading relations in the perception of speech by five-year-old children. *Journal of Experimental Child Psychology*, *37*, 231–250. doi: 10.1016/S0093-934X(75)80076-9

Morse, P. A. (1972). The discrimination of speech and nonspeech stimuli in early infancy. *Journal of Experimental Child Psychology*, *14*, 477–492. doi: 10.1016/0022-0965(72)90066-5

Mulak, K. & Best, C. T. (2013). Development of word recognition across speakers and accents. In L. Gogate & G. Hollich (Eds.) *Theoretical and Computational Models of Word Learning: Trends in Psychology and Artificial Intelligence*, pp. 242–269. Hershey PA: IGI Global-Robotics.

Mulak, K., Best, C. T., Tyler, M. D., Kitamura, C. & Irwin, J. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify familiar words spoken in a non-native regional accent. *Child Development*, *84*, 2064–2078. doi: 10.1111/cdev.12087

Peperkamp, S. (2003). Phonological acquisition: Recent attainments and new challenges. *Language and Speech*, *46*(2-3), 87–113. doi: 10.1177/00238309030460020401

Polka, L., Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, *100*, 577–592. doi: 10.1121/1.415884

Polka, L. & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, *41*, 221–231. doi: 10.1016/S0167-6393(02)00105-X

Polka, L. & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*, 467–478. doi: 10.1016/j.wocn.2010.08.007

Polka, L., Bohn, O.-S. &Molnar, M. (2005). Natural referent vowels guide the development of vowel perception. *Journal of the Acoustical Society of America*, *117*, 2398. doi: 10.1121/1.4785978

Polka, L., Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 421–435. doi: 10.1037/0096-1523.20.2.421

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 10598–10602. doi: 10.1073/pnas.0904134106

Popper, K. (1957). *The poverty of historicism*. Harper & Row: New York NY

Remez, R. E. (1989). When the objects of perception are spoken. *Ecological Psychology*, *1*, 161–180. doi: 10.1207/s15326969eco0102_4

Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophysics*, *59*, 347–357. doi: 10.1207/s15326969eco0102_4

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*(6640), 381–382.

Studdert-Kennedy, M. (1989). Cues to What? A Comment on Diehl and Kluender 'On the objects of speech perception'. *Ecological Psychology*, *1*, 181–193. doi: 10.1207/s15326969eco0102_5

Studdert-Kennedy, M. & Goldstein, L. (2003). Launching language: The gestural origin of discrete infinity. In M. H. Christiansen & S. Kirby (Eds.) *Language evolution*, pp. 235–254. Oxford: Oxford University Press.

Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*, *108*, 850–855.

Tyler, M. D., Best, C. T., Goldstein, L. & Antoniou, M. (2014). Contributions of native-language tuning and articulatory organs to infants' discrimination of native and non-native consonant contrasts. *Developmental Psychobiology*, *56*, 210–227. doi: 10.1002/dev.21195

Walton, G. E., & Bower, T. G. (1993). Amodal representations of speech in infants. *Infant Behavior & Development*, *16*, 233–243. doi: 10.1016/0163-6383(93)80019-5

Watson, T. L., Robbins, R. A. & Best, C. T. (2014). Development of face and speech perception: An integrated approach. *Developmental Psychobiology*, *56*, 1454–1481. doi: 10.1002/dev.21243

Werker, J. F. (1989). Becoming a native listener. *American Psychologist*, *77*, 54–59.

Werker, J. F. (1991). The ontogeny of speech perception. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception* (pp. 91–109). Hillsdale, NJ: Erlbaum.

Werker, J. F., & Curtin, S. A. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197–234. doi: 10.1080/15475441.2005.9684216

Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*, 348–355. doi: 10.2307/1129249

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*(1), 49–63. doi: 10.1016/S0163-6383(84)80022-3

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences*, *9(*11*)*, 519–527. doi: 10.1016/j.tics.2005.09.003

Werker, J. F., Yeung, H. H. & Yoshida, K. A. (2012). How do infants become experts at native-speech perception? *Current Directions in Psychological Science*, *21*, 221–226. doi: 10.1177/0963721412449459

# 22 Children's Performance Abilities: Language Production

CECILE McKEE[1], DANA McDANIEL[2], AND MERRILL F. GARRETT[3]

[1] University of Arizona
[2] University of Southern Maine
[3] University of Arizona

## Introduction

Despite the emphasis in the field of language acquisition on how children talk, study of the production system itself is somewhat rare. There is some irony in this because strong claims about children's linguistic competence often rely on production data. This situation is muddied by the fact that comprehension generally precedes production (e.g., Benedict, 1979; Bornstein & Hendricks, 2012; Golinkoff, Ma, Song, & Hirsh-Pasek, 2013). Some scholars remark on this irony (e.g., Marshall, 1979; McKee, Rispoli, McDaniel, & Garrett, 2006; Wijnen, 1990), and some study the development of this performance system. We will sample that research here, describing the production system and illustrating how it influences the data that acquisitionists use to buttress conclusions about children's linguistic competence.

Adult production data is generally considered an insufficient basis for analyses of linguistic competence for good reasons: The rarity of a structure doesn't necessarily bear on its grammaticality, and speech errors can yield ungrammatical sentences. Such factors are arguably of more concern in children, who are new to speaking and have a smaller vocabulary. Both continuity and discontinuity accounts typically ignore the possibility that the production system might separately influence utterances. Three types of claims along these lines are: (A) The lack of a structure in children's speech means that the structure isn't part of their

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

competence. (B) An occasionally occurring non-adult structure indicates a non-adult grammar. (C) A consistently occurring non-adult structure indicates a non-adult grammar.

Regardless of the theoretical details, the logic in such claims is the same: Non-adult speech indicates non-adult knowledge. In our view, this logic is flawed because it doesn't consider contributions of the production system. Even in the case of (C), where children's speech is consistent, pressures on the developing production system could yield consistent non-adult structures that are ungrammatical for the child. Note that (A), (B), and (C) are negative claims about children's knowledge. In our view, Occam's Razor makes positive claims less problematic: Adult speech indicates adult knowledge, regardless of the speaker's age. But positive claims should also be treated with caution, since the developing production system could play a role here too. Our point is not that production data should be avoided when we study linguistic competence or that the claims of research based on such data are necessarily wrong. Rather, caution is needed whenever we interpret production data. It is the output of a performance system, and so it is not a direct indicator of competence.

The developing production system is interesting in its own right. It involves relations between declarative and procedural knowledge and the effects on one domain of resource limitations in another domain. As Charest and Johnston (2011) observe, three characteristics of language production raise developmental questions. First, the resources required for production under certain circumstances (e.g., when an utterance is long) might exceed the child's capacities. Second, costs associated with different components of the production process may vary independently (e.g., the costs associated with the syntactic formulation of an utterance differ from those associated with retrieving the lexical items). Third, the cognitive load from one component might affect another component; in other words, there are trade-offs among the different components of the production process. In this chapter, we emphasize the findings on individual components of the production system, but we note that there is also intriguing evidence for trade-offs (e.g., Bloom, Miller, & Hood, 1975; Boyle & Gerken, 1997; Masterson, 1997; Masterson & Kamhi, 1992; Nelson & Bauer, 1991).

Although we won't pursue this here, it is also the case that some language impairments are fruitfully analyzed in terms of the production system per se (see Charest & Johnston, 2011, for discussion). Research on impairments therefore both contributes to and benefits from greater understanding of the developing production system.

## Theoretical considerations

Before turning to children's production abilities, we briefly describe the theoretical framework that dominates in this area of psycholinguistics (Bock, 1987; Dell, 1986; Fromkin, 1971; Garrett, 1980a; Levelt, 1989). It provides a foundation for evaluating children's developing capacity for the integration of lexical, syntactic, morphological, and phonological knowledge in real time as they produce sentences.

This framework describes early and late stages of processing. The earliest stage generates a preverbal representation of the content to be uttered. This message reflects discourse and semantic objectives. Message-level representations are then mapped onto language-specific forms in processes that engage both lexical and grammatical representations. Lexical representations include different types of information, which are retrieved at different stages. The lemma associated with an open class word, for example, contains semantic and syntactic information that is integrated with morphosyntactic information during grammatical encoding. The latter includes what will eventually be instantiated as order and agreement relations. The next stage determines the phonological organization for the preceding stage's output, and a word form retrieval step recovers the remaining lexical information, which is pertinent to sound structure and related metrical processes.[1] These operations are incremental and overlapping. For example, consider this simple clause: *Chris likes water*. The information that will be produced as "Chris," "like," and "water" is represented early. The information that will be produced as "-s" is also represented early, but it is realized later in the series of operations than is its host "like" (see, e.g., Levelt, 1989, pp. 321–326). Finally, the framework reflects the fact that speakers monitor the output of production processes, sometimes modifying their utterances accordingly. Figure 22.1, from Levelt (1989, p. 9), shows the processes and information types just described. This figure also indicates relations between the comprehension and production systems.

Although production models include a great deal more than phonology, lexicon, and syntax, we will limit the following to research emphasizing these three domains. But some of the research that we discuss below touches on other components in Figure 22.1 as well, such as message-level representations and monitoring of the output.

## Methodological considerations

Like production research with adult speakers, developmental research uses spontaneously produced or naturally occurring speech as well as experimentally elicited speech. Spontaneous speech has the advantage of being generated naturally, whereas experimentally elicited speech has the advantage of being generated under controlled circumstances. Speech can be elicited with or without modeled utterances. In an imitation task, the model consists of the target sentence itself, whereas the model in a priming task is a particular structure. (Syntactic priming occurs when the comprehension or production of a structure facilitates its subsequent production. See Bock, 1986a, 1986b.) Other elicited production tasks provide a context that encourages the targeted structure without modeling it. In fact, all elicitation protocols can be accompanied by a context, using toys or pictures. Such contexts constrain the message. For example, in some studies of naturalistic speech, subjects narrate a wordless picture book, and, in some imitation studies, the imitated sentence refers to an enacted scenario. In structural priming studies, both the prime and the target sentence describe a picture.

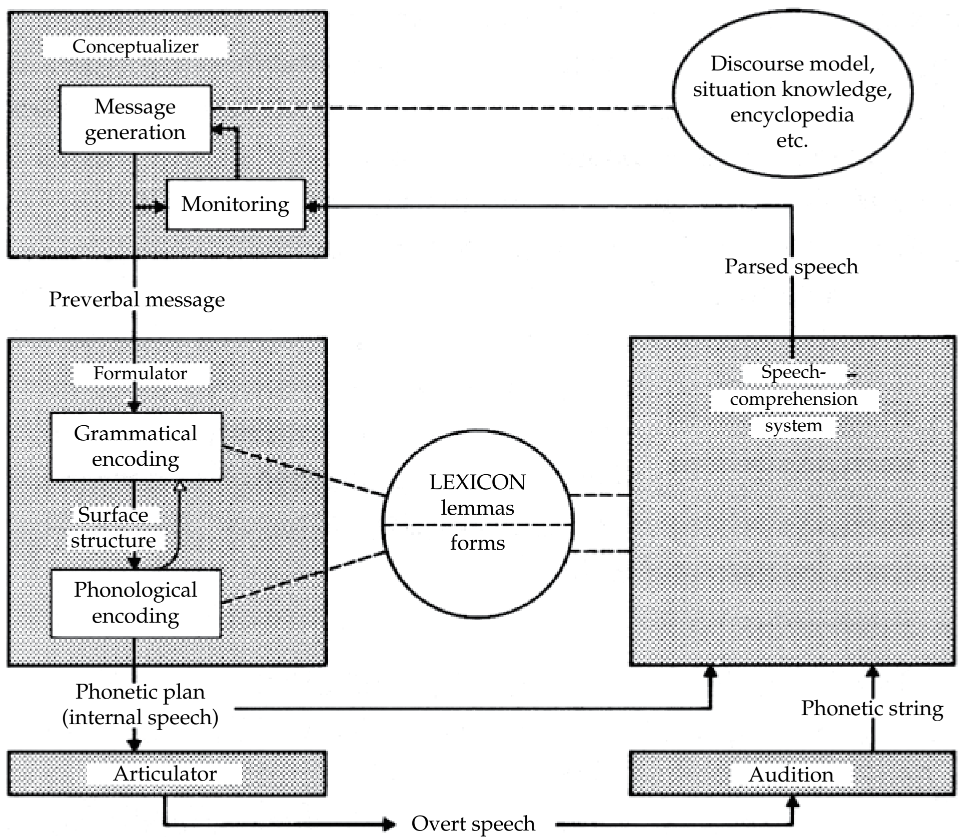



**Figure 22.1** Model of the production processes (Levelt, Willem J. M., *Speaking: From Intention to Articulation*, figure 1.1, page 9, © 1989 Massachusetts Institute of Technology, by permission of The MIT Press).

Production research also includes a range of different types of data. Among the most common are speech errors or slips of the tongue, dysfluency patterns, priming, and measures of rate.[2] These types of data can reflect different aspects of the production process.

Regarding speech errors, note that they occur only occasionally (e.g., Fromkin, 1971; Jaeger, 2005; MacKay, 1970 [who analyzed errors from Meringer, 1908]; Stemberger, 1989; Wijnen, 1992). Research on phonological and lexical development often analyzes errors in children's speech that might be either speech errors if encountered in adult study or reflect limitations in linguistic competence. With child speakers, it is challenging yet essential to distinguish slips of the tongue from errors reflecting non-adult knowledge. Stemberger (1989) discussed this problem, reporting this morphological error from an adult: *I had never connect them*. He illustrated how a similar utterance produced by a child is harder to analyze unless the child's mastery of past tense is known or if the child self-corrects, as in one example he reported: *Now it's end – Now it's ended*. Slips can be categorized by the linguistic elements involved (e.g., syllable onsets, words) and by the processes that operate on those elements (e.g., exchanges, blends). A comprehensive classification can be found in Jaeger (2005, pp. 511-515), with references to her corpus of slips produced by children aged 1;4 to 6;0.

Sustained output that is both fluent and error-free is most decidedly not the norm. During typical speech, all speakers slow or disrupt their output in a variety of ways. Such dysfluencies are multiply determined. They can, for example, reflect difficulties in lexical retrieval, increased load in sentence planning, phonological complexity, and so on. Studies of children's utterances in terms of dysfluencies include McDaniel, McKee, and Garrett (2010); Rispoli, Hadley, and Holt (2008); and Wijnen (1990).

Rate can also be used to study the production system. It is usually reported in syllables per second, which responds to the problem of phonological variation across words and phrases. Measures of *speech* rate (often used in research on atypical speech, e.g., stuttering) include dysfluencies in the sample, while measures of *articulation* rate are based on only fluent segments of speech. Here, we emphasize work with the latter because the former confounds rate effects with the many other factors that affect speech errors, hesitation phenomena, and self-corrections. Studies on articulation rate include Logan, Byrd, Mazzocchi, and Gillam (2011) and McKee, McDaniel, Garrett, Lozoraitis, and Mutterperl (2013).

Finally, priming can probe the production system. Used with adults for some time now (Bock, 1986a, 1986b), this measure has recently been exercised to study the developing production system (Anderson & Conture, 2005; Branigan, McLean, & Jones, 2005; Huttenlocher, Vasilyeva, & Shimpi, 2004; Savage, Lieven, Theakston, & Tomasello, 2003). It's a tool with significant potential and is generating important findings, as we discuss below.

## Phonological encoding and articulation

One obvious difference between child and adult speech is pronunciation. Children begin producing speech sounds around six months of age, and it often takes some years beyond that for them to master the pronunciation of their target language. Several factors drive this progression, including neuro-motor development, experience with the mapping between the perceptual experience and motor control, and the acquisition of the target language's phonological system. And, teasing these factors apart is extremely challenging. But clearly, not all of mispronunciation is due to misrepresentation. For example, children can recognize the correct pronunciation of the very words that they themselves mispronounce. This is called the *fis* phenomenon, based on a child who pronounced *fish* as [fɪs] but corrected the researcher's imitation of his pronunciation (Berko & Brown, 1960).

The role of neuro-motor development in children's articulation is supported by research that investigates children's oral motor control in relation to language production in both younger children (e.g., Alcock & Krawchyk, 2010, on 21 month olds) and older children (e.g., Goffman, Hiesler, & Chakraborty, 2006, on 4–6 year olds). The general finding is that children with better oral motor control produce more sophisticated language. Studies that compare production within children find trade-offs between articulation and structural complexity (e.g., Nelson & Bauer, 1991, for typically developing children; Masterson, 1997, for clinical populations). Utterances with more demanding phonetics are structurally simpler and vice versa. In addition to demonstrating trade-offs in the production system, these findings emphasize the importance of not judging children's syntactic knowledge solely on the basis of the structures that they produce.

Gerken (1992) attributed children's systematic non-adult pronunciations to the developing production system. Based on similarities between mispronunciations that children are aware of (i.e., the *fis* phenomenon) and adult and child speech errors, she suggested that mispronunciations result from weak points in the production system. Gerken did not clarify what she meant by such points. But an example might be the insertion of the products of lexical retrieval into structural environments. This integration requires processing at two major points, and there is evidence of considerable fall-out at both points. (Examples follow in the next two sections.) Gerken further proposed that children's production system might develop templates (similar to what we will refer to as "procedures" below) to overcome these weaknesses, which would explain the consistency of their mispronunciations.[3]

Regardless of the nature of children's systematic non-adult pronunciations, children also produce speech errors. As illustrated in (1), these errors involve the same elements (e.g., feature, phoneme, syllable) and operations on those elements (e.g., anticipation, exchange) as adult speech errors do. The examples below are from Jaeger's (2005) corpus; the speaker's age (and page number in Jaeger, 2005) is given after each example; where it's essential to the analysis, the target words are given after the speech error.

(1) Examples of children's sound-based speech errors

| | | |
|---|---|---|
| a. anticipation of phonetic feature, place of articulation in this example | | |
| *Don't touch my [mægɪt] sword* | magic | 2;10 (607) |
| b. sound exchange | | |
| *Give me my [staptʃiks]* | chopsticks | 2;9 (605) |
| c. anticipation of consonant, creating new cluster | | |
| *I want my [blɪg] blanket.* | big | 2;5 (519) |
| d. anticipation of syllable | | |
| *Why did you invite [deʲbrə] … David and Barbara over?* | | 5;4 (22) |
| e. phonologically-based word substitution | | |
| *Hey, Collection is on!* | Contraption | 2;7 (635) |

Phonological stages of children's production are mainly studied through errors in spontaneous speech. (As with adults, speech errors can also be elicited from children. See, e.g., Smith, 1990.) The most important finding from this research is that children's error patterns generally parallel those of adults. Their errors reflect the same phonological organization as in the adult. For example, a large proportion of children's and adults' slips involve a single feature (Smith, 1990; Stemberger, 1989; Poulisse, 1999), most often place of articulation. In (1a), for example, anticipation of the alveolar feature in [s] affects the preceding word. In general, syllable structure, stress pattern, and initial and final consonants matter (Aitchison & Straf, 1980; Jaeger, 1992; MacKay, 1970). Child errors, like adult errors, observe syllable position constraints for consonants (Stemberger, 1989); that is, interacting consonants are predominantly in parallel syllabic positions. In (1b), for example, the exchanged elements are both onsets. Child errors, like adult errors, rarely result in "impossible" sound sequences, such as [sf] in *sfilly face* (Stemberger, 1989). Children, like adults, tend to create consonant clusters rather than to delete them (Stemberger, 1989). This is illustrated in (1c), where the anticipation of [l] from *blanket* creates the consonant cluster [bl] in *blig*. The similarities between child and adult speech errors suggest that the production system, at the level of phonological planning, functions in much the same way for adults and children.

There are also interesting differences between children and adults. Some are attributed to the child's production process being less practiced and less automated. For example, most studies conclude that children produce more speech errors than adults do. Also, Jaeger (1992) and Stemberger (1989) have claimed that the average number of feature mismatches is lower for children than for adults. In other words, adult errors are more likely to affect whole segments than individual features. Stemberger (1989) suggested that this is because adults' segmental representations are more integrated than children's, entailing that the features of a phoneme behave more independently in children and in a more unified manner in adults. He attributed this difference to less feedback from the phonological elements to the lexical item in children than in adults. Finally, and in a related vein, Wijnen (1992) found that children produce fewer phonologically-motivated lexical substitutions (so-called malapropisms) than adults do. An example of this is (1e). Wijnen suggested that backward activation spreading (which could result in malapropisms) would require prolonged practice. We will return to this aspect of children's errors in the next section. Phonological effects also interact with functional morphology, a topic that we will take up in the syntactic encoding section.

## Lexical retrieval

We focus here on the processes that guide retrieval of specific lexical items. These provide real-time access to the several classes of information required by successive stages of utterance generation. As we will see, major features of children's lexical retrieval are adult-like. Some interesting departures reflect development of control, as well as facility with the integration of lexical and structural constraints.

As in other areas of acquisition research, children's spoken words are sometimes taken as a direct indication of their lexical knowledge, and contributions of retrieval processes may be unclear or ignored. Consider, for example, the observation that children's first words tend to be nouns. This has been attributed to learning per se: Either nouns are easier to learn because of conceptual predispositions (e.g., Gentner, 1982), or they are easier to learn because of patterns in the input (e.g., Tardif, Gelman, & Xu, 1999). Whereas learning alone is a possibility, it is complicated by the fact that children's comprehension vocabulary exceeds their production vocabulary. Another example of a link between production patterns and claims regarding lexical competence is overextension, where children use a word in apparent reference beyond its extension, such as *dog* to refer to all animals. In this case, the competence/performance distinction is not ignored; there is debate in the literature regarding whether overextension is due to immature lexical knowledge or to difficulties with lexical retrieval (e.g., Gershkoff-Stowe, Connell, & Smith, 2006; Hoek, Ingram, & Gibson, 1986; Naigles & Gelman, 1995). Research that independently investigates the development of the retrieval processes could resolve such questions.

A pioneering study of children's lexical retrieval is Dapretto and Bjork (2000), henceforth *D & B*. This study addressed the apparent vocabulary spurt at around age two, when the child is producing approximately 50 words.[4] D & B tested 30 children between 14 and 24 months, and grouped them as pre-spurt, undergoing a spurt, or post-spurt based on each child's ratio of production and comprehension vocabulary derived from parental reports. Using words from the children's spontaneous speech, D & B had each child watch an experimenter hide objects in one of two boxes. One box had pictures on it that included the objects hidden in that box, and the other box did not. Once the objects were hidden, an experimenter asked, "*What* is in this box?" Then, she turned the box with the pictures around so that the child could not see the pictures, and asked about each object, "*Where* is the X?" The three groups of children were equally successful in response to the *where* questions, showing that they could both access the words in their lexicons and remember where the objects had been hidden. But the groups differed in response to the *what* questions. The post-spurt group was more successful than the pre-spurt group; the spurt group performed in the middle. Further, children were more successful when the object was depicted on the box, but the difference between the picture and no-picture conditions was reliable only for the pre-spurt group. These findings suggest an early point when lexical retrieval is difficult and relies on perceptual cues, as proposed by Bloom (1993). D & B suggested that development in lexical retrieval accounts for other aspects of children's early speech as well, in particular the tendency for their first words to emphasize objects that are physically present. An important question about D & B's finding is *why* lexical retrieval processes change. It could be that non-lexical aspects of the production system are immature (e.g., how message-level representations synchronize with the lexicon). Or, some aspect of the early lexicon—its organization and/or the lexical representations—might not allow the usual lexical retrieval processes to function efficiently. This may relate to proposals about lack of phonetic detail in early lexical representations (e.g., Ferguson & Farwell, 1975;

see also Swingley & Aslin, 2002). Also potentially relevant is Jaeger's (2005) report of *no* phonological speech errors in the period when children produced fewer than 50 words.

Pursuing a phenomenon reported by Gerschkoff-Stowe and Smith (1997), D & B also studied naming errors. In a picture-book reading session, their spurt group made significantly more naming errors than their pre- and post-spurt groups. Most of the naming errors were categorical (i.e., using a word from the same semantic category as the intended word). D & B suggested that pre-spurt children produce too few words for competition to be a problem, and that, as children practice producing a larger number of words, the connections involved in retrieving each one strengthen, leading to a decrease in naming errors after the spurt. Gerschkoff-Stowe (2001, 2002) provided additional evidence for this account. Gerschkoff-Stowe (2001), which was longitudinal, found that during the vocabulary spurt, a large proportion of children's naming errors showed perseveration effects: The incorrect label was a word that the child had recently produced. Such interference, which diminished after the vocabulary spurt, is indicative of a real-time process.[5] Investigating the effects of practice, Gerschkoff-Stowe (2002) showed that, after children had practiced producing a word, it was more resistant to interference from competitors.

Retrieval processes might also explain some of children's word choices. For example, Clark and Johnson (1994) compared instance naming and superordinate naming (e.g., *dog* versus *animal*) in preschoolers and first graders. For both groups, and in both comprehension and production tasks, performance was better on the instance than on the superordinate items. But the largest effect was in the production task with the preschoolers. In both accuracy and reaction time data, this group performed more poorly with the superordinate category. Assuming that superordinate naming entails suppression of competing instance names, Clark and Johnson interpreted their findings as showing a development in inhibitory mechanisms. This suggestion is consistent with D & B's and Gerschkoff-Stowe's accounts of naming errors during the vocabulary spurt. Although these claims apply to different periods in development, both suggest that lexical retrieval in children is particularly susceptible to interference.

Finally, some research on naming shows that children—like adults—have an easier time retrieving words that are early-acquired, frequent in the language, and have fewer synonyms; and that lexical retrieval is aided by repetition priming (see Anderson, 2008, and citations therein.) These findings are interpreted as indicating development in the strength of lexical connections.

Data bearing on more detailed retrieval issues complements the adult/child correspondence for the general properties of lexical processing as just discussed. These properties address the design and control for access to the different classes of lexically coded information. Thus, children make speech errors that indicate a two-stage lemma/word form retrieval process, as in adults. Examples of their word substitutions from Jaeger (2005), given in (2), are instructive.

(2) Word substitution errors
   a. phonologically-based (repeated from (1e) above)
      *Hey, Collection is on!* Contraption 2;7 (635)
   b. semantically-based
      *Daddy, sit.* Mommy 1;9 (629)
   c. semantic/phonologically based
      *Thank you for the cookie.* candy 3;0 (636)

Examples (2a) and (2b) show the broad contrast between form-driven and meaning-driven retrieval errors. However, for children, the error data shows an interesting evolution in this contrast: It unfolds over the period between 2 and 5 years old reported by Jaeger. The adult data from a number of research reports clearly dissociate the two error classes; the child data does not. For adults, the majority of semantically based errors are not phonologically similar, and the errors with strong phonological similarities do not typically share significant meaning similarities (Fay & Cutler, 1977). Exceptions to these regularities are like (2c), in which both types of overlap occur. These latter occur often in Jaeger's corpus

Two linked observations regarding the child data are germane. First, young children do produce pure form errors like (2a), but form-driven errors occur in significant numbers later than do meaning-driven errors like (2b) and mixed errors like (2c). The form-based processes in younger children may not be strong enough to frequently trigger error on their own. (This may relate to the earlier discussion. For example, with a small vocabulary, the number of strongly form-similar competitors is small and purely form-driven processes have smaller scope for error influence.) Second, though errors like (2b) are a dominant type for all speakers, this is particularly so for younger children—semantic similarity is powerful. In addition, errors like (2c) are common in Jaeger's child error corpus. To some degree, this may be affected by robust environmental effects: Many of the child errors were produced in contexts with an environmental trigger for the intruding word (person, picture, immediately prior discourse, etc.). In these circumstances, form-driven processes activated by the speech environment will contribute to the strength of the error word. The result enhances the opportunity for mixed meaning-and-form errors to arise.

Another feature of word substitution errors, of all three types in (2), is a constraint on grammatical category. Target and intrusion are of the same grammatical class; exceptions to this for substitution errors are rare for adults. It is striking that the child error data displays this same constraint, and to very much the same degree. The effect in Jaeger's data is powerful even at the younger age ranges: Correspondence of grammatical class for target and error is the rule. The appropriate fit of the lexical substitution error word to their structural environments is evident for all major classes at over 90%. Note too that content and function words do not substitute for each other. Indeed, the functional vocabulary shows category correspondence for target and intrusion word at levels equal to or greater than that for content words and does so from the outset of their appearance (Jaeger, 2005, p. 236). Such data urge the view that children's lexical organization is accurately keyed to the syntactic constraints of the language production system from early on.

In summary, the trajectory suggested by research on children's lexical retrieval is as follows. When children first start producing words, lexical retrieval is challenging and perceptual cues can help (Dapretto & Bjork, 2000). But a shift occurs soon, at which point the lexical retrieval system, though susceptible to challenges, seems functionally adult-like. As Gerschkoff-Stowe (2001, 2002) pointed out, the factors that interfere with children's lexical retrieval (i.e., perseverance and semantic similarity) also affect lexical priming in adults.

## Syntactic encoding

The irony we began with is easily found in the study of syntactic development, where a focus on children's competence dominates, and related claims are often based on the frequency and type of their utterances. Most research exploring children's performance systems concerns comprehension rather than production (e.g., Clahsen & Felser, 2006; Golinkoff *et al*., 2013). In our view, understanding the sentence formulator is essential to understanding children's syntactic competence. Some familiar features of child speech may well reflect sentence-planning processes. These include the omission of functional morphology, null subjects, optional infinitives, optional subject-auxiliary inversion in English *wh*-questions, lack of (or infrequent use of) passives, and non-adult relative clause structures. Research on the developing sentence formulator points to a number of explanations for such errors.

Regarding functional morphology, we noted above that the representations are determined early in the production process, but the realization occurs much later in the detailed phonological expression of phrasal forms. Children's early speech is characterized by the omission of many functional morphemes. Both their omission and the intriguing finding that these morphemes emerge in approximately the same order across children (first reported by Brown, 1973) are generally characterized as reflecting *learning*. However, evidence suggests that such phenomena also implicate the developing production system. Several studies have shown that children comprehend functional morphemes before they reliably produce them (Gerken & McIntosh, 1993; Petretic & Tweney, 1977; Shady, 1996; Shipley, Smith, & Gleitman, 1969), and the omission of these elements in both children with Specific Language Impairment and typically developing children occurs more in longer and more complex sentences (Charest & Johnston, 2011). Further, children learning languages with rich inflectional morphology produce some functional morphemes and omit others (e.g., McKee & Emiliani, 1992). Speech error patterns also distinguish these elements. Comparing spontaneous speech errors produced by adults and 2 and 3 year olds, Wijnen (1992) found a difference in function words. While errors involving function words were under-represented in the adult data, this was not true of the child data. Wijnen suggested that young children have not yet developed the specialized retrieval procedure needed for function words. Along similar lines, McKee and Iwasaki (2001) suggested that children's omission of function morphemes might be because of their late realization (i.e., after the phonological encoding of content morphemes; see also Garrett (1975, 1980b) on accommodation errors and order of processing steps).

Speech errors involving exchanges of sounds, morphemes, and words have played a central role in the evolution of claims for adult sentence planning. Though the data for such errors produced by children is more limited, it supports several similar claims. The word exchanges in (3) are from Jaeger (2005), again with the speaker's age and Jaeger's page number. In (3a), we see an interaction across the three successive elements *nose*, *is*, and *run*. This shows lexical retrieval and phrasal planning processes operating over multiple elements within a phrase. Children's planning is adult-like in that respect. Whole phrases are being developed in parallel regarding their constituent lexical elements. Both Jaeger and Stemberger reported several such errors. Word exchanges such as (3b) support somewhat stronger claims about planning, namely, an elaboration over larger spans of the projected utterance. In (3b), the interacting elements span two adjacent phrases, and the exchanged elements correspond in grammatical category. These features of exchange errors have been interpreted in adult research as reflecting early stage syntactic planning. Children's errors of this sort indicate more adult-like planning domains in the four- to six-year range.

(3) Examples of children's word exchanges
  a. word exchange with morpheme stranding
     *Her run is nosing.* 2;7 (706)
  b. cross-phrase word exchange
     *I can go in the deep part where the head's over my water.* 5;11 (708)

The fact that the error profiles involving function words preserve their grammatical class in the intended utterance also reinforces the idea, discussed above, that it is not a lack of information about these items that leads to their initial omission. It is, instead, some problem with the integration of the forms with their local environment.

We can augment the evidence from exchange errors with some contrasts in adult and child performance for errors of anticipation and perseveration, such as those in (4).

(4) Examples of children's anticipation and perseveration errors
  a. anticipation of syllable (repeated from (1d) above)
     *Why did you invite [deʲbrə] … David and Barbara over?* 5;4 (22)
  b. anticipation of word
     *Yeah, it likes it … I like it.* 3;8 (687)
  c. perseveration of word
     *Daddy, me watching Daddy cooking … no … Mommy's cooking.* 2;4 (695)

First, perseverations substantially outnumber anticipations in children (Jaeger, 2005; Wijnen, 1992). The fact that adults make proportionally many more anticipatory errors than children suggests less advance planning for children than for adults. The content of as yet unspoken elements must be represented in the planned, upcoming speech in order to trigger error. This condition holds for adult speakers to a greater extent than for child speakers. Note that this applies to both sound and word anticipations, indicating that the conclusion holds for syntactic and associated phonological planning. A related feature of anticipations in children and adults is the relative number of *incomplete* errors such as (4a). Such errors are interrupted and corrected before the locus of the anticipated error word is reached. The anticipation in (4a) is incomplete because the repair makes it unclear whether the speaker would have produced the anticipated word *Barbara* correctly, or whether an exchange would have occurred. The set of incomplete anticipations comprises both of these possible error outcomes. Stemberger (1989) found the proportion of incomplete anticipations to be lower in children than in adults. Assuming that adults are capable of longer lookahead for their planning, their ability to detect the product of an impending exchange could lead to the interruption and correction, thus weeding out a portion of the observations, and hence the observed difference between child and adult performance.

The research considered so far shows major similarities in the structure of child and adult production processes, but it also indicates important limitations in children's ability to implement the full resources of those systems. Potential limitations of memory, speed and efficiency of lexical retrieval, and ability to integrate the different levels of the production architecture may be involved. We will now take up additional classes of data that suggest ways in which resource limitations may be reflected in children's speech.

We begin with research on fluency. Wijnen (1990) studied changes in one two year old's fluency over seven months. The child was more fluent during a period when he produced a small number of syntactic structures and less fluent as the variety of structures expanded. (Rispoli & Hadley, 2001, reported similar findings but emphasized how long the child has been producing a structure.) Such findings suggest that the emergence of greater syntactic complexity may depend on the efficiency of the production system.

Rispoli *et al*. (2008) studied stalls and revisions in the naturalistic speech of 20 children from 21 to 33 months of age. They defined stalls as dysfluencies that slow down an utterance without changing its meaning (e.g., pauses filled by elements such as *um* or *uh*), and revisions as changes of at least one meaningful unit. Over the period that Rispoli *et al*. studied, the rate of revisions increased with age, but the rate of stalls did not. Further, sentence length affected stalls but not revision: The stall rate was higher for longer sentences than for shorter sentences. Rispoli *et al*. recognized, but did not focus on, a variety of causes for stalls. They focused on revisions, which they hypothesized reflect speakers' monitoring of their own overt speech. They proposed that, because monitoring is a complex, multifaceted activity, it is absent early on. They interpreted their stall-revision differences as indicating an increasing role for monitoring processes.

Using a sentence completion task to investigate the production of participial forms in German, Clahsen, Hadler, and Weyerts (2004) found that children aged 5 to 12 years and adults showed a frequency effect for irregular forms. Speakers took longer to produce low-frequency irregular participles than high frequency ones. The adults overall showed the same pattern for regular forms, whereas the children showed an anti-frequency effect for regular forms; they took longer to produce high-frequency regular participles than low-frequency ones. A subset of the adult group with overall long latencies showed the same effect. Clahsen *et al*. suggested that both adults and children have high-frequency regular forms stored in memory, causing retrieval of the whole form to compete with a rule-based retrieval mechanism. People with slower lexical processes (e.g., children and some adults) take longer to retrieve the stored regular forms and to inhibit one of the two mechanisms.

A widely discussed phenomenon is young children's occasional omission of the subject in non-null-subject languages. Several researchers (e.g., Hyams, 1986; Hyams & Wexler, 1993; Orfitelli & Hyams, 2012; Radford, 1990; Rizzi, 1994) have argued that the phenomenon reflects children's linguistic competence. But others have proposed production-based accounts (e.g., Bloom, 1990; Valian, 1991; Valian, Hoeffner, & Aubrey, 1996). Some studies showed that English-speaking children omit subjects more when producing longer sentences. Valian and Aubrey (2005) found a trade-off between pronominal subjects and verbs in an imitation study. Inclusion rates for a lexical subject and the verb were approximately equal, whereas the inclusion rate for the verb was higher than for a pronominal subject. They suggested that limited resources combined with considerations of informativeness determine the patterns of children's subject omissions. Gerken (1991) proposed that prosody is an additional factor, pointing to independent evidence that children have difficulty with initial weak syllables. If these performance-based accounts are on the right track, it is necessary to determine how the developmental progression relates to the production model. At what point is an omitted subject "lost"? Is it omitted at the very end due to challenges involving the output, or is it omitted at an earlier level?

Another structure whose emergence has been explored in performance terms is the oblique relative clause. We focus here on pied-piped genitive structures, such as *the robber whose rope Dorothy is swinging*. McDaniel, McKee, and Bernstein (1998) found that children aged 3 to 12 tended to avoid this structure in an elicited production task, while adults tended to produce it. Both age groups tended to accept the structure in a grammaticality judgment task. Children's avoidance of this structure, together with a tendency to repeat the complementizer *that* in their attempts ("the robber that that that …") suggested an account in terms of commitment to the complementizer. Since the complementizer is possible (and frequently used) in all other relative clause structures in English, we suggested that the production system has developed a procedure for relative clauses that includes the complementizer. Adults use this procedure as well (evidenced by the struggle that the adult control group had with the genitive structures), but have an easier time altering it than children do.[6] If correct, this account suggests that the syntactic encoding for a relative clause structure can begin before all the details of that part of the message have been determined.

Bunger, Trueswell, and Papafragou (2012) compared preschoolers' and adults' production of sentences describing motion events (e.g., *The boy skated into the net*). Eye-movement data indicated that preschoolers conceptualized the event the same way that adults did. They performed as well as adults on a memory task identifying changes in path and motion. They also structured their utterances like adults did, tending to encode the manner in the verb (*skate*) and the path in a modifier phrase (*into the net*). Children's eye movements, like those of adults in this and earlier studies, corresponded to the order of the planned utterance. However, the children and adults differed with respect to how often they included both manner and path in their utterances; the adults usually included both components, whereas the children usually omitted one or the other. These findings indicate that the omissions are not attributable to differences in the conceptualization of the event, but in the mechanisms involved in sentence planning. Bunger *et al.* discussed three accounts of the child/adult difference: It might be pragmatic in nature, involving the decision about which aspects of the event should be included in the message. Or, children may be more susceptible to limiting the utterance to the aspects of the scenario that most capture their attention. Finally, the children's sentence planning system may be less efficient than the adult system in various ways, possibly requiring more attention to the aspects of the event component that will be included in the utterance.

Another interesting area of investigation uses priming, a phenomenon extensively studied in adults (e.g., Bock, 1986b; Bock & Loebell, 1990; Pickering & Branigan, 1998). Whereas research on syntactic priming in adults emphasizes priming itself and what it means for language production, studies of children's syntactic priming are generally designed to investigate syntactic competence (e.g., Bencini & Valian, 2008; Branigan *et al.*, 2005; Goldwater, Tomlinson, Echols, & Love, 2011; Huttenlocher *et al.*, 2004; Savage *et al.*, 2003; Shimpi, Gámez, Huttenlocher, & Vasilyeva, 2007). For example, Bencini and Valian's experiment on the passive indicated that three year olds know the structure and, more generally, that their representation is abstract. They also found greater use of the passive over the course of the experiment, possibly indicating some learning. As with adults, syntactic priming in children is also interesting because it reflects aspects of the production system. Anderson and Conture (2004), in a priming study of children who do and do not stutter, found that syntactic priming increased fluency. They suggested that children's sentence planning processes are inefficient and benefit from repetition, which reduces some of the burden of planning.[7] The fact that syntactic priming occurs in children is also relevant to the possibility, raised earlier, that non-adult structures consistently produced by children could be due to a sentence planning procedure. A child using such a procedure will create a self-priming effect for future utterances of the same type.

We now turn to some research on children's formulation of multi-clause structures based on dysfluency patterns (McDaniel *et al.*, 2010) and articulation rate (McKee *et al.*, 2013).

McDaniel *et al*. (2010) reported analyses of utterances collected with an elicitation protocol that was originally designed to study children's grammatical knowledge (Hamburger & Crain, 1982). The targeted sentence types, exemplified in (5), manipulated various aspects of the relative clause structure such as gap position and distance between filler and gap.

(5) a. one-clause subject gap relative clause:
*Pick up the baby that is pulling the hen*.
b. one-clause object gap relative clause:
*Pick up the bear that the king is pushing*.
c. two-clause subject gap relative clause:
*Pick up the pirate that Dorothy said was tapping the horse*.
d. two-clause object gap relative clause:
*Pick up the duck that Big Bird thinks the princess is kissing*.

Subjects in this study were 47 children and 30 adults. The Young group of children included 23 three to five year olds, and the Older group included 24 six to eight year olds. Our question in this study was not *whether* children could produce these structures (earlier work had shown that), but rather *how* they produce them. Specifically, we were interested in the patterns of dysfluencies occurring during the utterance of such structures as an indication of processing complexity. We divided the utterances into sections and categorized the dysfluencies in them. The types we discuss here are illustrated in (6), with the speaker's age after each example. Note that these types occur in adult speech as well. All dysfluencies, whether they involve revisions or not, slow speech down.

(6) Examples of dysfluencies
a. silent or unfilled pause
*The one that* [960 msec] *Big Bird thought the princess was kissing*. 7;1
b. pause filled with elements such as *um* or *uh*
*The one that Dorothy said* <u>um</u> *was tapping the horse*. 5;3
c. stop and restart
i. with repetition of stopped material
*The one that's – the one that's pulling the hen*. 6;10
ii. with repair of stopped material
*Pick up the one – pick up the bear that the evil king is pushing*. 5;11

Not surprisingly, we found that overall fluency increased with age; that is, Young children's production of the same structures was less fluent than Older children's, which was in turn less fluent than that of adults. Despite these differences, the effects of sentence type were similar across ages. Overall, the two-clause structures were less fluent than the one-clause structures, and the two-clause object items were less fluent than the two-clause subject items. This is an interesting, and to our knowledge unique, indicator for similar detailed complexity profiles in syntactic planning for child and adult speakers.

Another finding concerned the distribution of the dysfluencies. Unfilled pauses, such as the 960 milliseconds in (6a), tended to occur at clause boundaries across the three age groups. This comports with findings on spontaneous speech in adults (Beattie, 1980; Boomer, 1965; Butterworth, 1980; Ford, 1978). But filled pauses, such as the one with *um* in (6b), differed across age groups. Young children's filled pauses tended to distribute like the unfilled pauses, whereas Adults used filled pauses almost exclusively before the onset of the utterance. Older children's filled pauses occurred both in the unfilled pause loci and utterance initially. The difference in distribution of filled and unfilled pauses suggests that they mark different aspects of sentence planning (e.g., Levelt, 1983; Maclay & Osgood, 1959; Smith & Clark, 1993). We also examined restarts. As illustrated in (6c), we included both repetitions and repairs in this category. Restarts patterned differently in children and adults. Adults' (rare) restarts occurred almost exclusively within the lower clause, whereas children's (frequent) restarts occurred throughout the utterance and affected larger units.

We took our findings on these dysfluencies to indicate that the architecture of the production system for children and adults is similar with regard to planning units, but that processing resources differ. The unfilled pause loci may mark structurally determined planning units for all ages. Filled pauses, on the other hand, may indicate effects of early stage planning, as in the mapping from a message to a specific sentence's features. When adults plan difficult utterances, filled pauses indicate work done while they plan the sentence corresponding to the message. Children may do such planning at multiple points, with both message and linguistic content considered at pre-designated stopping points. This view fits the restart data as well. The child/adult discrepancies in the locations of filled pauses and restarts indicate that children do less pre-utterance planning at the message level than adults and thus must return to higher level structural constraints more frequently than adults.

McKee *et al*. (2013) explored the difficulty of producing embedded clauses using articulation rate. Recall that articulation rate is the rate (in syllables per second) of fluent speech. This study included 130 three to eight-year-old children and 24 adults. The targeted structures were sentences with relative clauses or conjoined clauses, such as those in (7). The relative clause items included structures with the relative modifying the main clause subject or object, and with the gap in subject or object position. The conjoined clause items included a pronoun in the second conjunct in subject or object position that referred to the subject or object of the first conjunct.

(7) a. relative clause: *The witch that the spider tasted covered the box*.
  b. conjoined clauses: *The boy followed the car and the dog kissed him*.

We elicited these utterances with an imitation task. The sentences given for repetition were prerecorded and played for the subjects after an enacted scenario. The purpose in the scenario was to provide the message corresponding to the repeated sentence, making the task more natural. Since our measure was articulation rate, we used only utterances with no dysfluencies, including pauses over 200 milliseconds. Filtered this way, our data included contributions from all 24 adults and from 69 of the children, ranging in age from 3;9 to 8;11. We report here on articulation rate itself in children vs. adults, and on articulation rate in different structures.

We found that children's articulation rate is slower than adults' and that articulation rate in the children was reliably correlated with age in months. This result is not surprising and corresponds to earlier reports (e.g., Logan *et al.*, 2011, and references cited therein). However, some earlier work emphasized speech rate (which confounds articulation and dysfluency) and naturalistic data (which doesn't control syntactic complexity). The question for future research is why children speak more slowly than adults when such confounding factors are eliminated or controlled. This could reflect lesser efficiency in the integration of the production system's levels, maturity of the motor system, and/or working memory limitations.

We also compared articulation rate across relative clause and coordinate sentences overall, as well as the first and second conjuncts of the conjoined sentences, and the main clause and relative clause of the relative clause sentences. The rate patterns were the same for adult and six to eight-year-old speakers. There was no overall difference in rate between conjoined and relative clause sentences, but the relative clause was spoken faster than the main clause, whereas the two conjuncts of the conjoined structure were spoken at the same rate. The pattern for three to five year olds was different. In this group, the overall articulation rate for the relative clause sentences was slower than for the conjoined sentences, and there was no acceleration on the relative clause part.

We suggested that adults adjust rate of articulation as a function of planning pressures. Relative clauses are structurally complex, but they function to modify an element of the main clause and consist of presupposed material. The focus of the message is the main clause. It is thus plausible that a speaker would accelerate the relative clause in order to keep this focus. The fact that younger children did not show this pattern suggests that their production system is not yet adept at making such adjustments. This kind of finding should figure into accounts of children's difficulties with relative clauses and other structures.

In summary, with the possible exception of functional morphology early on, the sentence formulation process in children appears to be adult-like. However, limited resources affect children's output in various ways. These findings, as well as future research on the production system per se, should help to identify the contributions of competence and performance to the non-adult syntactic structures that are characteristic of children's speech.

## Conclusion

Research on the developing production system fills a significant gap in the field of language acquisition. Better understanding of this system helps us factor out its effects in children's utterances, which are an important type of data for the field. Phenomena that remain after such filtering can credibly serve as evidence for accounts of linguistic competence. Further, research on children's production systems could shed light on the nature of speech and language disorders that affect language output.

The developing production system is of intrinsic interest. It is an important piece of the language acquisition process, one that interacts with the developing comprehension system, the grammar, the lexicon, and other aspects of cognition. It is also true that the theoretical framework described above could be improved/expanded by attention to its development. Specific models along the lines of that framework have been tested with different types of data (e.g., aphasic speakers, speech errors from normally functioning adults, computational modeling), and modifications have been made on the basis of such investigations. But little attention has gone to the question of how this performance system comes into place. We might ask, for example, how this system operates when speakers have minimal knowledge of their target language. In the sister domain of parsing, Fodor (1998) argued for an innate parser that applies across languages, differing only in its response to differences in grammar and lexicon. The research we have reviewed suggests that a similar account might apply to the production system. If this line of reasoning is correct, then one of the child's tasks would be to map the specifics of the target language onto the universal production system. For example, the syntactic encoding process would differ for head-final and head-initial languages. The rest would consist of practice and development in other neurological and cognitive areas.

What seems clear at this point is that the production system is architecturally adult-like early on, but less efficient. The same pressures that create production challenges for adults, such as lexical retrieval of semantically similar items and certain syntactic structures, can be disruptive for children. Children are less able to attend to all the parts of the system, as evidenced by trade-offs, and are less able to plan in advance.

## Acknowledgment

Support for the authors' research includes National Science Foundation grants BCS-0822558 and BCS-0822457.

## NOTES

1 The term "lexeme" is also used to refer to word form.
2 The term "slip of the tongue" has the virtue of the signing counterpart "slip of the hand." It is important to acknowledge that language production occurs in multiple modalities.
3 Gerken (1992) distinguished production templates from the meta-rules in other accounts (e.g., Menn, 1978). Though meta-rules also simplify the output, they add a layer of complexity to the architecture and seem to lack a counterpart in the adult system. Note that Gerken's account, as well as those based on meta-rules, though focused on phonology, also apply to non-adult aspects of morphology, syntax, and vocabulary.

4 It is widely assumed that all children undergo a vocabulary spurt. See Bloom (2000) and Ganger and Brent (2004) for challenges to this assumption.
5 This is not to say that naming errors can't also reflect a lack of lexical knowledge. In fact, Gerschkoff-Stowe (2001) included objects whose labels the children were unfamiliar with, and found that the children used semantically related known words to label those too.
6 The commitment account is language-specific; it is predicted to hold only in a language that prefers the complementizer as a relativizer. McDaniel and Lech (2003) studied the production of oblique relative clauses in Polish, a language that generally uses a *wh*-word rather than a complementizer as the relativizer. As predicted, the Polish-speaking children were more successful at producing oblique relative clauses than English-speaking children.
7 Similarly, McDaniel, McKee, and Garrett (2011) reported fewer dysfluencies and more adult-like dysfluencies in children's utterances in an imitation task than in a parallel elicited production task that had no prompt.

## REFERENCES

Aitchison, J. & M. Straf (1980). Lexical storage and retrieval: A developing skill? In A. Cutler (Ed.) *Slips of the tongue and language production*, 751–795. The Hague: Mouton. Reprinted in *Linguistics* 2009. doi: 10.1515/ling.1981.19.7-8.751

Alcock, K. & K. Krawczyk (2010). Individual differences in language development: relationship with motor skill at 21 months. *Developmental Science*, 13, 677–691. doi: 10.1111/j.1467-7687.2009.00924.x

Anderson, J. (2008). Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter. *Journal of Fluency Disorders*, 33, 135–155. doi: 10.1016/j.jfludis.2008.04.001

Anderson, J. & E. Conture (2005). Sentence structure priming in young children who do and do not stutter. *Journal of Speech, Language, and Hearing Research*, 47, 552–571. doi: 10.1044/1092-4388(2004/043)

Beattie, G. (1980). The role of language production processes in the organization of behavior in face to face interaction. In B. Butterworth (Ed.) *Language production: Vol. 1. Speech and talk*, 69–107. London: Academic.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, 6, 183–200. doi: 10.1017/S0305000900002245

Berko, J. & R. Brown (1960). Psycholinguistic research methods. In P. H. Mussen (Ed.) *Handbook of research methods in child development*, 517–557. New York: Wiley.

Bencini, G. & V. Valian (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, 59, 97–113. doi: 10.1016/j.jml.2007.12.007

Bloom, L. (1993). *The transition from infancy to language*. Cambridge: Cambridge University Press.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, 21, 491–504. http://www.jstor.org/stable/4178692

Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.

Bloom, L., P. Miller, & L. Hood (1975). Variation and reduction as aspects of competence in language development. In A. Pick (Ed.) *Minnesota symposium on* *child psychology (Vol. 9)*. Minneapolis: University of Minnesota Press.

Bock, K. (1986a). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Language, Memory, and Cognition*, 12, 575–586. doi: 10.1037/0278-7393.12.4.575

Bock, K. (1986b). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387. doi: 10.1016/0010-0285(86)90004-6

Bock, K. (1987). Coordinating words and syntax in speech plans. In A. Ellis (Ed.) *Progress in the psychology of language*, 337–390. London: Erlbaum.

Bock K. & H. Loebell (1990). Framing sentences. *Cognition*, 35, 1–39. doi: 10.1016/0010-0277(90)90035-I

Boomer, D. (1965). Hesitation and grammatical encoding. *Language and Speech*, 8, 145–158.

Bornstein, M. & C. Hendricks (2012). Basic language comprehension and production in >100,000 young children from sixteen developing nations. *Journal of Child Language*, 39, 899–918. doi: 10.1017/S0305000911000407

Boyle, M.K. & L. Gerken (1997). The influence of lexical familiarity on children's function morpheme omissions: A nonmetrical effect? *Journal of Memory and Language*, 36, 117–128. doi: 10.1006/jmla.1996.2478

Branigan, H., J. McLean, & M. Jones (2005). A blue cat or a cat that is blue? Evidence for abstract syntax in young children's noun phrases. *The Proceedings of the 29th Annual Boston University Conference on Language Development*, ed. by A. Brugos, M. R. Clark-Cotton, & S. Ha, 109–121. Somerville, MA: Cascadilla Press. http://hdl.handle.net/10373/2076

Brown, R. (1973). *A First Language: The Early Stages*. Oxford, England: Harvard University Press.

Bunger, A., J. Trueswell, & A. Papafragou (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition*, 122, 135–149. doi: 10.1016/j.cognition.2011.10.002

Butterworth, B. (1980). Some constraints on models of language production. In B. Butterworth (Ed.) *Language production: Vol. 1. Speech and talk*, 423–459. London: Academic.

Charest, M. & J. Johnston (2011). Processing load in children's language production: A clinically oriented review of research. *Canadian Journal of Speech-Language Pathology and Audiology*, 35, 18–31.

Clahsen, H. & C. Felser (2006). Grammatical processing of spoken language in child and adult language learners. *Journal of Psycholinguistic Research*, 38, 305–319. doi: 10.1017/S0142716406060024

Clahsen, H., M. Hadler, & H. Weyerts (2004). Speeded production of inflected words in children and adults. *Journal of Child Language*, 31, 683–712. doi: 10.1017/S0305000904006506

Clark, J. & C. Johnson (1994). Retrieval mechanisms in the development of instance and superordinate naming of pictures. *Journal of Experimental Child Psychology*, 57, 295–326. doi: 10.1006/jecp.1994.1015

Dapretto, M. & E. Bjork (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development*, 71, 635–648. http://www.jstor.org/stable/1132382

Dell, G. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321. doi: 10.1037/0033-295X.93.3.283

Fay, D. & A. Cutler (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, 8, 505–520. http://www.jstor.org/stable/4177997

Ferguson, C.A. & C.B. Farwell (1975). Words and sounds in early language acquisition. *Language*, 51, 419–439. http://www.jstor.org/stable/412864

Fodor, J.D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27, 285–319. doi: 10.1023/A:1023258301588

Ford, M. (1978). Planning units and syntax in sentence production. Doctoral dissertation, University of Melbourne.

Fromkin, V. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27–52. http://www.jstor.org/stable/412187

Ganger, J. & M. Brent (2004). Re-examining the vocabulary spurt. *Developmental Psychology*, 40, 621–632. doi: 10.1037/0012-1649.40.4.621

Garrett, M.F. (1975). The analysis of sentence production. In G. Bower (Ed.) *Psychology of learning and motivation* (Vol. 9), 133–177. New York: Academic Press.

Garrett, M.F. (1980a). Levels of processing in sentence production. In B. Butterworth (Ed.) *Language production: Vol. 1. Speech and talk*, 177–220. London: Academic.

Garrett, M.F. (1980b). The limits of accommodation: Arguments for independent processing levels in sentence production. In V. Fromkin (Ed.) *Slips of Tongue, Ear and Hand: Errors in Linguistic Performance*, 263–371. New York: Academic Press.

Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S.A. Kuczaj (Ed.) *Language development. Vol. II, Language, thought, and culture*, 301–334. Hillsdale, NJ: Lawrence Erlbaum.

Gerken L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, 30, 431–451. doi: 10.1016/0749-596X(91)90015-C

Gerken, L. (1992). A slip-of-the-tongue approach to language development. Unpublished manuscript. SUNY.

Gerken, L. & B. McIntosh (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, 29, 448–457. doi: 10.1037/0012-1649.29.3.448

Gerschkoff-Stowe, L. (2001). The course of children's naming errors in early word learning. *Journal of Cognition and Development*, 2, 131–155. doi: 10.1207/S15327647JCD0202_2

Gershkoff-Stowe, L. (2002). Object naming, vocabulary growth, and the development of word retrieval abilities. *Journal of Memory and Language*, 46, 665–687. doi: 10.1006/jmla.2001.2830

Gershkoff-Stowe, L., B. Connell, & L. Smith (2006). Priming overgeneralizations in two- and four-year-old children. *Journal of Child Language*, 33, 461–486. doi: 10.1017/S0305000906007562

Gershkoff-Stowe, L. & L. Smith (1997). A curvilinear trend in naming errors as a function of early vocabulary growth. *Cognitive Psychology*, 34, 37–71. doi: 10.1006/cogp.1997.0664

Goffman, L., L. Hiesler, & R. Chakraborty (2006). Mapping of prosodic structure onto words and phrases in children's and adults' speech production. *Language and Cognitive Processes*, 21, 25–47. doi: 10.1080/01690960400001820

Goldwater, M., M. Tomlinson, C. Echols, & B. Love (2011). Structural priming as structure-mapping: Children use analogies from previous utterances to guide sentence production. *Cognitive Science*, 35, 156–170. doi: 10.1111/j.1551-6709.2010.01150.x

Golinkoff, R., W. Ma, L. Song, & K. Hirsh-Pasek (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on Psychological Science*, 8, 316–339. doi: 10.1177/1745691613484936

Hamburger, H. & S. Crain (1982). Relative acquisition. In S.A. Kuczaj (Ed.) *Language development. Vol. I, Syntax and semantics*, 245–274. Hillsdale, NJ: Lawrence Erlbaum.

Hoek, D., D. Ingram, & D. Gibson (1986). Some possible causes of children's early word overextensions. *Journal of Child Language*, 13, 477–494. doi: 10.1017/S0305000900006838

Huttenlocher, J. M. Vasilyeva, & P. Shimpi (2004). Syntactic priming in young children. *Journal of Memory and Language*, 50, 182–195. doi: 10.1016/j.jml.2003.09.003

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.

Hyams, N. & K. Wexler (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry*, 24, 421–459. http://www.jstor.org/stable/4178822

Jaeger, J. (1992). 'Not by the chair of my hinny hin hin': Some general properties of slips of the tongue in young children. *Journal of Child Language*, 19, 335–366. doi: 10.1017/S0305000900011442

Jaeger, J. (2005). *Kids' Slips: What young children's slip of the tongue reveal about language development*. Mahwah, NJ: Lawrence Erlbaum.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41–104. doi: 10.1016/0010-0277(83)90026-4

Levelt, W.J.M. (1989). *Speaking. From intention to articulation*. Cambridge, MA: MIT.

Logan, K.J., C.T. Byrd, E.M. Mazzocchi, & R.B. Gillam (2011). Speaking rate characteristics of elementary-school-aged children who do and do not stutter. *Journal of Communication Disorders*, 44, 130–147. doi: 10.1016/j.jcomdis.2010.08.001

MacKay, D. (1970). Spoonerisms of children. *Neuropsychologia*, 8, 315–322. doi: 10.1016/0028-3932(70)90077-1

Maclay, H. & C. Osgood (1959). Hesitation phenomena in spontaneous English speech. *Word*, 15, 19–44. doi: 10.1080/00437956.1959.11659682

Marshall, J.C. (1979). Language acquisition in a biological frame of reference. In P. Fletcher & M. Garman (Eds.) *Language acquisition*, 437–453. Cambridge: Cambridge University Press.

Masterson, J.J. (1997). Interrelationships in children's language production. *Topics in Language Disorders*, 17, 11–22.

Masterson, J.J. & A.G. Kamhi (1992). Linguistic trade-offs in school-age children with and without language disorders. *Journal of Speech and Hearing Research*, 35, 1064–1075. doi: 10.1044/jshr.3505.1064

McDaniel, D. & D. Lech (2003). The production system's formulation of relative clause structures: Evidence from Polish. *Language Acquisition*, 11, 63–97. doi: 10.1207/s15327817la1102_1

McDaniel, D., C. McKee, & J. Bernstein (1998). How children's relatives solve a problem for minimalism. *Language*, 74, 308–334. http://www.jstor.org/stable/417869

McDaniel, D., C. McKee, & M.F. Garrett (2010). Children's sentence planning: Syntactic correlates of fluency variations. *Journal of Child Language*, 37, 59–94. doi: 10.1017/S0305000909009507

McDaniel, D., C. McKee, & M.F. Garrett (2011). Fluency markers for children's sentence planning: Early and late stage processing. In N. Danis, K. Mesh, & H. Sung (Eds.) *Online Supplement to Proceedings of the 35th Boston University Conference on Language Development*.

McKee, C. & M. Emiliani (1992). Il clitico: C'è ma non si vede (The clitic: It's there but you don't see it). *Natural Language and Linguistic Theory*, 10, 415–437. http://www.jstor.org/stable/4047774

McKee, C. & N. Iwasaki (2001). Lemma structure in language learning: Comments on representation and realization. In J. Weissenborn & B. Hoehle (Eds.), *Approaches to bootstrapping: Phonological, syntactic, and neurophysiological aspects of early language acquisition (From signal to syntax, Volume I)*, 125–144. Amsterdam-Philadelphia: John Benjamins.

McKee, C., M. Rispoli, D. McDaniel, & M.F. Garrett (2006). How do children become adult sentence producers? *Journal of Applied Psycholinguistics*, 27, 74–81. doi: 10.1017/S014271640631003X

McKee, C., D. McDaniel, M.F. Garrett, C. Lozoraitis, & M. Mutterperl (2013). Articulation rate: Effects of age, fluency, and syntactic structure. *Revista de Logopedia, Foniatría, y Audiología*, 33, 55–63. doi: 10.1016/j.rlfa.2013.04.001

Menn, L. (1978). Phonological units in beginning speech. In A. Bell & J.B. Hooper (Eds.) *Syllables and segments*, 157–171. Amsterdam: North Holland.

Meringer, R. (1908). *Aus dem Leben der Sprache: Versprechen, Kindersprache, Nachahmungstrieb*. Berlin: Behr's Verlag.

Naigles, L. & S. Gelman (1995). Overextensions in comprehension and production revisited: Preferential-looking in a study of dog, cat, and cow. *Journal of Child Language*, 22, 371–380. doi: 10.1017/S0305000900009612

Nelson, L. & H. Bauer (1991). Speech and language production at age 2: Evidence for tradeoffs between linguistic and phonetic processing. *Journal of Speech and Hearing Research*, 34, 879–892. doi: 10.1044/jshr.3404.879

Orfitelli, R. & N. Hyams, N. (2012). Children's grammar of null subjects: Evidence from comprehension. *Linguistic Inquiry*, 43, 563–590. doi: 10.1162/ling_a_00106

Petretic, P.A. & R.D. Tweney (1977). Does comprehension precede production? The development of children's responses to telegraphic sentences of varying grammatical adequacy. *Journal of Child Language*, 4, 201–209. doi: 10.1017/S0305000900001604

Pickering, M. & H. Branigan (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39, 633–651. doi: 10.1006/jmla.1998.2592

Poulisse, N. (1999). Slips of the tongue: Speech errors in first and second language production. Amsterdam: John Benjamins.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Oxford: Blackwell.

Rispoli, M. & P. Hadley (2001). The leading edge: The significance of sentence disruptions in the development of grammar. *Journal of Speech, Language, and Hearing Research*, 44, 1131–1143. doi: 10.1044/1092-4388(2001/089)

Rispoli, M., P. Hadley, & J. Holt (2008). Stalls and revisions: A developmental perspective on sentence production. *Journal of Speech, Language, and Hearing Research*, 51, 953–966. doi: 10.1044/1092-4388(2008/070)

Rizzi, L. (1994). Early null subjects and root null subjects. In B. Lust, G. Hermon, & J. Kornfilt (Eds.) *Syntactic theory and first language acquisition: Cross-linguistic perspectives. Vol. 2: Binding, dependencies, and learnability*, 249–272. Hillsdale, NJ: Erlbaum.

Savage, C., E. Lieven, A. Theakston, & M. Tomasello (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, 6, 557–567. doi: 10.1111/1467-7687.00312

Shady, M.E. (1996). Infants' sensitivity to function morphemes. Doctoral dissertation, SUNY Buffalo.

Shimpi, P., P. Gámez, J. Huttenlocher, & M. Vasilyeva (2007). Syntactic priming in 3- and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology*, 43, 1334–1346. doi: 10.1037/0012-1649.43.6.1334

Shipley, E., C. Smith, & L. Gleitman (1969). A study in the acquisition of language: Free responses to commands. *Language*, 45, 322–342. http://www.jstor.org/stable/411663

Smith, B.L. (1990). Elicitation of slips of the tongue from young children: A new method and preliminary observations. *Applied Psycholinguistics*, 11, 131–144. doi: 10.1017/S0142716400008730

Smith, V.L. & Clark, H.H. (1993). On the course of answering questions. *Journal of Memory and Language*, 32, 25–38. doi: 10.1006/jmla.1993.1002

Stemberger, J. (1989). Speech errors in early child language production. *Journal of Memory and Language*, 28, 164–188. doi: 10.1016/0749-596X(89)90042-9
Swingley, D. & R.N. Aslin (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13, 480–484. doi: 10.1111/1467-9280.00485
Tardif, T., S. Gelman, & F. Xu (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development*, 70, 620–635. doi: 10.1111/1467-8624.00045
Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, 40, 21–81. doi: 10.1016/0010-0277(91)90046-7
Valian, V., J. Hoeffner, & S. Aubrey (1996). Young children's imitation of sentence subjects: Evidence of processing limitations. *Developmental Psychology*, 32, 153–164. doi: 10.1037/0012-1649.32.1.153
Valian, V. & S. Aubrey (2005). When opportunity knocks twice: two-year-olds' repetition of sentence subjects. *Journal of Child Language*, 32, 617–641. doi: 10.1017/S0305000905006987
Wijnen, F. (1990). The development of sentence planning. *Journal of Child Language*, 17, 651–675. doi: 10.1017/S030500090001093X
Wijnen, F. (1992). Incidental word and sound errors in young speakers. *Journal of Memory and Language*, 31, 734–755. doi: 10.1016/0749-596X(92)90037-X

# 23 Language Comprehension in Monolingual and Bilingual Children

KRISTA BYERS-HEINLEIN[1] AND
CASEY LEW-WILLIAMS[2]

[1] Concordia University
[2] Princeton University

## Introduction

Long before they speak their first words, children begin to understand the language that they hear around them. Indeed, language comprehension—extracting meaning from speech—outpaces language production throughout development (Fenson *et al.*, 2007). Nevertheless, understanding speech is a challenging and multi-faceted task. At minimum, children must identify and perceive speech sounds, parse the speech stream into its constituent words, identify the meaning of these words, consider their order in the context of a language's grammar, and link the entire message to the speakers' intended meaning (see Figure 23.1). All of this happens quickly and in real time: even "slow" infant-directed speech occurs at a rate of several syllables per second (Fernald & Simon, 1984). When and how do children come to understand what is spoken to them?

Recent research has shown that infants extract meaning from speech much earlier than previously thought. By age six to nine months, infants understand the meanings of many common words like *feet*, *juice*, and *spoon* (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 2011). However, the path to language comprehension neither begins nor ends there. Throughout development, successful language comprehension intertwines with developing linguistic, cognitive, and social abilities.

The development of language comprehension depends on the specific nature of children's language environments. Monolingual children hear one of the world's many languages, bidialectal children hear two varieties of the same language, and

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

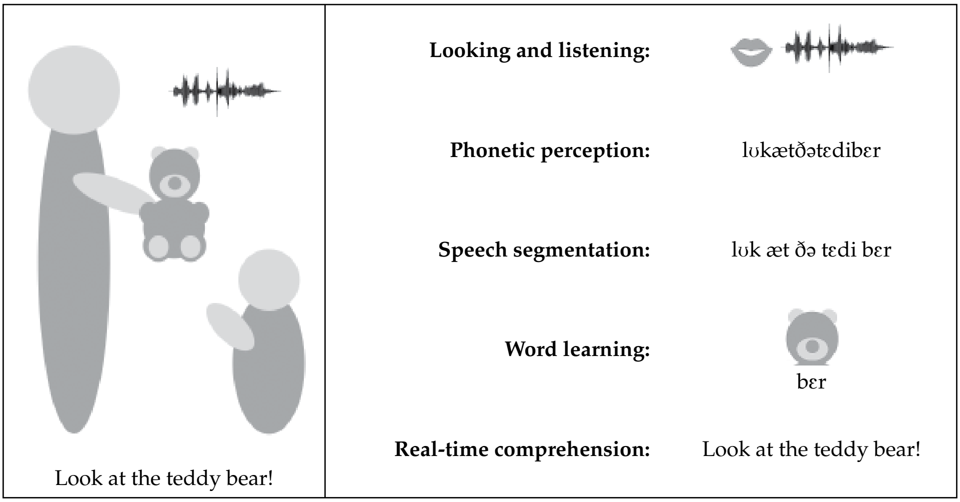



**Figure 23.1** Developing language skills such as looking and listening, phonetic perception, speech segmentation, and word learning together contribute to real-time language comprehension in the first few years of life.

bilingual and multilingual children hear two or more languages. Diversity in language experiences can take many other forms as well. Children growing up in poverty often have fewer opportunities to hear words and sentences relative to children growing up with more resources (Hart & Risley, 1995; Weisleder & Fernald, 2013). Some children are exposed to signed rather than to spoken languages. Children with cochlear implants hear speech that is degraded relative to children with acoustic hearing (see Pisoni, this volume; Grieco-Calub, Saffran, & Litovsky, 2009). And relative to typically developing children, children with developmental language disorders experience a complex interaction between cognition and language input (Rice, Warren, & Betz, 2005). All children must adapt to the specific challenges presented by their environments. Children whose experience reduces the quantity and quality of language exposure are often slower in language acquisition. Other children, such as bilinguals or sign language learners, develop language differently but are not delayed (see Kegl, this volume; Peña, Gillam, Bedore, & Bohman, 2011; Petitto *et al.*, 2001; Werker & Byers-Heinlein, 2008).

In this chapter, we focus on cross-linguistic research with monolingual and bilingual infants and toddlers to explore how they navigate the path from hearing to understanding. While most research to date has investigated monolingual children, there is growing interest in understanding language acquisition in the many children around the world who encounter multiple languages early in life. Note that while we use the blanket term "bilingual" to refer to children acquiring two or more languages, this is anything but a one-size-fits-all category. There are vast differences in the quality and quantity of language experiences across different households and populations, the timing of exposure to different languages, as well as many other demographic, cultural, and linguistic differences (McCabe *et al.*, 2013). Considering both monolingual and bilingual learners enriches what we know about language learning in general.

Each section of the chapter begins with an overview of research on monolingual children and then discusses relevant findings from research on bilingual children. We start by describing how infants' looking and listening facilitate their entry into language, and we then discuss phonetic development, speech segmentation, word learning, and real-time language processing. We conclude with a section on how monolingual and bilingual infants learn from the imperfect speech that is inherent in the complexities of natural language environments.

## Looking and listening

Language acquisition begins with looking at and listening to native speakers of the ambient language(s). From very early in life, infants attend to speech over other types of sounds (Vouloumanos & Werker, 2007), and quickly target their attention to the native language or languages (Byers-Heinlein, Burns, & Werker, 2010; Moon, Cooper, & Fifer, 1993). Infants' preference for language is not limited to the spoken modality: six-month-old hearing infants with no exposure to sign language look more at linguistic signs than non-linguistic gestures (Krentz & Corina, 2008). Young infants may also be sensitive to the notion that language can convey information between speakers. For example, 12 month olds understand that speech, but not other types of vocalizations such as coughing, can communicate information to a listener (Martin, Onishi, & Vouloumanos, 2012). However, non-speech signals, such as tones, can quickly approximate the special status if infants witness them being used to communicate in a natural dialogue (Ferguson & Lew-Williams, 2016).

For infants growing up bilingual or multilingual, it is not enough to simply attend to their languages in an undifferentiated fashion. Instead, they must acquire each as an independent communicative system, which hinges on an ability to detect the differences between languages (Byers-Heinlein, 2014b). While fully disentangling their two languages might be a somewhat gradual process (Byers-Heinlein, 2014b), there is evidence that the ability to differentiate two languages emerges early in life. At birth, monolingual and bilingual infants can discriminate between languages that differ in rhythm, such as English and French (Byers-Heinlein, Burns, & Werker, 2010; Mehler *et al.*, 1988; Nazzi, Bertoncini, & Mehler, 1998). By age four to five months, monolinguals and bilinguals can also discriminate between rhythmically similar languages that belong to the same category as their own native language (Bosch & Sebastián-Gallés, 2001; Molnar, Gervain, & Carreiras, 2014; Nazzi, Jusczyk, & Johnson, 2000). Infants can also distinguish languages using visual cues available on the lips and face of their interlocutors. Both English-monolingual and French-English bilingual four- and six-month-old infants can tell apart visual English and French when they see talking faces with the sound turned off. However, only bilingual infants retain this sensitivity at eight months (Weikum *et al.*, 2007). Eight-month-old bilinguals also show enhanced abilities to visually discriminate unfamiliar languages (Sebastián-Gallés, Albareda-Castellot, Weikum, & Werker, 2012). Such sensitivities could help bilinguals extract meaning in their complex linguistic environments. Thus, infants use a range of auditory and visual cues to break into multiple languages, laying the foundation for discovering the sounds and sequences of sounds that comprise each language (Byers-Heinlein, Morin-Lessard, & Lew-Williams, in press).

## Phonetic development

Words are built from sounds, and languages vary in terms of which sound differences are meaningful. These meaningful differences group speech sounds into phonetic categories. For example, the phonetic difference between/r/and/l/is meaningful in English, as in the words *rake* and *lake*. This difference is not meaningful in Japanese, and so Japanese speakers tend to ignore it and group/r/and/l/ into the same phonetic category. Infants cannot know at birth whether they will be growing up in an English, Japanese, or bilingual English-Japanese environment. As such, newborn infants are sensitive to most sound differences that are meaningful across the worlds' languages. Important development occurs within the first year of life, when monolinguals lose sensitivity to non-native sound distinctions (Werker & Tees, 1984), but gain sensitivity to native language distinctions (Kuhl *et al.*, 2007). This developmental pattern is often referred to as *perceptual narrowing*, and is thought to be driven in part by infants' innate sensitivity to distributional regularities available in the input (Maye, Werker, & Gerken, 2002; Thiessen & Pavlik, 2013). Developing phonetic categories, together with growing knowledge of native language words, help children interpret whether a speech sound difference is meaningful or not (Dietrich, Swingley, & Werker, 2007).

Evidence from bilingual infants suggests that early language experience can affect phonetic development in unexpected ways (Byers-Heinlein & Fennell, 2014). For example, Spanish-Catalan bilingual infants show a U-shaped developmental pattern for their perception of vowels that exist only in Catalan (/e/-/ɛ/): they readily discriminate this phonetic difference at 4 and 12 months of age, but sometimes fail to do so at eight months of age (Bosch & Sebastián-Gallés, 2003). Monolingual Catalan-learning infants discriminate the same phonetic difference throughout the first year of life. While there are numerous different explanations for this finding (Byers-Heinlein & Fennell, 2014), one explanation focuses on how Spanish and Catalan link sound to meaning. Spanish and Catalan are both Romance languages with many cognates, which have similar meanings and differ on only a few sounds, which are often vowels (for example Catalan *pilota* and Spanish *pelota*, both meaning "ball"). Bilinguals acquiring these close languages may learn to ignore some vowel variability, and to focus on the invariant consonants (Sebastián-Gallés & Bosch, 2009). Studies with populations of bilinguals learning languages that are not closely related, such as French-English and Spanish-English bilinguals, have found patterns of phonetic development that are similar to those of monolinguals (Burns, Yoshida, Hill, & Werker, 2007; Sundara & Scutellaro, 2010; Sundara, Polka, & Molnar, 2008). More research with bilinguals is needed to investigate a wider variety of phonetic contrasts and language pairs.

Mature language-specific phonetic categories can help infants interpret meaning in speech, but there is also evidence that consistent links between sound and meaning can actually help infants interpret speech sounds. In laboratory studies, infants who hear two sounds consistently paired with two different objects are more likely to discriminate these sounds than infants who hear the sounds paired randomly with the objects (Yeung & Werker, 2009; Yeung, Chen, & Werker, 2013).

## Finding words in the speech stream

While infants are learning about the sounds of their native language(s), they also begin learning which sounds go together to form words (see Levine, Strother-Garcia, Hirsh-Pasek, & Michnick Golinkoff, this volume). Spaces signal word boundaries in written language, but silent pauses are not reliable cues to word boundaries in spoken language as they often occur in the middle of words. Children do sometimes hear words in isolation or at the edge of an utterance, and these words are relatively easy for them to pick out of the speech stream (Brent & Siskind, 2001; Johnson, Seidl, & Tyler, 2014; Lew-Williams, Pelucchi, & Saffran, 2011; Shukla, Nespor, & Mehler, 2007). However, most words occur in the middle of utterances, and infants must locate these words in order to eventually learn word meanings and interpret word combinations.

Infants can recognize familiar sound combinations in running speech—word forms—during the middle of their first year (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005). In a typical study, infants hear a list of familiar words, and later hear passages that either do or do not contain those words. Infants prefer listening to passages with the familiar list of words (Houston & Jusczyk, 2003), and are not fooled by similar-sounding words (Jusczyk & Aslin, 1995). This demonstrates their ability to segment the speech stream. Infants are especially skilled at segmenting words from familiar talkers and languages, particularly when words adhere to patterns typical of the native language (Brent & Cartwright, 1996; Houston & Jusczyk, 2000; Jusczyk, Houston, & Newsome, 1999; Polka & Sundara, 2012). Bilingual infants can flexibly and efficiently recognize word forms in each of their two native languages (Polka & Sundara, 2003; Vihman, Thierry, Lum, Keren-Portnoy, & Martin, 2007).

How do infants locate word forms in the speech stream? A learning mechanism known as *statistical learning* allows infants to track sounds and syllables that occur together with the most consistency (see Romberg & Saffran, 2010). The central idea is as follows: sounds that occur together often in a language (such as b-a-b-y in English) are likely to be words, and sounds that rarely occur together (b-a-g-u) are less likely to be words. After even brief opportunities to learn, eight month olds can detect words hidden in artificially constructed languages (Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996) and in carefully controlled passages of natural but unfamiliar languages (Lew-Williams *et al*., 2011; Pelucchi, Hay, & Saffran, 2009). Some sources of variation, such as varying word lengths, can make statistical learning more difficult (Johnson & Tyler, 2010).

Over time statistical learning begins to interact with children's growing native language experience (Graf Estes & Bowen, 2013; Lew-Williams & Saffran, 2012). For example, infants sometimes use language-specific cues (e.g., in English, paying attention to the stressed syllables that often occur at the beginning of words) rather than statistical cues when segmenting speech (Johnson & Jusczyk, 2001; Johnson & Seidl, 2009; Thiessen & Saffran, 2003). Moreover, there is a coupling between language input and the dynamic nature of caregiver-child interaction: tactile cues from adults (Seidl, Tincoff, Baker, & Cristià, 2014), and highly familiar word forms such as the child's own name also aid in segmenting the speech stream (Bortfeld *et al.*, 2005; Mersad & Nazzi, 2012). To date, most research on speech segmentation has studied monolingual infants, although recent research suggests that bilingual infants outperform monolingual infants in tracking regularities embedded in two interleaved artificial speech streams (Antovich & Graf Estes, in press). Future research will need to investigate the complexities of segmentation in bilingual contexts.

## Word learning

Once children locate a word in the speech stream, how do they figure out its intended meaning? Despite the potential difficulty of this task, children are powerful word learners, deploying a myriad of cognitive, linguistic, and social resources (Hollich, Hirsh-Pasek, & Golinkoff, 2000). As they gradually gain familiarity with common sequences of sounds, they begin to link those sequences to meaning (Graf Estes, Evans, Alibali, & Saffran, 2007). Children can sometimes infer the basic meaning of a word from a single example, a process called *fast mapping* (Carey & Bartlett, 1978). These processes set the stage for more protracted learning of a word's full meaning (Horst & Samuelson, 2008; Swingley, 2010). The rest of this section will discuss some of the many contributors to successful word learning, as well as the ultimate outcome of this learning: a child's vocabulary.

### *Associative learning mechanisms*

The ability to form associations between words and their referents is foundational to mature word learning. One year olds can successfully associate a picture of an object with a repeated word (Mackenzie, Curtin, & Graham, 2012b; 2012a; Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Even six month olds can do so if given appropriate prosodic information (Shukla, White, & Aslin, 2011). Associative word learning abilities are robust regardless of whether children are growing up monolingual or bilingual (Byers-Heinlein, Fennell, & Werker, 2012). Young infants can also associate words and objects in more challenging conditions. Even when the same word is paired with several pictures, or when the same picture is paired with several words, infants are able to track which words and pictures co-occur most reliably (Smith & Yu, 2008; Vouloumanos & Werker, 2009).

A number of different perceptual and attentional cues can support the formation of these word-object links. For example, English learners tend to learn concrete words such as nouns before other types of words (Bergelson & Swingley, 2013; Gentner, 1982), suggesting that some types of words are easier to learn than others. Low-level information can also affect how easily infants learn a new word. Infants are better at forming associations if an object is labeled synchronously with its motion (Gogate & Bahrick, 2001; Matatyaho-Bullaro, Gogate, Mason, Cadavid, & Abdel-Mottaleb, 2014), or if the labeled object is dominant in the infant's field of view (Yu & Smith, 2012).

## *Word learning biases*

Children do not associate words with just any meaning. Instead, they expect new words to refer to whole objects, rather than to their parts, and expect newly learned words to refer to categories of objects of the same shape or kind (see Hollich, Golinkoff, & Hirsh-Pasek, 2007; Markman, 1991). The origin of such expectations continues to be an important area of inquiry. Researchers have proposed diverse explanations: that these biases are built into the word-learning system (Markman, 1991), that they arise from children's social understanding (Bloom & Markson, 1998), or that they are learned from regularities in the environment (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). Regardless of their origin, evidence from bilingual and multilingual infants suggests that early word learning environments can change infants' use of these biases. Using a word learning bias known as *mutual exclusivity*, children reject two labels for the same object, expecting each object to have only one basic-level label (Markman & Wachtel, 1988). This has been demonstrated in monolinguals as young as 16 to 18 months (Halberda, 2003; Markman, Wasow, & Hansen, 2003). However, children growing up in bilingual and multilingual environments do not show mutual exclusivity from the same age (Byers-Heinlein & Werker, 2009; 2013; Houston-Price, Caloghiris, & Raviglione, 2010). This difference is likely because bilingual and multilingual children, unlike monolinguals, hear multiple labels for the same object—one in each language. Thus, while monolinguals' experience supports the notion of one-to-one mappings between words and objects, bilinguals' experience could lead to more flexible word learning.

## *Social information*

Infants also exploit rich social cues available in the environment, such as pointing and eye gaze, to help determine a word's meaning (Hollich *et al.*, 2000). For example, 18-month-old infants are more likely to link a speaker's utterance with an object when the speaker is attending to that object (Baldwin *et al.*, 1996). Similarly, when several different objects are present, children use their interlocutor's eye gaze and pointing to figure out what she is referring to (Baldwin & Moses, 2001). Beyond simply providing cues to a word's meaning, infants' understanding of a speaker's referential intentions is foundational to learning new words (Frank, Goodman, & Tenenbaum, 2009; Waxman & Gelman, 2009).

There is also evidence that infants' language background can affect their sensitivity to different types of social information. For example, because different speakers use different languages, bilingual children might be particularly sensitive to communicative information provided by a speaker. Consistent with this possibility, three-year-old bilinguals are better than monolinguals in using a speaker's gaze to find where a toy is hidden (Yow & Markman, 2011).

## *Vocabulary*

Children's vocabulary size provides a key index of their language development. To measure their receptive vocabulary—the words they can understand—children as young as 2 1/2 years are typically asked to point at which picture corresponds to a particular word (Dunn & Dunn, 2007). For younger children, parents check off different words that their child understands from a predetermined list (Fenson *et al.*, 2007). In both cases, the number of words understood is compared to age-referenced norms to understand how a particular child compares to her peers. Studies that extrapolate from such measures suggest that the average monolingual 12 month old can understand about 100 words, which jumps to around 550 words for the average monolingual 18 month old (Mayor & Plunkett, 2011).

Children's receptive vocabulary is almost invariably larger than their productive vocabulary, as typically they understand all the words they can say, but do not say all the words they can understand. This appears to be especially true for bilingual children, who may have particularly disproportionate receptive vocabularies compared to their productive vocabularies (Gibson, Oller, Jarmulowicz, & Ethington, 2011).

Typically, bilingual children understand fewer words in either of their languages than monolingual children understand in their single language (Bialystok, Luk, Peets, & Yang, 2010; Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2012). This is thought to arise because bilingual children's language input is inherently split between two languages (Byers-Heinlein & Lew-Williams, 2014). Despite knowing fewer words in each language, bilingual children usually learn words at the same rate as monolinguals, and importantly, they understand a similar number of total words when both languages are considered (De Houwer, Bornstein, & Putnick, 2013; Marchman, Fernald, & Hurtado, 2010; Pearson, Fernández, & Oller, 1995; Thordardottir, 2011). Bilingual children also understand translation equivalents—cross-language synonyms like English *cat* and Spanish *gato*—from an early age (De Houwer, Bornstein, & De Coster, 2006; Pearson *et al.*, 1995; Umbel, Pearson, Fernández, & Oller, 1992).

While there can be imbalances in vocabulary across a bilingual's two languages, there is no consistent evidence that bilingual children are more likely than monolingual children to experience delays or deficiencies in language learning. Bilingualism is not considered a risk factor for language learning, and bilingualism does not impose an additional burden on children diagnosed with impairments such as specific language impairment and autism spectrum disorders (Paradis, Crago, Genesee, & Rice, 2003; Peterson, Marinova-Todd, & Mirende, 2012).

## Understanding language in real time

Listeners usually encounter the words they know, as well as those they have yet to learn, in the context of running speech. Imagine if it took minutes or hours to determine the meaning of each incoming sentence—conversation would be impossible. Instead, communication occurs in real time, and young children show a developing ability to process speech as it unfolds. Fernald and colleagues (1998) presented young children with simple sentences (*Where's the baby?*), and found that 15 month olds take approximately one second to move their eyes toward a picture of a baby, while 24 month olds do so considerably faster. Similar developmental gains in real-time processing have been observed in Spanish-learning children from low-income households (Hurtado, Marchman, & Fernald, 2007). These findings are also echoed in studies of children's neural responses to familiar words (Friedrich & Friederici, 2005; Mills, Plunkett, Prat, & Schafer, 2005). Young children even begin to recognize words after hearing partial phonetic information, such as the onset *ba-* of *baby* (Swingley, Pinto, & Fernald, 1999).

Counter-intuitively, words in sentences can be easier for children to understand than words in isolation (Fernald & Hurtado, 2006). One reason is that children can leverage information across different parts of an utterance. For example, Spanish-learning children can use gender-marked articles like *la* and *el* ("the") to predict whether a speaker will name an object with a masculine or feminine grammatical gender (Lew-Williams & Fernald, 2007). Other studies show how young children exploit color and size adjectives (Fernald, Thorpe, & Marchman, 2010). Young monolingual children can even use familiar verbs and visual scenes to learn novel nouns (Ferguson, Graf, & Waxman, 2014; Waxman, Lidz, Braun, & Lavin, 2009), and use sentence structure to learn the meanings of novel verbs (Naigles, 1990).

Experimental studies with young children are beginning to elucidate how these words are organized in the developing mind. Priming studies investigate whether hearing one word (e.g., *cat*) helps children access words that are related in meaning (*dog*) or sound (*mat*). Research shows that from around their second birthday, both monolingual and bilingual children indeed make links between words with related meanings (Arias-Trejo & Plunkett, 2009; Singh, 2013) or with overlap in their sounds (Holzen & Mani, 2012; Mani, Durrant, & Floccia, 2012).

Children's ability to process language in real-time matters for later development. Children who respond faster to familiar words at age two have better language and cognitive outcomes in third grade, even when matched on overall vocabulary size (Marchman & Fernald, 2008). Similar longitudinal patterns have been observed in children with autism spectrum disorders (Venker, Eernisse, Saffran, & Weismer, 2013). Moreover, the speed of children's processing predicts which 18-month-old "late talkers" will make gains in language learning over the subsequent year (Fernald & Marchman, 2012).

Children's developing language expertise is built on a foundation of exposure to high-quality, high-quantity child-directed speech (Hart & Risley, 1995; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). On average, children from high-income families hear three to four times as much language as children from families on welfare. Even within low-income families, there is striking variability in the use of language in the household. Latino children from low-income families who hear more child-directed speech at home are faster in real-time language processing, and less likely to fall behind in language learning (Weisleder & Fernald, 2013). For bilingual children, relative exposure to each language shapes language processing and word learning in each language (Hurtado, Grueter, Marchman, & Fernald, 2014). Proficiency in one language does not carry over to the other language, as vocabulary size in one language is not related to the other, and processing efficiency in one language is not related to the other (Marchman *et al*., 2010).

## Challenges to language comprehension

While most research has tested infants' and toddlers' understanding of clearly articulated speech, real listening conditions are far from perfect. Everyday speech is replete with mispronunciations, accents, disfluencies, and background noise. Children have more difficulty understanding degraded speech than normal speech, but their ability to cope improves with age and vocabulary size (Zangl, Klarman, Thal, Fernald, & Bates, 2005).

Children's processing of mispronunciations provides a particularly interesting window into how they handle challenges to language comprehension. In typical laboratory studies of mispronunciations, children are shown pairs of object on a screen (e.g., a dog and a baby), and then hear a label either correctly pronounced (*Look at the baby!*), or mispronounced (*Look at the vaby!*). As early as age 12 months, monolinguals detect the mispronunciation, by looking less often and/or more slowly at the labeled object (Bailey & Plunkett, 2002; Swingley, 2005; White & Morgan, 2008). However, they still successfully identify the target object, demonstrating considerable flexibility in language comprehension. Experience improves children's word recognition: infants notice small sound changes more easily in familiar words than in newly learned words (Fennell, 2011; Stager & Werker, 1997).

Interestingly, there is evidence that Spanish-Catalan bilingual infants show a different pattern of processing mispronunciations than monolingual infants. As discussed previously, Spanish and Catalan share a high proportion of cognate words, such that they could be considered variant pronunciations rather than *mis*pronunciations. Bilingual toddlers do not respond differently to correctly pronounced versus mispronounced cognates (Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009), likely because they have learned to ignore small sound variations in cognates, which do not change a word's meaning across the languages. Indeed, when non-cognate words are mispronounced, bilinguals respond like monolinguals, showing less robust recognition than when words are correctly pronounced (Ramon-Casas & Bosch, 2010). Similarly, infants exposed to two dialects of the same language show less sensitivity to variant pronunciations than those exposed to a single dialect (Durrant, Delle Luche, Cattani, & Floccia, 2014), perhaps mirroring infants' ability to ignore surface variation across speakers and instead attend to underlying structure (see Pardo, this volume, and Pisoni, this volume; Graf Estes & Lew-Williams, 2015).

Accents are another type of variation that alters the phonetic form of speech. Young infants have difficulty learning and recognizing words spoken in a non-native accent, although children improve with age and through experience with a particular accent (Best, Tyler, Gooding, Orlando, & Quann, 2009; Schmale, Cristià, & Seidl, 2012; Schmale, Hollich, & Seidl, 2011). It is not surprising that children show some difficulty processing unfamiliar accents, as adults often show similar difficulties (Cristià *et al*., 2012).

Despite some parallels across the lifespan, infants are sometimes sensitive to phonetic variation that adults ignore. In one study, monolinguals learned new words best from a monolingual speaker, and bilinguals learned new words best from a bilingual speaker, even though the differences between the two speakers' pronunciations were very subtle (Fennell & Byers-Heinlein, 2014). This suggests that children are highly tuned to their language learning environments—something that researchers must take into account as they design studies comparing infants from different language backgrounds (Byers-Heinlein, 2014a). However, there is other evidence that children gravitate away from their parents' accent towards the accent of their wider communities. Twenty-month-old children exposed to two different English accents, one from their parents at home and one in their community, are best at identifying words pronounced in the accent of their communities (Floccia, Luche, Durrant, Butler, & Goslin, 2012). While these studies provide somewhat conflicting patterns of results, they underscore how sensitive children's comprehension can be to subtle sound changes.

Some types of imperfect speech can actually boost children's comprehension. Much of everyday speech contains disfluencies such as *um*, *ah*, and silent pauses. Interestingly, young children can exploit this information to their advantage, by capitalizing on the fact that disfluencies are particularly likely before unfamiliar and infrequent words. In one study, when two year olds heard a target noun preceded by a disfluency, for example, *Look at thee, uh, …*, they expected that the next word referred to an unfamiliar object, rather than to a familiar object (Kidd, White, & Aslin, 2011). Adults share similar expectations that disfluencies signal new information (Arnold, Fagnano, & Tanenhaus, 2003).

## Conclusions

While we often take particular joy in children's first words, children's early language comprehension constitutes an equally important, albeit somewhat hidden, side of language development. Early language comprehension sets the stage for successful development in many other areas, including language production (Fenson *et al*., 2007), and school-aged cognitive and language skills (Marchman & Fernald, 2008). This chapter has followed children's path to language comprehension: from orienting to their native language(s), to picking out the sound patterns of words and learning their meanings, to understanding speech in real-time. Children show remarkable flexibility in adapting to their language-learning environments, whether they are monolingual or multilingual.

## Acknowledgments

We thank the members of the Concordia Infant Research Lab / Laboratoire de Recherche sur l'Enfance de Concordia, and the Princeton Baby Lab. We also thank the National Institute of Child Health and Human Development for funding our recent bilingualism research (R03HD079779).

## REFERENCES

Antovich, D. M., & Graf Estes, K. (in press). Learning across languages: Bilingual experience supports dual language statistical word segmentation. *Developmental Science*.

Arias-Trejo, N., & Plunkett, K. (2009). Lexical-semantic priming effects during infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3633–3647. doi: 10.1098/rstb.2009.0146

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research*, 32(1), 25–36. doi: 10.1023/A:1021980931292

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321–324. doi: 10.1111/1467-9280.00063

Bailey, T., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, 17(2), 1265–1282.

Baldwin, D. A., & Moses, L. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development*, 10(3), 309–329. doi: 10.1111/1467-9507.00168

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, 67(6), 3135–3153. doi: 10.2307/1131771

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9), 3253–3258. doi: 10.1073/pnas.1113380109

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, 127(3), 391–397. doi: 10.1016/j.cognition.2013.02.011

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native-and Jamaican-accented words. *Psychological Science*, 20(5), 539–542.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13(04), 525–531. doi: 10.1017/S1366728909990423

Bloom, P., & Markson, L. (1998). Capacities underlying word learning. *Trends in Cognitive Sciences*, 2(2), 67–73. doi: 10.1016/S1364-6613(98)01121-8

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech segmentation. *Psychological Science*, 16(4), 298–304. doi: 10.1111/j.0956-7976.2005.01531.x

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2(1), 29–49. doi: 10.1207/S15327078IN0201_3

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46(2-3), 217–243. doi: 10.1177/00238309030460020801

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, 61(1), 93–125. doi: 10.1016/S0010-0277(96)00719-6

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33–B44. doi: 10.1016/S0010-0277(01)00122-6

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28(03), 455–474. doi: 10.1017/S0142716407070257

Byers-Heinlein, K. (2014a). Bilingual advantages, bilingual delays: Sometimes an illusion. [Peer commentary on "Moving toward a neuroplasticity view of bilingualism, executive control and aging" by S. Baum & D. Titone]. *Applied Psycholinguistics*, 35(05), 902–905. doi: 10.1017/S0142716414000204

Byers-Heinlein, K. (2014b). Languages as categories: Reframing the "one language or two" question in early bilingual development. *Language Learning*. doi: 10.1111/lang.12055

Byers-Heinlein, K., & Fennell, C. T. (2014). Perceptual narrowing in the context of increased variation: Insights from bilingual infants. *Developmental Psychobiology*, 56(2), 274–291. doi: 10.1002/dev.21167

Byers-Heinlein, K., & Lew-Williams, C. (2014). Bilingualism in the early years: What the science says. *LEARNing Landscapes*, 1–18.

Byers-Heinlein, K., Morin-Lessard, E., & Lew-Williams, C. (in press). Bilingual infants control their languages as they listen. *Proceedings of the National Academy of Sciences*.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, 12(5), 815–823. doi: 10.1111/j.1467-7687.2009.00902.x

Byers-Heinlein, K., & Werker, J. F. (2013). Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants. *Cognition*, 128(3), 407–416. doi: 10.1016/j.cognition.2013.05.010

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21(3), 343–348. doi: 10.1177/0956797609360758

Byers-Heinlein, K., Fennell, C. T., & Werker, J. F. (2012). The development of associative word learning in monolingual and bilingual infants. *Bilingualism: Language and Cognition*, 16(1), 198–205. doi: 10.1017/S1366728912000417

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. Proceedings of the Stanford Child Language Conference, 15, 17–29.

Cristià, A., Seidl, A. H., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology*, 3. doi: 10.3389/fpsyg.2012.00479/abstract

De Houwer, A., Bornstein, M. H., & De Coster, S. (2006). Early understanding of two words for the same thing: A CDI study of lexical comprehension in infant bilinguals. *International Journal of Bilingualism*, 10(3), 331–347. doi: 10.1177/13670069060100030401

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2013). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, 1–23. doi: 10.1017/S0142716412000744

Dietrich, C., Swingley, D., & Werker, J. F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of* *the National Academy of Sciences*, 104(41), 16027–16031. doi: 10.1073/pnas.0705270104

Dunn, L. M., & Dunn, L. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test*. Minneapolis, MN: Pearson Assessments.

Durrant, S., Delle Luche, C., Cattani, A., & Floccia, C. (2014). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language*. doi: 10.1017/S0305000914000063

Fennell, C. T. (2011). Object familiarity enhances infants' use of phonetic detail in novel words. *Infancy*, 17(3), 339–353. doi: 10.1111/j.1532-7078.2011.00080.x

Fennell, C. T., & Byers-Heinlein, K. (2014). You sound like Mommy: Bilingual and monolingual infants learn words best from speakers typical of their language environments. *International Journal of Behavioral Development*, 38(4), 309–316. doi: 10.1177/0165025414530631

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Steven Reznick, J., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories* (2nd ed.). Baltimore, MD: Brookes.

Ferguson, B., Graf, E., & Waxman, S. R. (2014). Infants use known verbs to learn novel nouns: Evidence from 15- and 19-month-olds. *Cognition*, 131(1), 139–146. doi: 10.1016/j.cognition.2013.12.014

Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. Scientific Reports, 6, 25434. doi:10.1038/srep25434

Fernald, A., & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, 9(3), F33–F40. doi: 10.1111/j.1467-7687.2006.00482.x

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, 83(1), 203–222. doi: 10.1111/j.1467-8624.2011.01692.x

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104–113. doi: 10.1037/0012-1649.20.1.104

Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9(3), 228–231. doi: 10.1111/1467-9280.00044

Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective-noun phrases. *Cognitive Psychology*, 60(3), 190–217. doi: 10.1016/j.cogpsych.2009.12.002

Floccia, C., Luche, C. D., Durrant, S., Butler, J., & Goslin, J. (2012). Parent or community: Where do 20-month-olds exposed to two accents acquire their representation of words? *Cognition*, 124(1), 95–100. doi: 10.1016/j.cognition.2012.03.011

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, 20(5), 578–585.

Friedrich, M., & Friederici, A. D. (2005). Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, 17(11), 1785–1802. doi: 10.1162/089892905774589172

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj, (Vol. 2, pp. 301–334). Hillsdale, NJ: Language Development.

Gibson, T. A., Oller, D. K., Jarmulowicz, L., & Ethington, C. A. (2011). The receptive–expressive gap in the vocabulary of young second-language learners: Robustness and possible mechanisms. *Bilingualism: Language and Cognition*, 1–15. doi: 10.1017/S1366728910000490

Gogate, L. J., & Bahrick, L. E. (2001). Intersensory redundancy and 7-month-old

infants' memory for arbitrary syllable-object relations. *Infancy*, 2(2), 219–231. doi: 10.1207/S15327078IN0202_7
Graf Estes, K., & Bowen, S. (2013). Learning about sounds contributes to learning about words: Effects of prosody and phonotactics on infant word learning. *Journal of Experimental Child Psychology*, 114(3), 405–417. doi: 10.1016/j.jecp.2012.10.002
Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 13(3), 254–260. doi: 10.1111/j.1467-9280.2007.01885.x
Graf Estes, K., & Lew-Williams, C. (2015). Listening through voices: Infant statistical word segmentation across multiple speakers. *Developmental Psychology*, 51, 1517–1528. doi: 10.1037/a0039725
Grieco-Calub, T. M., Saffran, J. R., & Litovsky, R. Y. (2009). Spoken word recognition in toddlers who use cochlear implants. *Journal of Speech, Language and Hearing Research*, 52(6), 1390–1400.
Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, 87(1), B23–B34. doi: 10.1016/S0010-0277(02)00186-5
Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes.
Hollich, G., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Young children associate novel words with complex objects rather than salient parts. *Developmental Psychology*, 43(5), 1051–1061. doi: 10.1037/0012-1649.43.5.1051
Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). The emergentist coalition model. *Monographs of the Society for Research in Child Development*, 65(3), 17–29. doi: 10.1111/1540-5834.00092
Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, 113(4). doi: 10.1016/j.jecp.2012.08.001
Horst, J., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13(2), 128–157. doi: 10.1080/15250000701795598
Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1570–1582. doi: 10.1037/0096-1523.26.5.1570
Houston, D. M., & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance*, 29(6), 1143–1154. doi: 10.1037/0096-1523.29.6.1143
Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, 15(2), 125–150. doi: 10.1111/j.1532-7078.2009.00009.x
Hurtado, N., Grueter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. *Bilingualism: Language and Cognition*, 17(01), 189–202. doi: 10.1017/S136672891300014X
Hurtado, N., Marchman, V. A., & Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language*, 34(02), 227. doi: 10.1017/S0305000906007896
Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27(2), 236–248. doi: 10.1037/0012-1649.27.2.236
Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*,

44(4), 548–567. doi: 10.1006/jmla.2000.2755
Johnson, E. K., & Seidl, A. H. (2009). At 11 months, prosody still outranks statistics. *Developmental Science*, 12(1), 131–141. doi: 10.1111/j.1467-7687.2008.00740.x
Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, 13(2), 339–345. doi: 10.1111/j.1467-7687.2009.00886.x
Johnson, E. K., Seidl, A. H., & Tyler, M. D. (2014). The edge factor in early word segmentation: Utterance-level prosody enables word form extraction by 6-month-olds. *PLoS ONE*, 9(1), e83546. doi: 10.1371/journal.pone.0083546.s001
Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29(1), 1–23. doi: 10.1006/cogp.1995.1010
Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159–207. doi: 10.1006/cogp.1999.0716
Kidd, C., White, K. S., & Aslin, R. N. (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. *Developmental Science*, 14(4), 925–934. doi: 10.1111/j.1467-7687.2011.01049.x
Krentz, U. C., & Corina, D. P. (2008). Preference for language in early infancy: The human language bias is not speech specific. *Developmental Science*, 11(1), 1–9. doi: 10.1111/j.1467-7687.2007.00652.x
Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2007). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 979–1000. doi: 10.1098/rstb.2007.2154
Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18(3), 193–198. doi: 10.1111/j.1467-9280.2007.01871.x
Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition*, 122(2), 241–246. doi: 10.1016/j.cognition.2011.10.007
Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science*, 14(6), 1323–1329.
Mackenzie, H., Curtin, S., & Graham, S. A. (2012a). 12-month-olds' phonotactic knowledge guides their word-object mappings. *Child Development*, 83(4), 1129–1136. doi: 10.1111/j.1467-8624.2012.01764.x
Mackenzie, H., Curtin, S., & Graham, S. A. (2012b). Class matters: 12-month-olds' word-object associations privilege content over function words. *Developmental Science*, 15(6), 753–761. doi: 10.1111/j.1467-7687.2012.01166.x
Mani, N., Durrant, S., & Floccia, C. (2012). Activation of phonological and semantic codes in toddlers. *Journal of Memory and Language*, 66(4), 612–622. doi: 10.1016/j.jml.2012.03.003
Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11(3), F9–F16. doi: 10.1111/j.1467-7687.2008.00671.x
Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. *Journal of Child Language*, 37(4), 817–840. doi: 10.1017/S0305000909990055
Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In *Perspectives on*

*Language and Thought: Interrelations in Development*. Cambridge Univ Press.
Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20(2), 121–157. doi: 10.1016/0010-0285(88)90017-5
Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, 47(3), 241–275. doi: 10.1016/S0010-0285(03)00034-3
Martin, A., Onishi, K. H., & Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12 months. *Cognition*, 123(1), 50–60. doi: 10.1016/j.cognition.2011.12.003
Matatyaho-Bullaro, D. J., Gogate, L., Mason, Z., Cadavid, S., & Abdel-Mottaleb, M. (2014). Type of object motion facilitates word mapping by preverbal infants. *Journal of Experimental Child Psychology*, 118(C), 27–40. doi: 10.1016/j.jecp.2013.09.010
Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–B111. doi: 10.1016/S0010-0277(01)00157-3
Mayor, J., & Plunkett, K. (2011). A statistical estimate of infant and toddler vocabulary size from CDI analysis. *Developmental Science*, 14(4), 769–785. doi: 10.2307/25057972?ref=no-x-route:58a0ca288783055a71a92e8d3b06ac76
McCabe, A., Tamis-LeMonda, C., Bornstein, M., Golinkoff, R. M., Hirsh-Pasek, K., Hoff, E., Kuchiro, Y., Melzi, G., Mendelson, A., Paez, M., Song, L., & Wishard, A. (2013). Multilingual children: Beyond myths towards best practices. *Social Policy Report for the Society for Research in Child Development*, 27(4).
Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143–178. doi: 10.1016/0010-0277(88)90035-2
Mersad, K., & Nazzi, T. (2012). When mommy comes to the rescue of statistics: Infants combine top-down and bottom-up cues to segment speech. *Language Learning and Development*, 8(3), 303–315. doi: 10.1080/15475441.2011.609106
Mills, D. L., Plunkett, K., Prat, C., & Schafer, G. (2005). Watching the infant brain learn words: Effects of vocabulary size and experience. *Cognitive Development*, 20(1), 19–31. doi: 10.1016/j.cogdev.2004.07.001
Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, 19(3), 326–337. doi: 10.1111/infa.12041
Moon, C., Cooper, R., & Fifer, W. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16(4), 495–500. doi: 10.1016/0163-6383(93)80007-U
Naigles, L. R. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17(02), 357–374. doi: 10.1017/S0305000900013817
Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756–766. doi: 10.1037/0096-1523.24.3.756
Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, 43(1), 1–19. doi: 10.1006/jmla.2000.2698
Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). French-English bilingual children with SLI: How do they compare with their monolingual peers? *Journal of Speech, Language, and Hearing Research*, 46(1), 113–127. doi: 10.1044/1092-4388(2003/009)

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1995). Cross-language synonyms in the lexicons of bilingual infants: One language or two? *Journal of Child Language*, 22(2), 345–368. doi: 10.1017/S030500090000982X

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, 80(3), 674–685. doi: 10.1111/j.1467-8624.2009.01290.x

Peña, E. D., Gillam, R. B., Bedore, L. M., & Bohman, T. M. (2011). Risk for poor performance on a language screening measure for bilingual preschoolers and kindergarteners. *American Journal of Speech-Language Pathology*, 20(4), 302. doi: 10.1044/1058-0360(2011/10-0020)

Petersen, J. M., Marinova-Todd, S. H., & Mirenda, P. (2012). Brief report: An exploratory study of lexical skills in bilingual children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 42(7), 1499–1503. doi: 10.1007/s10803-011-1366-y

Petitto, L. A., Katerelos, M., Levy, B. G., Gauna, K., Tetreault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language*, 28(2), 453–496.

Polka, L., & Sundara, M. (2003). Word segmentation in monolingual and bilingual infant learners of English and French. *Proceedings of the 15th International Congress of Phonetic Sciences*, 1021–1024.

Polka, L., & Sundara, M. (2012). Word segmentation in monolingual infants acquiring Canadian English and Canadian French: Native language, cross-dialect, and cross-language comparisons. *Infancy*, 17(2), 198–232. doi: 10.1111/j.1532-7078.2011.00075.x

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2012). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, 17(1), 1–15. doi: 10.1177/1367006911431198

Ramon-Casas, M., & Bosch, L. (2010). Are non-cognate words phonologically better specified than cognates in the early lexicon of bilingual children? In M. Ortega-Llebaria, *Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 31–36). Sommerville, MA: Cascadilla.

Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, 59(1), 96–121. doi: 10.1016/j.cogpsych.2009.02.002

Rice, M. L., Warren, S. F., & Betz, S. K. (2005). Language symptoms of developmental language disorders: An overview of autism, Down syndrome, fragile X, specific language impairment, and Williams syndrome. *Applied Psycholinguistics*, 26(01), 7–27. doi: 10.1017/S0142716405050034

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6), 906–914. doi: 10.1002/wcs.78

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. doi: 10.1126/science.274.5294.1926

Schmale, R., Cristià, A., & Seidl, A. H. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science*, 15(6), 732–738. doi: 10.1111/j.1467-7687.2012.01175.x

Schmale, R., Hollich, G., & Seidl, A. H. (2011). Contending with foreign accent in early word learning. *Journal of Child Language*, 1–13. doi: 10.1017/S0305000910000619

Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science*, 12(6), 874–887. doi: 10.1111/j.1467-7687.2009.00829.x

Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science*, 23(9), 994–999. doi: 10.1177/0956797612436817

Seidl, A. H., Tincoff, R., Baker, C., & Cristià, A. (2014). Why the body comes first: Effects of experimenter touch on infants' word finding. *Developmental Science*. Advance online publication. doi: 10.1111/desc.12182

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, 54(1), 1–32. doi: 10.1016/j.cogpsych.2006.04.002

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, 108(15), 6038–6043. doi: 10.1073/pnas.1017617108

Singh, L. (2013). One world, two languages: Cross-language semantic priming in bilingual toddlers. *Child Development*. doi: 10.1111/cdev.12133

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. doi: 10.1016/j.cognition.2007.06.010

Smith, L. B., Jones, S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. K. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, 13(1), 13–19. doi: 10.1111/1467-9280.00403

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388(6640), 381–382. doi: 10.1038/41102

Sundara, M., & Scutellaro, A. (2010). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, 39(4), 505–513. doi: 10.1016/j.wocn.2010.08.006

Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108(1), 232–242. doi: 10.1016/j.cognition.2007.12.013

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8(5), 432–443. doi: 10.1111/j.1467-7687.2005.00432.x

Swingley, D. (2010). Fast mapping and slow mapping in children's word learning. *Language Learning and Development*, 6(3), 179–183. doi: 10.1080/15475441.2010.484412

Swingley, D., Pinto, J., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, 71(2), 73–108. doi: 10.1016/S0010-0277(99)00021-9

Thiessen, E. D., & Pavlik, P. I. (2013). iMinerva: A mathematical model of distributional statistical learning. *Cognitive Science: a Multidisciplinary Journal*, 37(2), 310–343.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, 39(4), 706–716. doi: 10.1037/0012-1649.39.4.706

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, 1–21. doi: 10.1177/1367006911403202

Tincoff, R., & Jusczyk, P. W. (2011). Six-month-olds comprehend words that refer to parts of the body. *Infancy* 17(4), 432–444. doi: 10.1111/j.1532-7078.2011.00084.x

Umbel, V., Pearson, B. Z., Fernández, M., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development*, 63(4), 1012–1020. doi: 10.2307/1131250

Venker, C. E., Eernisse, E. R., Saffran, J. R., & Weismer, S. E. (2013). Individual differences in the real-time comprehension of children with ASD.

*Autism Research*, 6(5), 417–432. doi: 10.1002/aur.1304

Vihman, M. M., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. (2007). Onset of word form recognition in English, Welsh, and English-Welsh bilingual infants. *Applied Psycholinguistics*, 28(3), 475–493. doi: 10.1017/S0142716407070269

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10(2), 159–164. doi: 10.1111/j.1467-7687.2007.00549.x

Vouloumanos, A., & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, 45(6), 1611–1617. doi: 10.1037/a0016134

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, 13(6), 258–263. doi: 10.1016/j.tics.2009.03.006

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, 59(1), 67–95. doi: 10.1016/j.cogpsych.2009.02.001

Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, 316(5828), 1159. doi: 10.1126/science.1137686

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 1–22. doi: 10.1177/0956797613488145

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Sciences*, 12(4), 144–151. doi: 10.1016/j.tics.2008.01.008

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49–63. doi: 10.1016/S0163-6383(84)80022-3

Werker, J. F., Cohen, L., Lloyd, V., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, 34(6), 1289–1309. doi: 10.1037/0012-1649.34.6.1289

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, 59(1), 114–132. doi: 10.1016/j.jml.2008.03.001

Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-Month-olds use distinct objects as cues to categorize speech information. *Cognition*, 1–10. doi: 10.1016/j.cognition.2009.08.010

Yeung, H. H., Chen, L. M., & Werker, J. F. (2013). Referential labeling can facilitate phonetic learning in infancy. *Child Development*. doi: 10.1111/cdev.12185

Yow, W. Q., & Markman, E. M. (2011). Young bilingual children's heightened sensitivity to referential cues. *Journal of Cognition and Development*, 12(1), 12–31. doi: 10.1080/15248372.2011.539524

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, 125(2), 244–262. doi: 10.1016/j.cognition.2012.06.016

Zangl, R., Klarman, L., Thal, D. J., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, 6(2), 179–208. doi: 10.1207/s15327647jcd0602_2

# 24 Names for Things… and Actions and Events: Following in the Footsteps of Roger Brown

DANI LEVINE[1], KRISTINA STROTHER-GARCIA[2], KATHY HIRSH-PASEK[1], AND ROBERTA MICHNICK GOLINKOFF[2]

[1] Temple University
[2] University of Delaware

*I cannot convince myself that there is any principle of correctness in names other than convention and agreement; any name which you give, in my opinion, is the right one, and if you change that and give another, the new name is as correct as the old.*
*–Plato, Cratylus*

Debates about how we learn names have occupied philosophers for centuries and psychologists for decades. Plato proposed two ways we might acquire names for things. The first possibility is that naming is a social convention derived from the culture of use. The other is that names are intrinsically linked to that which they represent. This discussion continued in the Confucian Xunzi (ca. 310–ca. 210 BCE) with the publication of the *Right Use of Names* and enjoyed a resurgence of interest in the Chomskian period of the mid-twentieth century. It was then that Roger Brown (1958) wrote his now classic book, *Words and Things*. Brown not only offered a theoretical treatment of how words map onto world, but also provided substantial diary data in *A First Language* (1973), which still serves as a foundation for research in word learning today. This chapter reviews this vibrant empirical enterprise and demonstrates just how far we have come in understanding how children learn words. Brown's work represented the first modern day treatment of this topic in psychology, and we think he would be pleased with what our field has

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

accomplished since. We offer but a portion of the research on word learning over the past 50 years. Interestingly, most of the current debate still sits with the age-old questions of how children learn to link words to world and how word learning processes change with development and experience (Hollich *et al.*, 2000).

## Introduction

As Paul Bloom (2002) noted in *How Children Learn the Meanings of Words*, word learning seems like it should be strikingly simple. Say the word *dog* in the presence of a dog and a child is sure to associate the word with its referent. Pilley and Reid (2011) demonstrated that their border collie, Chaser, could do precisely the same thing. Chaser correctly identified 1,022 objects by name. If a dog can learn over 1,000 words in three years, why is it any surprise that human children learn 14,000 words by the age of six (O'Grady & Archibald, 2010)? Obviously, children—and dogs—associate words they hear with the objects, actions, or events that are most salient at the time. Or is it so obvious?

As it turns out, *associationistic* learning supports the beginnings of word learning (Pruden *et al.*, 2006), but this mechanism alone cannot cleanly explain lexical development (Hollich *et al.*, 2000; Quine, 1960). Even a seasoned linguist could be thrown by the ambiguity in the *dog* scene. Might the word refer to the dog's ears or the dog's panting rather than the whole dog? Quine (1960) suggested that, given the vast number of options, there must be some way to resolve the indeterminacy of reference. Indeed, the problem is even more staggering: children do not merely learn perceptually salient object names—they also learn words for categories like *furniture*, abstract concepts like *truth*, relations like *connection*, and actions like *poking*. Words like *savage* are rooted in a cultural context of social and linguistic information above and beyond simple associative cues. The resulting variety in word types is necessary to achieve the level of complexity found in human language, but it makes the task of discovering a word's meaning that much harder for the child. Might word learning be influenced by constraints that bias the child toward certain interpretations (e.g., assuming a novel label refers to a *whole object* rather than a part or property of it), or perhaps a set of guiding social cues, or even the use of statistical computations that support the kinds of cross-situational learning that enables lexical acquisition? All of these possibilities have been posited in the literature.

This chapter explores a variety of theories and endeavors to explain how young children ultimately converge on lexical acquisition. We present the current research in five sections. In the first, we tackle the thorny question of what counts as a word. Using this definition, the second section identifies the processes children use to learn new words, and how these processes change over time. The required inputs for word learning—both linguistic and nonlinguistic—are discussed in Section 3. The fourth section offers readers a timeline of lexical acquisition and reveals some near-universal patterns observed in word learning across many of the world's languages. Finally, Section 5 revisits Plato's and Brown's questions by focusing on contemporary models of word learning and the ways in which they solve the mapping problem.

## What does it mean to know a word?

Word knowledge is defined, broadly, in two important ways. One definition focuses on the sound patterns, semantics, pragmatics, and syntax that specify adult linguistic knowledge (Brandone *et al*., 2006). The second definition reflects the real-time processes (i.e., perception, action, learning, attention, and memory) required for understanding or producing a word in the particular context (e.g., social cues, task demands) of a given moment (Bates & MacWhinney, 1989; Colunga & Smith, 2008; Regier, 2005). Modern theories affirm that word knowledge and the processing necessary for that knowledge are indistinguishable; that is, "knowledge is an abstraction over many underlying processes," (Smith, Colunga, & Yoshida, 2010). We begin here by outlining the various components of that abstract knowledge, as a lead-in to understanding the motley of processes that, together, ground lexical acquisition.

Certainly, a prerequisite to word knowledge involves recognition of a patterned string of sounds, for example, "dawg," (or handshapes and movements, in the case of signed languages). Beyond identifying a consistent phonological shape, however, there are additional criteria that must be met for a sound sequence to count as a word.

Knowing a word requires at least some *semantic information*. Children must minimally know that the sound unit "dawg" is associated with the particular dog that is present when they first hear the word. Many writers (Bloom, 2002; Golinkoff, Mervis, & Hirsh-Pasek, 1994; Golinkoff & Hirsh-Pasek, 1999) contend that a sound unit does not achieve word status until the child can extend the label to other members of the same category (e.g., to other dogs). Expanding on the basic referential meaning of a word, children later make connections between a given word and other related ones, forming a *semantic network*. For example, a *ball* is part of a category of objects called *toys*, and may be used to *play* a *game*. It is also part of a category of objects that are *round*, along with *oranges* and *marbles*. Thus the word *ball* is part of a network (including words like *toy, play, game, round, orange*, and *marble*) based on semantic relations.

*Pragmatics*, which includes social and cultural information about how to use a word in a given context, is another building block for word knowledge. Relatively early in the course of lexical development, words begin to take on communicative functions, being understood as a means to socially and intentionally share information. As word learning progresses, the pragmatics of a word become more intricate. This includes understanding how a word can influence other people's actions; for example, knowing that the simple word *stop* conveys the desire that another person halt their behavior. Additionally, pragmatics encompasses knowledge of what Tomasello (2008) called *common ground*—that is, the information shared between people in a conversation based on past experiences, cultural knowledge, and topics discussed earlier. In these ways, pragmatic information connects the literal meaning of a word to its real-world implications, as when "bad" paradoxically means good to members of a given subculture.

Finally, *syntactic information* dictates how a word combines structurally with other words in a phrase or sentence. Although children's earliest word representations likely lack this information, it is essential for complete word knowledge. Once children have acquired even a partial understanding of abstract syntactic structures, they use syntactic contexts to learn new words and to solidify the meanings of old words. One of the first demonstrations of how children do this was provided by Roger Brown (1957). He showed children a drawing of a person manipulating a substance in a bowl. When he asked children to point to "some sib" they pointed to the substance, but when he asked where the person was "sibbing" they pointed to the kneading action. This process is called *syntactic bootstrapping* (Gertner & Fisher, 2012; Gleitman *et al.*, 2005) and entails using the argument structure a word appears in to glean something of its meaning. As children's linguistic knowledge matures, the syntactic representation of a word comes to include specific information about part of speech (noun, verb, adjective, adverb, preposition, etc.) and about the types of *syntactic arguments* the word requires. The verb *kiss*, for instance, requires two arguments: an *agent* to perform the action of kissing and a *recipient* of the kiss. A sentence with the verb *kiss* will be ungrammatical if one of the arguments is missing, as illustrated in (1) below.

(1) a. Sally kissed the baby
    b. *Sally kissed.
    c. *Kissed the baby.

Note that the semantics dictate what types of things can be the arguments of a given verb; part of the meaning of *kiss* also includes the fact that only people (and perhaps certain animals) can be the *agent*. This semantic requirement of *kiss* explains why sentence (1a) is understandable while sentence (2) is not, except perhaps in a poetic sense. In many theories, semantic and syntactic information are thought to be stored with the word's lexical representation (Bresnan, 1978).

(2) The door kissed John.

The information that constitutes word meaning is complex, even during the first few years of life. The earliest words may enter the lexicon with only their phonology and a basic understanding of their semantics. That is, first words may initially be "things heard most often in the presence of a particular object," acquired via cross-situational learning mechanisms (see below, Models of Word Learning). These words might therefore constitute partial or incomplete lexical entries, not yet representing the entire reach of the word's meaning (Yurovsky *et al.*, 2014). For instance, Seston and colleagues (2009) found that 6 year olds evince protracted word development when extending words to odd, metaphorical uses as in, "He *vacuumed* with his mouth." Later on, lexical entries for earlier learned words will be expanded to include more semantic, pragmatic, and syntactic information. What it means to know a word, and the processes that support that burgeoning knowledge, develop gradually, alongside the growing lexicon.

## The process of word learning

With all these component pieces, it is little wonder that linking the word *dog* with its referent is far from a simple process. Children must first segment units of speech from strings of *sounds*, which are not well punctuated with stops and starts. That is, they have to isolate the phrases and individual words. Second, they have to segment a continuous stream of *events* into the objects, actions and event units that will be labeled by those words and phrases. Third, children must map linguistic units onto the objects, actions and events they refer to—often called the *mapping problem*. This latter challenge has turned out to be somewhat intractable and is the subject of most theoretical debates on word learning today.

### *Speech segmentation*

Before children can begin to learn what words mean, they must first recognize where one word ends and another begins. Though this segmentation seems obvious to adults, there are actually no pauses or reliable acoustic signals to indicate word boundaries in natural speech. So how do infants begin to parse the speech stream? Shortly after birth, sleeping neonates' brain responses to speech reveal a precocious sensitivity to the statistical structure underlying language (Teinonen *et al.*, 2009). Statistical cues, such as the likelihood of certain syllables being adjacent, are crucial for early word segmentation. For example, "bee" may be heard more often after "bay" (as in *baby*) than after "go" (as in *go before*), indicating that "bay-bee" is a word while "go-bee" is not. Newborns are also sensitive to the *prosody* or rhythmic patterns of language, as evidenced by changes in their sucking rate in response to hearing alternations between stress-timed languages (e.g., English) and mora-timed languages (e.g., Japanese; Nazzi, Bertoncini, and Mehler, 1998).

As infants gain experience with the language(s) they are exposed to, they develop language-specific biases that facilitate a more fine-tuned approach to word segmentation. Given the consistency of prosodic changes at clause boundaries in English (e.g., rises and falls in fundamental frequency; see Jusczyk, 1986 for a review), infants rapidly develop a sensitivity to phrase boundaries (Hirsh-Pasek *et al.*, 1987). By seven to nine months, infants show a listening preference for speech with pauses inserted at clausal boundaries relative to speech containing pauses within syntactic units (Hirsh-Pasek *et al.*, 1987). This demonstrates infants' remarkable ability to home in on important linguistic structures before they can understand what the words that form these structures actually mean. In this way, infants identify linguistic patterns early on that will help them learn words later in development.

Similarly, infants quickly acquire a parsing heuristic based on the lexical stress patterns of their language. By 7.5 months, English-learning infants segment strong/weak bisyllabic units (e.g., "*cray*on") but not weak/strong units (e.g., "sur*prise*"; Jusczyk, Houston, & Newsome, 1999), and are only able to extract trisyllabic words when the first syllable is stressed ("*par*achute" versus "tambou*rine*"; Houston, Santelmann, & Jusczyk, 2004). The ability to identify likely words from the speech stream before those words carry meaning is critical for ultimately mapping those word segments onto referents. Indeed, stress-based segmentation abilities at seven months predict vocabulary size at age three (Kooijman *et al*., 2013).

Statistical segmentation of speech also matures with language experience. As early as eight months, infants use statistical regularities to distinguish coherent syllabic units from non-units in a monotone, nonsense speech sample (Saffran, Aslin, & Newport 1996). Seventeen month olds capitalize on this ability for word learning; they learn a word-referent mapping if the label was previously presented in fluid speech, but not if the label is a novel syllabic sequence (Graf Estes *et al*., 2007).

In addition to these developments in bottom-up speech segmentation, stored knowledge of words becomes a tool for infants who use these words to conduct top-down analyses of the speech stream. This begins with the child's own name, which infants recognize at 4.5 months of age (Mandel, Jusczyk, & Pisoni, 1995). They then can use their name to isolate a novel word appearing after their name (but not someone else's name) by six months of age (Bortfeld *et al*., 2005). Speech segmentation, via developing bottom-up and top-down mechanisms, is clearly a critical step in word learning.

## *Segmentation of events*

Just as children must segment the sound stream, they must also segment events into meaningful units. Imagine a parent picking up a toy and putting it on a shelf. This sounds like two events as written here, but it also could be viewed as one ("putting the toy away") or even three ("grabbing the toy, moving it to the shelf, and placing it"). Infants are faced with the challenge of unitizing the rich, continuous stream of nonlinguistic events into meaningful categorical units that will be labeled by language.

This area of research is in its infancy, but it suggests a developmental trajectory similar to that of word segmentation. Newborns evince a very limited sensitivity to statistical event structure (Bulf, Johnson, & Valenza, 2011), maturing into more sophisticated visual statistical learners by seven to nine months of age (Roseberry *et al*., 2011; Stahl *et al*., 2014). Infants' detection of event goals (Lakusta *et al*., 2007) may also be crucial for the parsing of continuous events (Levine *et al*., 2017). Experience identifying event goals early in life may facilitate identification of actor intent in events later on, which in turn simplifies and aids in the process of segmenting events (Baldwin *et al*., 2001). Critically, segmentation of events is a foundational prerequisite for learning verbs, which map onto transient units of events (Friend & Pace, 2011; Golinkoff & Hirsh-Pasek, 2008).

### *Language is special*

Mapping word to world requires the understanding that words (and *not* other types of sounds) carry meaning as symbols; this understanding is gradually fine-tuned with language experience. By three months, infants can use a novel speech segment paired with a series of objects (e.g., fish exemplars), to form a category of those objects (i.e., *fish*; Ferry, Hespos, & Waxman, 2010). When this speech is replaced by a matched sequence of sine-wave tones, infants fail to form the object category (Ferry, Hespos, & Waxman, 2010). However, lemur vocalizations succeed at facilitating categorization similar to human speech at this age, and it is not until six months that the effect of nonhuman primate vocalizations disappears (Ferry, Hespos, & Waxman, 2013).

By 12 months, infants demonstrate their understanding that non-linguistic human noises (e.g., coughing), unlike words, do not communicate information about a target object (Martin, Onishi, & Vouloumanos, 2012). Infants at this age can learn a word-object pairing following habituation to the coupling, but given the same procedure, are unable to learn pairings of objects with nonlinguistic communicative sounds (e.g., "oooh") or consonantal sounds (e.g., "/l/"; MacKenzie, Graham, & Curtin, 2011). Twelve month olds also recognize that different languages use different labels for a given object, and do not expect a speaker of another language to use the same label for a given object as a speaker of their native language (Scott & Henderson, 2013). Still, given sufficient attentional cues, infants aged 12–18 months will map almost any symbol to an object—from non-native language sounds (e.g., "tsk-tsk"; May & Werker, 2014) to gestures (Namy & Waxman, 1998) to whistles and digitized sounds (Woodward & Hoyne, 1999; Hollich *et al*., 2000). By 20–26 months, however, infants fail to map anything but native-sounding words to objects, even with referential cues (May & Werker, 2014; Namy & Waxman, 1998; Woodward & Hoyne, 1999). Thus, the selectivity of words as symbols becomes greater over the first two years of life, leading children to develop more specialized means of language learning, beyond the general associative mechanisms they start out with (Namy, 2012).

## What it takes to learn a word: Quantity and quality of input

On a fundamental level, infants must receive input to learn, through their exposure to a language (i.e., perceptual input that is symbolic and communicative) and non-linguistic information (i.e., all other perceptual input as well as action experiences). This section explores the input children require (and that which they do not require) in order to acquire a lexicon.

### *Language input*

Receiving some type of language input is a guarantee for almost every infant (but see Fromkin *et al*., 1974). Thus, the vast majority of children become competent users of their native language. Despite the near universality of lexical acquisition, there is a great deal of variation in language input that is reflected in children's vocabulary outcomes. While a child from a family on welfare hears 616 words per hour, a child brought up by a professional family hears more than three times that amount (Hart & Risley, 1995). Considering the fact that 86% to 98% of the words in children's vocabularies at age three are words used by their parents, language input stands as a major determinant of children's lexical store (Hart & Risley, 1995). Hurtado, Marchman, and Fernald (2008) extended this research, demonstrating that the amount of language input at 18 months predicts vocabulary size and lexical processing efficiency at 24 months. This suggests that input quantity affects not only which words children acquire, but also how rapidly they understand the words they hear.

If lexical development was simply determined by the quantity of input, we could set infants up with books on tape and walk away. To assess the potential importance of input *quality*, one study asked a sample of adults to watch muted vignettes of a variety of parent-child interactions and to guess what the parents were saying at select moments in the videos. The children of parents whose words could be readily guessed by naive adult viewers had significantly larger vocabularies three years later, as compared to children of parents whose words were more difficult to infer from the socio-visual context (Cartmill *et al.*, 2013). Providing disambiguating social and visual cues during speech may therefore be critical to vocabulary acquisition.

The importance of unambiguous word learning situations for lexical development is also evidenced by situations in which children are unable to learn words. For example, Weisleder and Fernald (2013) demonstrated that language input that is not specifically directed toward the child (i.e., overheard words) does not contribute to vocabulary outcomes. Although laboratory experiments have suggested children could learn word mappings by overhearing speech (Akhtar, 2005; Floor & Akhtar, 2006; Yuan & Fisher, 2009), more naturalistic studies indicate that this is only possible with experimental constraints narrowing children's attentional focus (Shneidman *et al.*, 2013; Shneidman & Goldin-Meadow, 2012; Weisleder & Fernald, 2013). Children are also typically unable to learn words from video prior to age three (e.g., Zimmerman, Christakis, & Meltzoff, 2007). However, when video is live (e.g., over Skype) and involves socially contingent interactions, even verbs—harder to learn than nouns—can be learned from video as early as age two (Roseberry, Hirsh-Pasek, & Golinkoff, 2014). A growing literature emphasizes that adult talk must not only be directed toward the child, but must also be appropriate to the specific interaction in terms of timing, content, and intensity in order to resolve ambiguity in word learning situations (Bornstein *et al.*, 2008; Roseberry *et al.*, 2014; Tamis-LeMonda, Kuchirko, & Song, 2014).

In addition to these overall effects of linguistic quantity and quality, the importance of different aspects of linguistic input changes (or should change) as the child becomes a more sophisticated user of language. After assessing parental language in parent-child interactions, Rowe (2012) found that the most critical aspect of input contributing to vocabulary growth at 18 months was the quantity of parental speech; at 30 months, diversity and sophistication of vocabulary were the largest contributors to children's vocabulary development; and at 42 months, it was parents' use of decontextualized language (i.e., language removed from the immediate environment) that most significantly contributed to vocabulary advancement. Thus, children rely on different aspects of language input over the course of development, from building a foundational vocabulary of common words, to adding uncommon words, to practicing the language necessary for extended narratives (Rowe, 2012).

## *Infant-directed speech*

The acoustic properties of language input also make a difference for vocabulary development (Ma *et al.*, 2011; Yurovsky, Yu, and Smith, 2012). Originally called *motherese*, infant-directed speech (IDS), describes a particular register used by adults (Newport, 1975) and even by children without siblings of their own (Shatz & Gelman, 1973) when addressing infants and younger children (Broesch & Bryant, 2015; Fernald *et al.*, 1989). This register involves slower rates of speaking, longer vowels and pauses, shorter phrases, and higher and more variable pitches as compared to adult-directed speech (ADS; Andruski & Kuhl, 1996; Fernald & Simon, 1984; Graf Estes & Hurley, 2013; McRoberts & Best, 1997). IDS is also characterized by certain sentence structures: in English, the label of a referent often occurs in the final position of the sentence and that label is typically preceded by a frequently used article (e.g., "Look at *the balloon*"; Yurovsky, Yu, & Smith, 2012).

Although IDS has not always been extolled (Dougherty, 2000), research has demonstrated its value for word learning in children (Graf Estes & Hurley, 2013; Ma *et al.*, 2011; Ramirez-Esparza, Garcia-Sierra, & Kuhl, 2014; Singh *et al.*, 2009) and even in adults (Golinkoff & Alioto, 1995). In one study, seven and eight month olds were familiarized with words delivered either in IDS or ADS (Singh *et al.*, 2009). Twenty-four hours later, infants recognized words presented in ADS that were originally heard in IDS, but did not recognize words originally heard in ADS (Singh *et al.*, 2009). A second study presented 17 month olds with novel label-object pairs using IDS or ADS (Graf Estes & Hurley, 2013). Infants learned the labels only in the IDS condition, and only when prosody was varied rather than constant (Graf Estes & Hurley, 2013).

Despite the early advantage of IDS over ADS for word learning, children do not rely on IDS forever. At 21 months, infants with larger vocabularies than their peers learn novel words from ADS, and by 27 months even those with below-average vocabularies can do the same (Ma *et al.*, 2011). These findings suggest a developmental progression in which IDS is crucial for word learning early on, when much of the speech stream is unfamiliar to the infant, but becomes less critical as the lexicon grows.

## *Nonlinguistic input*

Perhaps less intuitively, nonlinguistic information is also critical for lexical development. One important clue to word meaning is where the speaker is looking—their *eye gaze*. As early as 12 months, infants attend to a speaker's eye gaze for substantially longer periods of time when the word learning situation is ambiguous than when it is unambiguous (Baldwin, Bill, & Ontai, 1996). Infants at this stage also show a developing sensitivity to gestural cues; dynamic gestures synchronized with object labeling promote greater attention to the labeled object than asynchronous dynamic gestures or static gestures (Rader & Zukow-Goldring, 2012).

Beginning in the second year of life, visually available social cues affect the success of word-referent mapping. For example, 18- to 20-month-old infants can map a label to an object only if the adult labeling the referent is observed attending to the object; if the adult is out of sight, the mapping fails (Baldwin *et al*., 1996). This illustrates the importance of *joint attention*—or the situation in which a child and her caretaker are both focused on the same object or event. Mothers and children speak more during episodes of joint attention, and mothers' frequency of object labeling during these episodes predicts later vocabulary (Tomasello & Farrar, 1986). Additionally, more novel words are learned if parents simultaneously look at and label the object their child is focused on rather than looking at other objects during labeling (Akhtar, Dunham, & Dunham, 1991). The redundancy of visual socio-pragmatic cues also increases the probability that a child will correctly map a word to its referent. Toddlers are more likely to learn a word when pointing accompanies eye gaze than when gaze cues are provided alone (Booth, McGregor, & Rohlfing, 2008; Hollich *et al*., 2000).

## *Infant-directed action*

Just as adults modify their speech when addressing infants, they also modify their actions. This more salient form of nonlinguistic input is called *infant-directed action* (IDA) or *motionese*. When labeling objects for infants, adults use more exaggerated and repeated actions, less complex combinations of actions, and more attempts to elicit interaction than in adult-directed action (ADA; Brand, Baldwin, & Ashburn, 2002). Speech is often synchronized with IDA, such that when a mother moves an object in the presence of her infant, she is more likely to label it than to use other non-labeling words (Gogate, Bahrick, & Watson, 2000). Moreover, six to eight month olds are more likely to map a word onto a referent when mothers make use of this label-movement synchrony (Gogate, Bolzani, & Betancourt, 2006).

Even the type of object motion concurrent with labeling makes a difference in the success of the object-label mapping. Mothers use looming or shaking object motions more often than upward or sideways motions when teaching novel object labels to their six- to eight-month-old infants (Matatyaho & Gogate, 2008). Word learning is facilitated when infants view looming or shaking object motions relative to other types of motions, likely because these particular adult gestures highlight the object, bringing it into the foreground of the child's attention (Matatyaho & Gogate, 2008; Matatyaho-Bullaro *et al*., 2014).

Over time, at least in the Western families studied in this research, adults tailor their actions to the developmental level of the infant, similar to their changing use of IDS. Synchronizing object movement with labeling is extremely common at the earliest stages of word learning, when infants lack alternative tools for detecting word-to-world relations. As children progress from the prelexical (5 to 8 months) to early-lexical period (9 to 17 months) and from the early-lexical to advanced-lexical stage (21 to 30 months), mothers use this method less and less (Gogate, Bahrick, & Watson, 2000). By the advanced-lexical stage, toddlers use subtle social cues (e.g., eye gaze) as well as more sophisticated (and less infant-directed) pragmatic cues. For example, 27 month olds will differentially map a speaker's novel label to an action or to an object depending on the prior (rather than concurrent) actions of the speaker (Tomasello & Akhtar, 1995). Thus non-linguistic input plays a critical, albeit shifting role in word meaning disambiguation across development.

### *What is* not *required for word learning?*

Despite a wealth of research supporting the role of eye gaze and IDA in language development (Baldwin, Bill, & Ontai, 1996; Carpenter *et al.*, 1998; Tomasello & Akhtar, 1995), vision is clearly not a prerequisite for lexical acquisition. Blind children learn words much the same as their sighted counterparts, including visual terms like *look* and *see* (Landau & Gleitman, 1985), even if the meanings they store for these lexical items are somewhat distinct from the meanings acquired by sighted children.

Lexical acquisition is most often discussed in terms of spoken language, but speech and hearing are not necessary for language development, either. Stromswold (1994) tested an anarthric child who could not produce speech and showed that he, too, comprehended many words and sentences. Furthermore, children learn signed languages just as easily as spoken languages (with the right input), regardless of whether they can hear or not. Indeed, hearing infants of hearing parents come prepared to find the "phonemes" in infant-directed *sign* at four months of age, an ability they lose by 14 months of age (Palmer *et al.*, 2012). Remarkably, the milestones for lexical acquisition are very similar for children learning signed languages and spoken languages (Bonvillian, Orlansky, & Novack, 1983; Schick, 2010). In the following section, we describe these milestones and discuss the implication, that certain aspects of word learning are universal.

## The timeline of lexical acquisition

Across the globe, children reach major vocabulary milestones at the same time and show similar patterns in learning words. Whether children are learning French or Chinese, they tend to comprehend more words than they can produce. Furthermore, children show a tendency to learn nouns before they learn verbs—even in what are termed *verb-friendly* languages (Bornstein *et al.*, 2004; Waxman *et al.*, 2013), in which verbs can appear alone or at the ends of sentences.

### ***Major milestones***

Although there is some variation among individuals and among languages, children typically experience a remarkably similar trajectory of lexical growth (Bleses *et al.*, 2008). It takes about 12 months for children to produce their first word, but from then onward, their expressive vocabulary grows to approximately 50 words in the following six months (O'Grady & Archibald, 2010). The lexicon rapidly expands after this point, during a period often referred to as the *vocabulary spurt* (Fernández & Cairns, 2010). Some research suggests this spurt may simply be a by-product of learning words, of varying difficulty, in parallel (McMurray, 2007). However, specialized learning processes do emerge, and a large body of evidence suggests that word learning accelerates across development because children discover regularities in referential mappings (e.g., the shape bias, Landau, Smith, & Jones, 1988) and increasingly make use of a variety of information when learning new words (Hollich *et al.*, 2000). This growth continues into adulthood, by which point most people know about 60,000 words. Table 24.1 summarizes some well-established milestones in lexical acquisition (Bornstein & Hendricks, 2012; Hollich *et al.*, 2000; O'Grady & Archibald, 2010).

## *Comprehension before and greater than production*

As Table 24.1 suggests, comprehension precedes and exceeds production throughout the early years of lexical development (Hirsh-Pasek & Golinkoff, 1996; O'Grady & Archibald, 2010). Some words are understood as early as six months, before any words can be produced (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 2012). Even once production begins, the rate of word learning for comprehension is nearly twice that of production (Benedict, 1979). Bornstein and Hendricks (2012) found that comprehension consistently exceeds production among two to nine year olds in 16 under-researched developing nations, indicating that this developmental pattern continues throughout childhood and may be universal.

The asymmetry between receptive and expressive vocabulary has sparked controversy over the potential independence of these two aspects of language. To explore this possibility, Gershkoff-Stowe and Hahn (2013) studied incremental changes in word knowledge for 12 novel objects over three weeks, in both two-year-old children and adults. The authors found a comprehension advantage in both age groups, but there was no clear pattern as children progressed from

**Table 24.1** Milestones of lexical acquisition

| *Age* | *Milestone* |
|---|---|
| 6–9 months | Understand first words |
| 12 months | Produce first words, understand 50 words |
| | Learn to produce 2 new words per week |
| 18 months | Produce 50 words, understand 150 words |
| | Learn to produce 10 new words per day |
| 6 years | Produce and understand 14,000 words |
| | Learn up to 20 new words per day |
| 17 years | Produce and understand 60,000 words |

comprehension to production. In other words, any given word need not be part of the child's receptive vocabulary *before* entering the expressive vocabulary (Gershkoff-Stowe & Hahn, 2013). Rather, comprehension and production are distinct processes with different requirements. Comprehension involves recognizing the target word, but the meaning of a recognized word can sometimes be inferred from context without retrieval from memory. Word production, on the other hand, requires the active generation (i.e., retrieval) of words to match a communicative intention, as well as the motivation to speak (Bock, 1995; Woodward, Markman, & Fitzsimmons, 1994). These processes likely share an overlapping knowledge store (Gershkoff-Stowe & Hahn, 2013), but word production seems to develop on its own timescale, somewhat independent of the earlier-developing comprehension. Mayor and Plunkett (2014) found that toddlers learning English, Dutch, Norwegian, and German all tend to understand the same set of words, but expressive vocabulary is highly variable among children (after the first 100 words), supporting the view that these two types of word knowledge progress differently during development.

## *The noun bias*

Children have been observed to learn more nouns than other types of words (Gentner, 1982; Goldin-Meadow, Seligman, & Gelman, 1976; Waxman *et al.*, 2013), but there has been some debate about the potential universality of this tendency (Tardif, Gelman, & Xu, 1999). Bornstein and colleagues (2004) found that children learning Spanish, Dutch, French, Hebrew, Italian, and Korean tend to exhibit a noun bias in expressive vocabulary. Still, certain environmental factors that vary substantially around the world may affect the strength of the noun bias in different linguistic communities. Goldfield (2000) reports, for example, that parents in New England elicit more nouns from their children than verbs and use verbs to elicit actions rather than speech. This suggests that children may understand many more verbs than they produce, and that the way parents use speech to interact with their children influences what types of words children tend to produce (Benedict, 1979; Goldfield, 2000; Waxman *et al.*, 2013). Korean (Choi & Gopnik, 1995) and Mandarin (Tardif, 1996) use verbs more frequently and in more prominent sentence locations than in English. Despite these differences, the noun bias is retained in these so-called *verb-friendly* languages (Waxman *et al.*, 2013; Bornstein *et al.*, 2004; Imai *et al.*, 2008), suggesting that nouns have a universally privileged status in lexical acquisition.

What causes nouns to be learned earlier and more easily? Gentner (1982) suggested that nouns are learned first because their meanings are easier to carve from the ever-changing world. Maguire, Hirsh-Pasek, and Golinkoff (2006) augmented this explanation, suggesting that all words lie on a continuum of abstractness, termed the SICI (shape, individuation, concreteness, and imageability) continuum (see Figure 24.1). SICI scores reflect the difficulty of learning a word based on four factors that have been discussed in the literature: the consistency of the referent's *shape*, the ease with which the referent concept can be *individuated*, the extent to

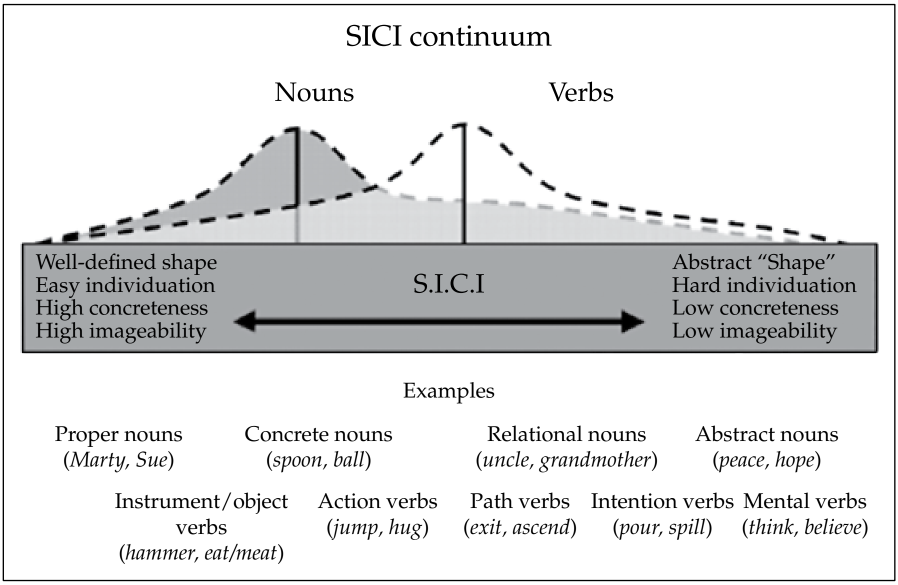



**Figure 24.1** The SICI Continuum. 'SICI' is an acronym for four factors (shape, individuation, concreteness, and imageability) that contribute to the ease or difficulty of learning nouns and verbs. The concepts that these words represent lie on a continuum defined by the reliability of the concept's shape, the ease with which the concept can be individuated from other items, the concreteness of the concept to sensory systems, and the degree to which the word elicits a mental image. Although nouns typically precede verbs in vocabulary acquisition, this pattern is a by-product of the SICI continuum. Reproduced, with permission, from Maguire, Hirsh-Pasek, and Golinkoff (2006) and Oxford University Press, USA.

which the referent is *concrete* to the senses, and the facility with which the word evokes a mental *image*. Although some verbs, such as *jump*, involve a consistent "shape" of motion and are easily imageable, and although some nouns are extremely opaque (e.g., *peace*), the average verb is more abstract (i.e., has a higher SICI score) than the average noun.

A number of sources support the accuracy of the SICI criteria in describing word difficulty. Shape consistency determines whether children will learn and extend both nouns (Landau, Smith, & Jones, 1988) and verbs (Golinkoff *et al.*, 2002) to other category members. Landau, Smith and Jones (1988) demonstrated a *shape bias* for extending count nouns: children easily extended a novel label to objects that had the same shape as the established referent (regardless of size or texture differences), but tended not to use the same label for objects of different shapes that had the same size or texture. For example, golf balls and tennis balls differ in size and texture, but both belong to the category of *ball* because of their spherical shape. This is untrue of things like tennis balls and ducklings, which have a similar size and a soft texture, but do not share a common label.

Additionally, concreteness of words predicts learnability: although infants cannot identify videos depicting abstract words like *wet* and *all-gone* until 10 to 14 months, six-month-old infants are already capable of recognizing pictures of several concrete words (e.g., *hand, banana*) in a similar task (Bergelson & Swingley, 2012, 2013). Further, imageability is one of the best predictors of age of acquisition for both nouns and verbs among English-learning children (Bird, Franklin, & Howard, 2001), and Ma and colleagues (2009) found that the increased imageability of Chinese compared to English verbs contributes to their being learned earlier. Thus, accruing evidence suggests that the noun bias may be an epiphenomenal by-product of the learnability of words, based on multiple dimensions of abstractness.

### Individual differences

Although there are some general milestones and patterns in word learning, well-known individual differences abound. Children who receive less input generally learn fewer words and tend to learn these words more slowly (Hart & Risley, 1995). Bilingual children might trail slightly behind in reaching milestones in either of their two languages (Hoff *et al.*, 2012), but combining the number of words known in both languages reveals that their overall vocabularies are as large as their monolingual peers' (Hoff *et al.*, 2012; see also Byers-Heinlein and Lew-Williams in this volume for a more detailed review of bilingual vocabulary development). Bilingualism is also associated with certain advantages in cognitive flexibility, even in infants (Bialystok & Viswanathan, 2009; Kovács & Mehler, 2009, but see Paap & Greenberg, 2013).

Nelson (1973) noted that individual variation may also result from children following one of two possible paths as they begin learning words. One group, the *referential learners*, fill their early lexicon with names for objects, such as *ball* and *milk*. More socially attuned children, called *expressive learners*, instead master non-referential, communicative words early on, such as *hi* and *want*. The expressive learners tend to reach the 50-word milestone slightly later than the referential learners (Nelson, 1973).

## Models of word learning: Solving the mapping problem

A heated debate surrounds the mechanisms behind word learning across development. In early word learning, theorists ask whether lexical acquisition is purely associationistic or whether it involves true referential learning via *fast mapping*—making a snap decision about the meaning of a novel word based on whatever information is available at the time of first exposure. Models of word learning after this initial "novice" phase diverge even further, with different researchers pointing to either perceptual, social, or linguistic information as the dominant force behind later lexical acquisition. In the last three decades, the field seems to have converged on a hybrid view, suggesting that all three types of cues are in play during advanced word learning. Hybrid models take a broader perspective, examining how the processes supporting lexical development change over the first few years of life (Hollich *et al.*, 2000).

## *Early word learning*

Although children eventually make use of complex social and linguistic cues to disambiguate word meaning, research suggests that they might not be able to recruit all these types of input from the outset. At first, they focus on perceptual salience as the main source of word meaning (Hollich *et al.*, 2000; Brandone *et al.*, 2007). Even with this narrow focus, two competing theoretical models propose distinct mechanisms for the acquisition of first words: cross-situational models and single-hypothesis models.

Cross-situational models propose that infants are robust statistical word learners, using similar methods to learn word meanings as they do to segment and identify words in the speech stream. To do this, infants must keep track of all the possible referents for a word, gleaned from experience within and across situations. At any given time, a child's representation of a word is considered to be manifold: first, the representation includes a mapping of the word to a single referent based on which referent co-occurs with the word most frequently; second, the representation requires partial knowledge of how frequently other referents have co-occurred with that word (Smith & Yu, 2008; Yu & Smith, 2007; Yurovsky *et al.*, 2014).

The single-meaning hypothesis offers an alternative to cross-situational models. According to this hypothesis, children fast map one and only one hypothetical meaning in any given word learning situation (Golinkoff *et al.*, 1992; Medina *et al.*, 2011; Trueswell *et al.*, 2013); no other possible meanings are stored, even when the word learning occurs in a highly ambiguous situation. Early instantiations of single-meaning hypothesis models postulated that a word meaning hypothesis was maintained until it was disconfirmed by experience, at which point a new hypothesis was posited and the old discarded (Medina *et al.*, 2011; Trueswell *et al.*, 2013).

However, behavioral evidence and computational modeling have led researchers to alter the single-meaning hypothesis. The revised version, termed *Pursuit*, proposes that disconfirmed hypotheses are maintained alongside new hypotheses for some time (Stevens *et al.*, 2017). After all, some words (i.e., homophones) can have multiple meanings (e.g., *bear*, *date*), and there must be a way for children to learn these. The Pursuit model takes a step toward acknowledging the infant's statistical learning abilities (Stevens *et al.*, 2017). Repeated encounters with a word that support the original (fast-mapped) meaning are thought to increase the child's confidence in this hypothesis. If new encounters suggest a different meaning instead, a new hypothesis is created and assigned its own confidence level (based on how informative the learning situation is), while confidence in the original hypothesis decreases. Thus, at any given time, the child's representation of a word includes the most probable hypothesis as well as hypotheses formed during prior exposures to the word.

The key difference between cross-situational and single hypothesis models is whether multiple possible meanings of a word (based on word-object co-occurrences) are retained, or whether the child maintains only a limited set of hypothetical word meanings (one from each experience with the word). Although cross-situational models seem to avoid errors by maintaining all competing possibilities, they crowd the hypothesis space for each word, requiring an enormous amount of memory for each entry in the lexicon (Stevens *et al.*, 2017). Single-meaning or Pursuit models, on the other hand, may be more prone to error due to mistakes in fast mapping.

So which theory is supported by the data? Co-occurrence statistics can be used to determine the meanings of novel words in constrained experimental settings, as proposed by cross-situational models (Vouloumanos & Werker, 2009; Yu & Smith, 2011), but more naturalistic studies are necessary to test whether this method of word learning works in real-life situations (Smith, Suanda, & Yu, 2014). For example, the human simulation paradigm tests adults' ability to learn a novel object label by watching videos of parent-child interactions that are muted. This simulates the vast ambiguity of natural labeling events to determine whether or not cross-situational experience with a word is sufficient for everyday word learning (Medina *et al.*, 2011; Yurovsky, Smith, & Yu, 2013). Yurovsky, Smith, and Yu (2013) found that adults perform significantly better on every subsequent trial, even if their hypothesized meaning on the preceding trial was incorrect, indicating an effect of other object co-occurrences. However, each trial in this study was only compared to the trial immediately before it, not to all prior trials. It is therefore possible that only participants who developed the correct meaning hypothesis at some earlier point (not necessarily the trial immediately prior) were eventually successful on a later trial. Indeed, Koehne, Trueswell, and Gleitman (2013) found this to be true, suggesting that participants must have retained the correct hypothesis from a previous mapping and did not simply happen upon it with repeated exposure to the word. This finding lends support to the Pursuit hypothesis as a more accurate model of perceptual word learning.

The growing acceptance that word meanings are learned probabilistically and gradually (i.e., partially; see Yurovsky *et al.*, 2014) across situations rather than instantaneously originates in these early word learning models and signifies a critical step in understanding the word learning process. It is also vital that word learning models take memory into account, as recent evidence demonstrates that the retention of fast mapped labels is remarkably poor (Bion, Borovsky, & Fernald, 2013; Horst & Samuelson, 2008; Twomey, Ranson, & Horst, 2014, but see Zosh, Brinster, & Halberda, 2013). Further, memory for newly learned words follows a curvilinear pattern, with rapid forgetting early on, and slower rates of forgetting as time passes (Vlach & Sandhofer, 2012). Perhaps counterintuitively, forgetting is crucial for successful word learning. Lexical representations for frequently experienced words are reactivated and strengthened with each subsequent experience of the word (Wojcik, 2013), but if a word-object pairing is not re-experienced, as might happen for erroneous mappings or rare words, the pairing is never retrieved (i.e., reactivated) from memory and is forgotten over time (Vlach & Sandhofer, 2012). Thus forgetting is necessary to weed out incorrect mappings and to extend object mappings to more general object categories (Vlach & Sandhofer, 2012). Importantly, the ability to retain fast mapped word meanings increases with language experience (Bion *et al.*, 2013), perhaps in part because later mappings are based on more than just perceptual information.

## *Word learning beyond the novice phase*

While early word learning relies on perceptual cues as the main source of information about word meaning, there are differing views on the role of these and other types of cues in *later* lexical acquisition. In addition to perceptual information, social and linguistic cues have been identified as potential indicators of meaning. Word learning models based on all three of these types of cues have found supporting evidence in experimental and observational research.

Evidence shows that perceptual cues remain important beyond the first year of life, with 18 month olds learning a novel object label more easily when the object has a consistent location than when its location varies (Benitez & Smith, 2012). Perceptual models of later word learning assume that social and linguistic cues simply function to increase or decrease the perceptual salience of possible word referents (Smith, 2000). In support of this view, Yoshida and Smith (2005) found that two year olds are more likely to learn a novel (i.e., non-native) semantic category when the linguistic information provided is redundant with perceptual cues. At this age, children can even learn a novel word when labeling occurs in the absence of the object referent, as long as labeling coincides with visual cues to the object's previous location (Baldwin, 1993; Smith, 2005). Three to four year olds and adults alike may learn new words through Bayesian inference, a type of statistical learning that requires general knowledge of word-to-world mapping and the ability to reweight the likelihood that fast-mapped hypotheses are correct, based on new experiences (Xu & Tenenbaum, 2007). In this way, Bayesian models of word learning straddle the cross-situational and Pursuit hypotheses, but still rely primarily on perceptual cues.

One major criticism of these perceptual models is that they seem to assume a nearly infinite number of tracked associations between words and their possible referents, as well as an infinite number of probabilistic calculations that must be computed to determine the correct referent of a given word. Arguing against this view, Yu and Smith (2012) suggest that word learning events are not as ambiguous as we (adults) believe, because children are visually selective in ways that adults are not. Not only do children move an object of interest so that it dominates their visual field, but they are more likely to learn the name for this object if their parents label it during a moment of visual focus (Yu & Smith, 2012). Still, perceptual models cannot easily explain how children map words to referents that are more abstract and lack perceptual salience (e.g., most verbs). The current evidence for these models comes from studies of noun learning, which generally involve mapping words to concrete, highly imageable referents (see Figure 24.1), for which perceptual cues are highly informative.

Another category of word learning models take a social-pragmatic approach. These models emphasize the importance of social-cognitive skills and socially contingent parent-child interactions (Tamis-LeMonda, Kuchirko, & Song, 2014; Tomasello, 2000). Word learning is thought to be facilitated by the child's understanding that language is used to exchange socially contextualized meanings, in conjunction with nonverbal communicative interaction. Additionally, these models suggest that social influences *gate*, or restrict, word learning processes, thus circumventing the unlimited number of calculations implicit in perceptual models (Kuhl, 2007). Studies have shown that even when a novel object is visually available at the time of labeling, this word-object mapping is learned more easily if the speaker and infant are jointly attending to the object (Baldwin, Bill, & Ontai, 1996; Bannard & Tomasello, 2012; Tomasello & Farrar, 1986). Booth, McGregor, and Rohlfing (2008) further demonstrated that word learning in 2.5 year olds could be enhanced by providing redundant socio-pragmatic cues, and this improvement resulted from increased attention to the communicative context rather than increased attention to the target referent.

Despite these findings, socio-pragmatic word learning models cannot explain the whole of lexical acquisition. Frank, Tenenbaum, and Fernald (2013) showed that socio-pragmatic cues alone are not reliable indicators of word meaning. Rather, these cues must be probabilistically combined to inform word reference (Frank, Tenenbaum, & Fernald, 2013). Moreover, socio-pragmatic approaches do not explain how infants who lack a repertoire of socio-cognitive skills (such as children who fall on the autistic spectrum) are able to learn words (e.g., Parish-Morris *et al.*, 2007), nor do they account for later word learning in the absence of social cues (e.g., when reading a text).

A third and final class of word learning models are the linguistic models, which attribute the child's later lexical acquisition to the developing knowledge of her native language's structure. Once a foundational vocabulary (of mostly basic-level nouns) is acquired, young children begin to use these "easy" words to learn new "hard" (i.e., less perceptually available) ones through syntactic bootstrapping (Gleitman *et al.*, 2005). For example two year olds infer that a novel verb in a two-argument (transitive) construction (e.g., "Look! The *duck* is gorping the *rabbit*!") is causal while a novel verb in a single-argument (intransitive) structure (e.g., "Look! *They* are gorping!") must refer to a self-caused act (Naigles, 1990; Hirsh-Pasek & Golinkoff, 1996). Similarly, Syrett and Lidz (2010) revealed that by 30 months, children even use syntactic bootstrapping to determine the meanings of novel adjectives, based on the type of adverbial modifier they appear with. Children used intensifiers (e.g., *too*) as a cue to *relative* adjective meaning (e.g., *small*) and proportional modifiers (e.g., *totally*) as a cue to *absolute* adjective meaning (e.g., *dry*).

Of course, linguistic models are also limited by their specificity. No single sentence is a reliable source for word meaning, and structural context alone is not enough to form an accurate mapping. Rather, children must experience a word in multiple sentential contexts and receive additional non-linguistic cues to a novel word's meaning (Rispoli, 1995; Yuan, Fisher, & Snedeker, 2012). Similar to socio-pragmatic word learning models, linguistic models fail to explain how word learning occurs prior to extensive language experience.

## *Modeling word learning as a complex developmental process*

With evidence supporting the importance of perceptual, social, *and* linguistic information, word learning may be better explained by hybrid models that emphasize the weighting of multiple cues. These multifaceted models are empirically testable and have the added advantage of allowing the field to examine the changing nature of word learning over time.

The first hybrid theory of lexical acquisition to acknowledge the complexity of the word learning task was the Emergentist Coalition Model (ECM; Golinkoff & Hirsh-Pasek, 2006; Hollich *et al.*, 2000). Initially proposed in response to competing theories that posited a single word learning mechanism, the ECM changed the question from what process underlies word learning to how the processes underlying word learning change and interact across development. Hybrid theories of lexical development have now become the norm, with many adopting the same basic views as the ECM (Booth & Waxman, 2008; Caza & Knott, 2012; Namy, 2012).

The ECM is founded on three tenets: 1) children are sensitive to multiple cues, including perceptual, social, and linguistic sources of information from the outset 2) there is a differential weighting of these cues over time such that perceptual cues are more salient at the start of word learning, and 3) children construct word learning principles from the combination of internal biases and attention to these interactive and weighted cues (Hollich *et al.*, 2000). These claims have been experimentally tested by using eye gaze, pointing, and enthusiastic speech to label only one of two novel objects in children's immediate view—one interesting (e.g., brightly colored) and the other boring (e.g., dull in appearance). When the interesting object is labeled (the *coincident* condition), perceptual and social cues converge. However, when the boring object is labeled (the *conflict* condition), children must override their natural preference for the interesting (perceptually salient) object to map the label correctly—that is, children must weight social cues over perceptual ones (Pruden *et al.*, 2006). While 10 month olds map the novel word to the interesting object regardless of condition, 12 month olds do not; they map successfully in the coincident condition, but fail to form any mapping in the conflict condition (Hollich *et al.*, 2000). Finally, by 19–24 months, children are successful at mapping in both conditions. These results implicate a gradual shift in the weighting of social cues with respect to perceptual ones. This shift may be explained, in part, by infants' accrual of multifaceted experiences with adults who respond contingently and appropriately to their pre-linguistic object-directed behaviors (e.g., vocalizing, pointing, eye gaze), reflecting the infants' perceptual interest (Goldstein & Schwade, 2009; Wu & Gros-Louis, 2014; Yu & Smith, 2012).

In addition to the transition toward socially cued mapping, the ECM posits a shift in the use of linguistic cues. This is reflected both in infants' increasing selectivity of using words (and not other sounds) as symbolic representations for objects (see discussion above on how language is special) and in the developmental changes in how linguistic information is utilized for learning words. The ECM argues that syntactic bootstrapping is available early, but not dominant when faced with competing cues to word meaning. Further, social and linguistic sources of information are sometimes leveraged against one another. Social cues such as eye gaze seem to be more critical for learning nouns than other word classes (Bergelson & Swingley, 2013), while linguistic cues are especially useful for verb learning, likely because verb referents are more abstract or fleeting (Gleitman *et al.*, 2005; Maguire, Hirsh-Pasek, & Golinkoff, 2006). Relatedly, social cues may be weighted more heavily earlier in the word learning process than linguistic cues (Caza & Knott, 2012; Hollich *et al.*, 2000).

In some cases, social and linguistic cues are integrated in a single word-learning situation, with neither type of information necessarily dominating over the other. Grassmann, Stracke, and Tomasello (2009) tested whether two year old's use of the *mutual exclusivity bias*, the tendency to map novel words onto referents lacking a known label, was influenced by social information. When an experimenter excitedly uttered a novel word in reference to an object she had never seen before, children, as expected, mapped the label to the object. However, if the adult and child jointly played with the object first, the subsequent excited labeling event was much less likely to lead children to this mapping. This indicates that children make use of common ground (social-pragmatic information) to determine whether the mutual exclusivity bias (a linguistic cue) will be useful in a given situation. In this way, social and linguistic cues may be especially useful in concert with one another during later word learning. The next wave of research in this area will involve testing hybrid models longitudinally to tease apart children's progressive reweighting of different cues to word meaning.

## Conclusion

Since Roger Brown's *Names for Things*, we have come to understand a great deal about the word learning process. We know that children, as master statisticians, can segment the fluid stream of sounds and events into coherent units. We even have some purchase on the mapping problem, which is compounded in the case of verbs and other relational and abstract words. Future research will surely continue to explore the mapping problem, but must do so in a way that nests the problem in a developmental and ecological framework. Gone are the days when researchers could seek simplistic single-mechanism answers to the "how" of lexical development. Any future solutions must embrace the complexity of the problem, including multiple inputs (i.e., linguistic and nonlinguistic), as well as the child's contribution in segmentation, symbolization, and the changes to mapping processes that occur over time. In short, the problems that plagued Plato remain contentious today; and P. Bloom was right—word learning is not a simple matching of word to world, but rather a window onto a multipronged cognitive problem.

## REFERENCES


Akhtar, N. (2005). The Robustness of learning through overhearing. *Developmental Science,* 8(2), 199–209. doi: 10.1111/j.1467-7687.2005.00406.x.

Akhtar, N., Dunham, F., & Dunham, P. J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language,* 18(1), 41–49. doi: 10.1017/S0305000900013283.

Andruski, J. E., & Kuhl, P. K. (1996). The acoustic structure of/i/,/u/, and/a/in mothers' speech to infants and adults. *The Journal of the Acoustical Society of America,* 100(4), 2725–2726. doi: 10.1109/ICSLP.1996.607913.

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology,* 29(5), 832–843. doi: 10.1037/0012-1649.29.5.832.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development,* 72(3), 708–717. doi: 10.1111/1467-8624.00310.

Baldwin, D. A., Bill, B., & Ontai, L. L. (1996). Infants' tendency to monitor others' gaze: Is it rooted in intentional understanding or a result of simple orienting? *Infant Behavior and Development,* 19, 270. doi: 10.1016/S0163-6383(96)90324-0.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, 67(6), 3135–3153. doi: 10.2307/1131771.

Bannard, C. & Tomasello, M. (2012). Can we dissociate contingency learning from social learning in word acquisition by 24-month-olds? *PloS One*, 7(11), e49881. doi: 10.1371/journal.pone.0049881.

Bates, E. & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates, (Eds.), *The Cross-linguistic Study of Sentence Processing*, (pp. 3–73). Cambridge, England: Cambridge University Press.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language,* 6(2), 183–200. doi: 10.1017/S0305000900002245

Benitez, V. L., & Smith, L. B. (2012). Predictable locations aid early object name learning. *Cognition,* 125(3), 339–352. doi: 10.1016/j.cognition.2012.08.006.

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America,* 109(9), 3253–3258. doi: 10.1073/pnas.1113380109.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, 127(3), 391-397. doi: 10.1016/j.cognition.2013.02.011.

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition,* 112(3), 494–500. doi: 10.1016/j.cognition.2009.06.014.

Bion, R. A. H, Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition,* 126(1), 39–53. doi: 10.1016/j.cognition.2012.08.008.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, &*

*Computers*, 33(1), 73–79. doi: 10.3758/BF03195349.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). The Danish communicative dvelopmental inventories: Validity and main developmental trends. *Journal of Child Language*, 35(3), 651–669. doi: 10.1017/S0305000907008574.

Bloom, P. (2002). *How Children Learn the Meanings of Words*. Cambridge, Mass: MIT Press.

Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition: Speech, language, and communication*, Vol. 11, (pp. 181–216). Orlando, FL: Academic Press.

Bonvillian, J. D., Orlansky, M. D., Novack, L. L., Bonvillian, J. D., Orlansky, M. D., & Lazin Novack, L. (1983). Developmental milestones: Sign language acquisition and motor development. *Child Development*, 54(6), 1435–1445. doi: 10.2307/1129806.

Booth, A. E., McGregor, K. K., & Rohlfing, K. J. (2008). Socio-pragmatics and attention: Contributions to gesturally guided word learning in toddlers. *Language Learning and Development*, 4(3), 179–202. doi: 10.1080/154754408021 43091.

Booth, A. E, & Waxman, S. R. (2008). Taking stock as theories of word learning take shape. *Developmental Science*, 11(2), 185–194. doi: 10.1111/j.1467-7687.2007.00664.x.

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., Pêcheux, M.-G., Josette Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75(4), 1115–1139. doi: 10.1111/j.1467-8624.2004.00729.x

Bornstein, M. H., & Hendricks, C. (2012). Basic language comprehension and production in >100,000 young children from sixteen developing nations. *Journal of Child Language*, 39(4), 899–918. doi: 10.1017/S0305000911000407.

Bornstein, M. H., Tamis-Lemonda, C. S., Hahn, C.-S., & Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, 44(3), 867–874. doi: 10.1037/0012-1649.44.3.867.

Bortfeld, H., Morgan, J. L., Michnick Golinkoff, R., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science*, 16(4), 298–304. doi: 10.1111/j.0956-7976.2005.01531.

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': Modifications in mothers' infant-directed action." *Developmental Science*, 5(1), 72–83. doi: 10.1111/1467-7687.00211.

Brandone, A. C, & Hirsh-Pasek, K. (2007). Action speaks louder than words: Young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development*, 78(4), 1322–1342. doi: 10.1111/j.1467-8624.2007.01068.x

Brandone, A. C., Salkind, S. J., Michnick Golinkoff, R., & Hirsh-Pasek, K. (2006). Language development. In G. G. Bear & K. M. Minke (Eds.), *Children's Needs: Development, Prevention, and Intervention*, (pp. 499–514). Bethesda, MD: National Association of School Psychologists.

Bresnan, J. (1978). A realistic transformational grammar. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality*. Cambridge, Mass: MIT Press.

Broesch, T. L., & Bryant, G. A. (2015). Prosody in infant-directed speech is similar across western and traditional cultures. *Journal of Cognition and Development*, 16(1), 31–43. doi: 10.1080/15248372.2013.833923.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology*, 55, 1–5. doi: 10.1037/h0041199.

Brown, R. (1958). *Words and Things*. New York: The Free Press.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, 121(1), 127–132. doi: 10.1016/j.cognition.2011.06.010.

Carpenter, M., Nagell, K., Tomasello, M., & Butterworth, G. (1998). Social cognition: Joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), i–174.

Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences of the United States of America*, 110(28), 11278–11283. doi: 10.1073/pnas.1309518110.

Caza, G. A., & Knott, A. (2012). Pragmatic bootstrapping: A neural network model of vocabulary acquisition. *Language Learning and Development*, 8(2), 113–135. doi: 10.1080/15475441.2011.581144.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, 22(3), 497–529. doi: 10.1017/S030500090 0009934.

Colunga, E., & Smith, L. B. (2008). Knowledge embedded in process: The self-organization of skilled noun learning. *Developmental Science*, 11(2), 195–203. doi: 10.1111/j.1467-7687.2007. 00665.x.

Dougherty, D. P. (2000). *How To Talk To Your Baby: A Guide To Maximizing Your Child's Language and Learning Skills*. New York: Penguin Putnam, Inc.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104–113. doi: 10.1037/0012-1649.20.1.104.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants." *Journal of Child Language*, 16(3), 477–501. doi: 10.1017/S0305000900010679.

Fernández, E. M. & Cairns, H. S. (2010). *Fundamentals of Psycholinguistics*. Hoboken, NJ: Wiley-Blackwell.

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3- and 4-month-old infants: An advantage of words over tones." *Child Development*, 81(2), 472–479. doi: 10.1111/j.1467-8624.2009.01408.x.

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2013). Nonhuman primate vocalizations support categorization in very young human infants. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), 15231–15235. doi: 10.1073/pnas.1221166110.

Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy*, 9(3), 327–339. doi: 10.1207/s15327078in0903_4

Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development*, 9(1), 1–24. doi: 10.1080/15475441.2012.707101.

Friend, M., & Pace, A. (2011). Beyond event segmentation: Spatial- and social-cognitive processes in verb-to-action mapping. *Developmental Psychology*, 47(3), 867–876. doi: 10.1037/a0021107.

Fromkin, V., Krashen, S., Curtiss, S., Rigler, D., & Rigler, M. (1974). The development of language in genie: A case of language acquisition beyond the "critical period." *Brain and Language*, 1(1), 81–107. doi: 10.1016/0093-934X(74)90027-3.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language Development: Language, Thought and Culture*, Vol. 2, (pp. 301–334). Hillsdale, NJ: Erlbaum.

Gershkoff-Stowe, L., & Hahn, E. R. (2013). Word comprehension and production asymmetries in children and adults. *Journal of Experimental Child Psychology*, 114(4), 489–509. doi: 10.1016/j.jecp.2012.11.005.

Gertner, Y., & Fisher, C. (2012). Predicted errors in children's early sentence comprehension. *Cognition*, 124(1), 85–94. doi: 10.1016/j.cognition.2012.03.010.

Gleitman, L. R., Cassidy, K., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, 1(1), 23–64. doi: 10.1207/s15473341lld0101.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, 71(4), 878–894. doi: 10.1111/1467-8624.00197.

Gogate, L. J., Bolzani, L. H., & Betancourt, E. A. (2006). Attention to maternal multimodal naming by 6-to 8-month-old infants and learning of word-object relations. *Infancy*, 9(3), 259–288. doi: 10.1207/s15327078in0903_1

Goldfield, B. A. (2000). Nouns before verbs in comprehension vs production: The view from pragmatics. *Journal of Child Language*, 27(3), 501–520.

Goldin-Meadow, S., Seligman, M. E. P., & Gelman, R. (1976). Language in the two-year old. *Cognition*, 4(2), 189–202. doi: 10.1016/0010-0277(76)90004-4.

Goldstein, M. H. & Schwade, J. A. (2009). From birds to words: Perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *The Oxford Handbook of Developmental and Comparative Neuroscience*, (pp. 708–729). New York, Oxford University Press. doi: 10.1093/oxfordhb/9780195314731.013.0035.

Golinkoff, R. M. & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language*, 22(3), 703–726. doi: 10.1017/S0305000900010011.

Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., Maguire, M. J., & Hennon, E. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology*, 38(4), 604–614. doi: 10.1037//0012-1649.38.4.604.

Golinkoff, R. M. & Hirsh-Pasek, K. (1999). *How Babies Talk: The Magic and Mystery of Language Acquisition*. New York: Dutton/Penguin.

Golinkoff, R. M., & Hirsh-Pasek, K. (2006). Baby wordsmith: From associationist to social sophisticate. *Current Directions in Psychological Science*, 15(1), 30–33. doi: 10.1111/j.0963-7214.2006.00401.x

Golinkoff, R. M. & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences*, 12(10), 397–403. doi: 10.1016/j.tics.2008.07.003.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, 28(1), 99–108. doi: 10.1037/0012-1649.28.1.99.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, 21(1), 125–155. doi: 10.1017/S0305000900008692

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 18(3), 254–260. doi: 10.1111/j.1467-9280.2007.01885.x.

Graf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy*, 18(5), 797–824. doi: 10.1111/infa.12006.

Grassmann, S., Stracke, M., & Tomasello, M. (2009). Two-year-olds exclude novel objects as potential referents of novel words based on pragmatics. *Cognition,* 112(3), 488–493. doi: 10.1016/j. cognition.2009.06.010.

Hart, B. & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Brookes.

Hirsh-Pasek, K. & Golinkoff, R. M. (1996). *The Origins of Grammar: Evidence From Comprehension.* Cambridge, MA: MIT Press.

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, 26(3), 269–286. doi: 10.1016/ S0010-0277(87)80002-1.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language,* 39(1), 1–27. doi: 10.1017/ S0305000910000759.

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., Hennon, E., *et al.* (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65(3), i–135.

Horst, J. S. & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy,* 13(2), 128–157. doi: 10.1080/15250000701795598.

Houston, D., Santelmann, L., & Jusczyk, P. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes*, 19(1), 97–136. doi: 10.1080/01690960344000143.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, 11(6), F31–F39. doi: 10.1111/j.1467-7687.2008.00768.x.

Imai, M., Haryu, E., Hirsh-Pasek, K., Li, L., Okada, H., Golinkoff, R. M., & Shigematsu, J., *et al.* (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development*, 79(4), 979–1000. doi: 10.1111/j.1467-8624.2008.01171.x.

Jusczyk, P. W. (1986). Representation of speech by infants. *The Journal of the Acoustical Society of America*, 79(S1), S52–S53. doi: 10.1121/1.2023271.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39(3),159–207. doi: 10.1006/cogp.1999.0716.

Koehne, J., Trueswell, J. C., & Gleitman, L. R. (2013). Multiple proposal memory in observational word learning. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, (pp. 805–810). Austin, TX: Cognitive Science Society.

Kooijman, V., Junge, C., Johnson, E. K., Hagoort, P., & Cutler, A. (2013). Predictive brain signals of linguistic development." *Frontiers in Psychology*, 4(25). doi: 10.3389/fpsyg.2013.00025.

Kovács, A. M. & Mehler, J. (2009). Flexible learning of multiple speech structures in nilingual infants. *Science*, 325, 611–612. doi: 10.1126/science.1173947.

Kuhl, P. K. (2007). Is speech learning "gated" by the social brain? *Developmental Science*, 10(1), 110–120. doi: 10.1111/j.1467-7687.2007.00572.x.

Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development*, 3(3), 179–197. doi: 10.1080/15475440701360168.

Landau, B. & Gleitman, L. R. (1985). *Language and Experience: Evidence from the*

*Blind Child*. Cambridge, MA: Harvard University Press.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3(3), 299–321. doi: 10.1016/0885-2014(88)90014-7.

Levine, D., Hirsh-Pasek, K., Pace, A., Golinkoff, R. M. (2017). A goal bias in action: The boundaries adults perceive in events align with sites of actor intent. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(6), 916–927. doi: 10.1037/xlm0000364.

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word kearning in infant- and adult-directed speech. *Language Learning and Development*, 7(3), 185–201. doi: 10.1080/15475441.2011.579839.

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language*, 36(2), 405–423. doi: 10.1017/S0305000908009008.

MacKenzie, H., Graham, S. A., & Curtin, S. (2011). Twelve-month-olds privilege words over other linguistic sounds in an associative learning task. *Developmental Science*, 14(2), 249–255. doi: 10.1111/j.1467-7687.2010.00975.x.

Maguire, M. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2006). A unified theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R. M. Golnikoff (Eds.), *Action Meets Word: How Children Learn Verbs*. New York: Oxford University Press.

Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science*, 6(5), 314–317. doi: 10.1111/j.1467-9280.1995.tb00517.x.

Martin, A., Onishi, K. H., & Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12 months. *Cognition*, 123(1), 50–60. doi: 10.1016/j.cognition.2011.12.003.

Matatyaho-Bullaro, D. J., Gogate, L., Mason, Z., Cadavid, S., & Abdel-Mottaleb, M. (2014). Type of object motion facilitates word mapping by preverbal infants. *Journal of Experimental Child Psychology*, 118, 27–40. doi: 10.1016/j.jecp.2013.09.010.

Matatyaho, D. J., & Gogate, L. J. (2008). Type of maternal object motion during synchronous naming predicts preverbal infants' learning of word-object relations. *Infancy*, 13(2), 172–184. doi: 10.1080/15250000701795655.

May, L., & Werker, J. F. (2014). Can a click be a word? Infants' learning of non-native words. *Infancy*, 19(3), 281–300. doi: 10.1111/infa.12048.

Mayor, J., & Plunkett, K. (2014). Shared understanding and idiosyncratic expression in early vocabularies. *Developmental Science*, 17(3), 412–423. doi: 10.1111/desc.12130.

McMurray, B. (2007). Defusing vocabulary the childhood explosion. *Science*, 317(5838), 631–631. doi: 10.1126/science.1144073.

McRoberts, G. W. & Best, C. T. (1997). Accommodation in mean F0 during mother–infant and father–infant vocal interactions: A longitudinal case study. *Journal of Child Language*, 24(3), 719–736. doi: 10.1017/S030500099700322X.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, 108(22), 9014–9019. doi: 10.1073/pnas.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17(2), 357–374. doi: 10.1017/S0305000900013817.

Namy, L. L. (2012). Getting specific: Early general mechanisms give rise to domain-specific expertise in word learning. *Language Learning and Development*, 8(1), 47–60. doi: 10.1080/15475441.2011.617235.

Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development*, 69(2), 295–308. doi: 10.2307/1132165.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756–766. doi: 10.1037/0096-1523.24.3.756.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 38(1–2), 1–135.

Newport, E. L. (1975). Motherese: The speech of mothers to young children. Doctoral dissertation, University of Pennsylvania.

O'Grady, W., Archibald, J., Aronoff, M., & Rees-Miller, J. (2010). *Contemporary Linguistics: An Introduction*, 6th ed. New York: Bedford/St. Martin's Press.

Quine, W. V. O. (1960). *Word and Object*. Cambridge, Mass: MIT Press.

Paap, K. R. & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232–258. doi: 10.1016/j.cogpsych.2012.12.002.

Palmer, S. B., Fais, L., Golinkoff, R. M., & Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Development*, 83(2), 543–553. doi: 10.1111/j.1467-8624.2011.01715.x.

Parish-Morris, J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M., & Tager-Flusberg, H. (2007). Children with autism illuminate the role of social intention in word learning. *Child Development*, 78(4), 126501287. doi: 10.1111/j.1467-8624.2007.01065.x.

Pilley, J. W. & Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes*, 86(2), 184–195. doi: 10.1016/j.beproc.2010.11.007.

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: ten-month-olds learn words through perceptual salience. *Child Development*, 77(2), 266–280. doi: 10.1111/j.1467-8624.2006.00869.x.

Rader, N. D. V. & Zukow-Goldring, P. (2012). Caregivers' gestures direct infant attention during early word learning: The importance of dynamic synchrony. *Language Sciences*, 34(5), 559–568. doi: 10.1016/j.langsci.2012.03.011.

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, 17(6), 880–891. doi: 10.1111/desc.12172.

Regier, T. (2005). The emergence of words: Attentional learning in form and meaning. *Cognitive Science*, 29(6), 819–865. doi: 10.1207/s15516709cog0000_31.

Rispoli, M. (1995). Missing arguments and the acquisition of predicate meanings. In W. E. Merriman & M. Tomasella (Eds.), *Beyond Names for Things: Young Children's Acquisition of Verbs*, (pp. 331–352). Hillsdale, NJ: Earlbaum.

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, 85(3), 956–970. doi: 10.1111/cdev.12166.

Roseberry, S., Richie, R., Hirsh-Pasek, K., Golinkoff, R. M., & Shipley, T. F. (2011). Babies catch a break: 7- to 9-month-olds track statistical probabilities in continuous dynamic events. *Psychological Science*, 22(11), 1422–1424. doi: 10.1177/0956797611422074.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child*

*Development*, 83(5), 1762–1774. doi: 10.1111/j.1467-8624.2012.01805.x.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928. doi: 10.1126/science.274.5294.1926.

Schick, B. (2010). The development of American Sign Language and manually coded English systems. In M. Marschark & P. E. Spencer (Eds.), *The Oxford Handbook of Deaf Studies, Language, and Education*, Vol. 2. New York: Oxford University Press.

Scott, J. C, & Henderson, A. M. E. (2013). Language natters: Thirteen-month-olds understand that the language a speaker uses constrains conventionality. *Developmental Psychology*, 49(11), 2102–2111. doi: 10.1037/a0031981.

Seston, R., Golinkoff, R. M., Ma, W., & Hirsh-Pasek, K. (2009). Vacuuming with my mouth? Children's ability to comprehend novel extensions of familiar verbs. *Cognitive Development*, 24(2), 113–124. doi: 10.1016/j.cogdev.2008.12.001.

Shatz, M. & Gelman, R. (1973). The development of communication skills: Modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development*, 38(5), 1–38.

Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, 15(5), 659–673. doi: 10.1111/j.1467-7687.2012.01168.x.

Shneidman, L. A., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, 40(3), 672–686. doi: 10.1017/S0305000912000141.

Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy*, 14(6), 654–666. doi: 10.1080/15250000903263973.

Smith, L. B. (2000). How to learn words: An associative crane. In: Golinkoff, R. M. & Hirsh-Pasek, K. (Eds.), *Breaking the word learning barrier*. Oxford: Oxford University Press.

Smith, L. B. (2005). Cognition as a dynamic system: Principles from embodiment. *Developmental Review*, 25(3–4), 278–298. doi: 10.1016/j.dr.2005.11.001.

Smith, L. B., Colunga, E., & Yoshida, H. (2010). Knowledge as process: Contextually cued attention and early word learning. *Cognitive Science*, 34(7), 1287–1314. doi: 10.1111/j.1551-6709.2010.01130.x.

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends in Cognitive Sciences*, 18(5), 251–258. doi: 10.1016/j.tics.2014.02.007.

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. doi: 10.1016/j.cognition.2007.06.010.

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development*, 85(5), 1821–1826. doi: 10.1111/cdev.12247.

Stevens, J. S., Gleitman, L. R., Trueswell, J. C., & Yang, C. (2017). The pursuit of word meanings. *Cognitive Science*, 41(S4), 638–676. doi: 10.1111/cogs.12416.

Stromswold, K. (1994). What a mute child tells us about language acquisition. Doctoral dissertation, Center for Cognitive Science, Rutgers University, New Brunswick, New Jersey.

Syrett, K. & Lidz, J. (2010). 30-month-olds use the distribution and meaning of adverbs to interpret novel adjectives. *Language Learning and Development*, 6(4), 258–282. doi: 10.1080/15475440903507905.

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science*, 23(2), 121–126. doi: 10.1177/0963721414522813.

Tardif, T. (1996). "Nouns Are Not Always Learned before Verbs: Evidence from Mandarin Speakers' Early Vocabularies." *Developmental Psychology*, 32(3), 492-504. doi:10.1037/0012-1649.32.3.492.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development*, 70(3), 620–635. doi: 10.1111/1467-8624.00045.

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, 10, 21. doi: 10.1186/1471-2202-10-21.

Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, 17(4), 432–444. doi: 10.1111/j.1532-7078.2011.00084.x.

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11(1–2), 61–82. doi: 10.1515/cogl.2001.012.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press.

Tomasello, M. & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development*, 10(2), 201–224. doi: 10.1016/0885-2014(95)90009-8.

Tomasello, M. & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 57(6), 1454–1463. doi: 10.2307/1130423.

Trueswell, J. C., Medina, T. N., Hafri, A. & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, 66(1), 126–156. doi: 10.1016/j.cogpsych.2012.10.001.

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's more like it: Multiple exemplars facilitate word learning. *Infant and Child Development*, 23(2), 105–122. doi: 10.1002/icd.

Vlach, H. A. & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology*, 3, 46. doi: 10.3389/fpsyg.2012.00046.

Vouloumanos, A., & J. F. Werker. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, 45(6), 1611–1617. doi: 10.1037/a0016134.

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H.-J. (2013). Are nouns learned before verbs? Infants provide insight into a longstanding debate. *Child Development Perspectives*, 7(3), 155–159. doi: 10.1111/cdep.12032.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143–2152. doi: 10.1177/0956797613488145.

Wojcik, E. H. (2013). Remembering new words: Integrating early memory development into word learning. *Frontiers in Psychology*, 4, 151. doi: 10.3389/fpsyg.2013.00151.

Woodward, A., & Hoyne, K. (1999). Infants' learning about words and sounds in relation to objects. *Child Development*, 70(1), 65–77. doi: 10.1111/1467-8624.00006.

Wu, Z., & Gros-Louis, J. (2014). Infants' prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language*, 34(1), 72–90. doi: 10.1177/01427237145 21925.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, 114(2), 245–272. doi: 10.1037/0033-295X.114.2.245.

Yoshida, H., & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science*, 16(2), 90–95. doi: 10.1111/j.0956-7976.2005.00787.x.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science,* 18(5), 414–420. doi: 10.1111/j.1467-9280.2007.01915.x.

Yu, C., & Smith, L. B. (2011). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science,* 14(2), 165–180. doi: 10.1111/j.1467-7687.2010.00958.x.

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition,* 125(2), 244–262. doi: 10.1016/j.cognition.2012.06.016.

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science,* 20(5), 619–626. doi: 10.1111/j.1467-9280.2009.02341.x.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, 83(4), 1382–1399. doi: 10.1111/j.1467-8624.2012.01783.x.

Yurovsky, D., Fricker, D. C., Yu, C., & Smith, L. B. (2014). The role of partial knowledge in statistical word learning. *Psychonomic Bulletin & Review,* 21(1), 1–22. doi: 10.3758/s13423-013-0443-y.

Yurovsky, D., Smith, L. B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science*, 16(6), 959–966. doi: 10.1111/desc.12036.

Yurovsky, D., Yu, C., & Smith, L. B. (2012). Statistical speech segmentation and word learning in parallel: Scaffolding from child-directed speech. *Frontiers in Psychology, 3,* 374. doi: 10.3389/fpsyg.2012.00374.

Zimmerman, F. J., Christakis, D. A., & Meltzoff, A. N. (2007). Associations between media viewing and language development in children under age 2 years." *Journal of Pediatrics*, 151(4), 364–368. doi: 10.1016/j.jpeds.2007.04.071

Zosh, J. M., Brinster, M., & Halberda, J. (2013). Optimal contrast: Competition between two referents improves word learning. *Applied Developmental Science*, 17(1), 20–28. doi: 10.1080/10888691.2013.748420

# 25 The Acquisition of Morphology

KAMIL UD DEEN

University of Hawaii at Manoa

## Introduction

What children do with morphology is quite simply adorable. Yes, they often mispronounce things and make pragmatic errors, but it is the morphological errors that really stand out in the minds of parents. But such errors show us more than how cute our kids are. They reveal the relationship between the child and human language—how children think about language, the kinds of hypotheses they entertain, the kinds of processes that are natural and unnatural to them. In this way, the study of morphology can be fundamentally revealing of the nature of the human child.

Morphology is the study of the smallest meaningful units of language (morphemes) and how they are organized in languages. In this chapter, we consider several of the more important categories of morphology, but the emphasis is on inflectional morphology. This is partly due to relative paucity of research on the acquisition of other kinds of morphology, and partly because of the interesting nature of inflectional morphology (see Penke, 2012). While the acquisition of lexical roots (technically a kind of morpheme) is a supremely important aspect to learning a language, this will not be addressed in this chapter (see Golinkoff & Hirsch-Pasek, this volume).

Putting aside lexical roots, then, morphemes can be split into two basic categories: inflectional morphemes and derivational morphemes. Inflection involves modification of the same basic word, resulting in slightly different meanings, while derivational morphology involves the creation of entirely new words. For example, the word *blend* is a verb root that can be inflected in numerous ways, such as *blends, blended, blending*, each carrying a different type of inflectional morphology (present tense, past tense, progressive, respectively). Moreover, derivational morphology may be added to *blend* to create a new word entirely, such as

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

*blender* (someone/thing that blends), *unblend* (to undo the blending), and *reblend* (to blend again). These latter examples are new words, derived from the original root, but able to carry their own inflectional morphology, e.g., *blenders* (plural of *blender*), *unblends* (present tense), *unblended* (past tense), *unblending* (progressive), and so on. Furthermore, the word *blend* may be combined with other lexical roots in a process called compounding, to create new words, for example, *a fruit blend*. As mentioned above, the focus of this chapter will not be on such derivational and word-formation processes, but rather on inflectional morphology.

A further important distinction in morphology relates to how closely the morphology is connected to its associates. For example, there is a sense in which the past tense *–ed* in English (e.g., *blended*) is more closely connected to the verb than, say, the definite article *the* (e.g., *the cars*). The past tense morpheme is referred to as a bound morpheme, since it is tightly bound to the verb, cannot be easily separated from it, and cannot occur without the verb root. The definite article, on the other hand, is referred to as a free morpheme, since it has more independence than its bound counterpart. It can, for example, be separated from the host morpheme by adjectives (*the fast cars*), or numerals (*the three cars*), or quantifiers (*the many cars*), etc. In this chapter, we focus on bound morphology.

There are various kinds of bound morphology, the most common of which are prefixes (occurring before the root or stem, e.g., *untie, retie*) and suffixes (occurring after the root or stem, e.g., *kicked, kicking, kicker*). Other kinds of bound morphology do occur, such as infixes and circumfixes. The bulk of research on inflection has been on prefixes and suffixes, although some work has been done on infixes.

These various kinds of morphology (inflectional versus derivational; bound versus free; prefix versus suffix versus infix versus circumfix[1]), as well as differences in the manner in which languages integrate morphemes into the rest of the sentence—isolating versus agglutinating versus (poly)synthetic languages—all make the acquisition of morphology a tremendously interesting research area. Most of what we know about the acquisition of inflection comes from the acquisition of Germanic and Romance languages. The research tradition on the acquisition of English dates back to the earliest days of the field (as reviewed below), but in recent years a large amount of research on other languages has been published, and we will include discussion of these languages as they become relevant. In fact, a (non-exhaustive) survey of the literature reveals significant gaps in our knowledge of how children acquire the morphology of the world's languages, as discussed in the conclusion. But we begin with the foundational study on the acquisition of morphology—Roger Brown's (1973) monograph on the acquisition of English morphology.

## The foundation: Roger Brown

Brown was interested in the question of whether children acquire the morphemes of their language in any kind of predictable order. This question was a very important one, especially in the context of the contemporary debate regarding the uniformity of the acquisition of language. Chomsky had argued that children acquire language in a relatively uniform manner, while his opponents argued that child language was far less predictable. It was far less the metronomic process of acquisition that Chomsky described, and more of an organic, piece-meal process of building larger pieces from smaller bits of language. The process was thought to be highly dependent on the input to children (which varies greatly from one moment to the next, from one interlocutor to the next, and from one home to the next), and as such, opponents to Chomsky argued that child language is a messy, error-filled process that is far less uniform than assumed.

Brown's question was therefore of crucial importance. Do English-speaking children acquire the morphemes in their language in the same order, or is there significant variation from child to child? To investigate this, he and his students collected longitudinal data from three children (given the pseudonyms Adam, Eve, and Sarah) over several years. His team then transcribed the data from each child (and their care-givers), and analyzed the acquisition of morphology. But a crucial problem Brown ran into was how to determine when a morpheme is acquired. It is not as obvious as it initially seems. Does the first production of a morpheme indicate acquisition? Not always, since children often imitate language without fully comprehending what they are imitating. Also, children often learn formulaic phrases (e.g., *all gone*, *who's that?*) and the morphemes therein are not actually analyzed by the child as independent units of meaning (at least not initially). Furthermore, children often use morphology sporadically (see next section), raising the question of whether such sporadic use of morphology indicates that the morpheme is acquired. Brown devised a method to address this question, which we discuss next.

## Brown's method for establishing a morpheme has been acquired

### *Obligatory contexts*

Brown's method begins with the recognition that the simple fact of a child using a particular morpheme is not sufficient to establish that the child has a deep and adult-like knowledge of that morpheme. For example, if a child says "the ball is in the kitchen," the use of the definite article "the" is only appropriate if there is a previously mentioned ball. If the child uttered this sentence without a previous mention of the ball, the use of the definite article is inappropriate and should not count as evidence that the child has acquired the definite article. Likewise, if there was a previously mentioned ball in the context and the child failed to use the definite article, then this would also indicate the child is missing some knowledge of the definite article. So Brown argued that when establishing whether a morpheme has been acquired or not, one must only consider obligatory contexts for that morpheme. That is, if it is unclear what the context is, or if the context is compatible with more than one morpheme, then one cannot consider what the child says as evidence for or against knowledge of a particular morpheme. So we are restricted to obligatory contexts for the analysis of the acquisition of morphology.

## *90% criterion*

What Brown found was that in obligatory contexts, at early ages, children omitted morphemes in very high proportions, sometimes producing the appropriate morphology in 0% of obligatory contexts. But as they mature, that rate gradually climbs, until at some point, children reach an adult-like 100% (or close to it) supply of morphology in obligatory contexts. So at what point does one give credit to the child for having acquired the morpheme? Must we wait until the child has achieved 100% supply in obligatory contexts? This seems overly stringent, since surely a child who produces a morpheme in 95% of obligatory contexts has significant knowledge of a morpheme. So Brown set the criterion at 90% of obligatory contexts. That is, a child has to produce a particular morpheme correctly in 90% or more of obligatory contexts before the child is considered to have acquired that morpheme. This number was an arbitrary number, and has since been criticized as being overly stringent. For example, if a child produces a morpheme in 80% of obligatory contexts (or even 60%), does the child not have some knowledge of what the properties of that morpheme are? As such, modern researchers often set their own (lower) criterion (often between 70% and 90%).

## *Consistency*

Brown noticed that there is a large amount of variation in supply of morphemes in obligatory context from transcript to transcript. So in one transcript, the child might produce third person singular *–s* in 70% of obligatory contexts, while in the very next transcript, the child produces *–s* in only 35% of obligatory contexts. And in the very next transcript, that figure may swing all the way up to 90%. This is in part due to changes in the child's mood/temperament, changes in who the interlocutors are, and also changes in the size of the sample in each transcript.

Consider Table 25.1, a hypothetical data set showing supply of third person *–s* in the speech of one child. In transcript 1, the rate of supply of *–s* in obligatory contexts is 55%, and in transcript 2 this rate rises to 100%. However, in the next two transcripts, that rate is well below the 90% threshold. It is clearly incorrect (according to Brown's 90% criterion) to conclude that the child has acquired third person singular morphology by transcript 2.

Because of such cases, Brown included as part of the criterion for acquisition the requirement that the 90% threshold be met across three consecutive transcripts. Thus if that threshold is crossed for one transcript, perhaps due to a sampling anomaly (transcript 2, Table 25.1), and then drops down below 90% (transcript 3, Table 25.1), then the child is not considered to have acquired the morpheme. But if the rate remains above 90% for three consecutive transcripts (as is the case in transcripts 5, 6, and 7), then the child is considered to have acquired the morpheme at the first transcript in which the threshold was met—transcript 5.

**Table 25.1** Hypothetical data set of a child's production of third person singular –*s* in obligatory contexts.

| *Transcript* | *Age of Child* | *Number of Utterances* | *Morphemes/ Obligatory Contexts* | *% Morpheme in Obligatory Context* |
|---|---|---|---|---|
| 1 | 1;10.04 | 1357 | 68/123 | 55 |
| 2 | 1;10.28 | 45 | 4/4 | 100 |
| 3 | 1;11.08 | 1022 | 70/97 | 72 |
| 4 | 1;11.27 | 973 | 85/105 | 81 |
| 5 | 2;0.03 | 1102 | 120/130 | 92 |
| 6 | 2;0.25 | 985 | 102/110 | 93 |
| 7 | 2;1.05 | 1001 | 132/139 | 95 |

### *Mean Length of Utterance*

Furthermore, Brown was intent on comparing the language of three different children to establish whether there was uniformity in the acquisition of the 14 morphemes. He and others (notably Brown, Cazden & Bellugi-Klima, 1968; Brown & Bellugi, 1964) had discovered that age is a very bad indicator of linguistic maturity, and so he used the measure now known as Mean Length of Utterance (MLU). This measure takes all the morphemes in a random selection of 100 utterances and divides that number by 100, thus resulting in the mean length of utterance for that sample. This was shown to be a better measure of linguistic growth, and has now been widely adopted in the field. There are problems with MLU (e.g., it does not allow for accurate comparison across languages), but it is widely acknowledged to be a better indicator of linguistic maturity than age (see Valian, 1991 for discussion).

## Results

With all these methodological innovations, what did Brown find? He found that across the three children that he studied, the fourteen morphemes that he targeted were indeed acquired in the same order—in fact, the order given in (1).

(1) Order of Morpheme Acquisition (Brown, 1973)
1. Progressive *-ing*
2. Proposition *in*
3. Preposition *on*
4. Regular plural *-s*
5. Irregular past
6. Possessive *-s*
7. Uncontractible copula
8. Articles
9. Regular past *-ed*

10. Regular third person -*s*
11. Irregular third person
12. Uncontractible auxiliary
13. Contractible copula
14. Contractible auxiliary

Several factors might be at play in determining this language-specific order of acquisition: frequency, semantic weight, saliency, etc., all of which have been shown to play a part (see, for example, de Villiers & de Villiers, 1973). But interestingly, no single factor is sufficient to explain this order. For example, the order of morphemes in (1) is not in descending order of frequency: the articles *the* and *a* are the most frequent items in the input to a child, but these are acquired later than possessive -*s*. See Peters (1997) for a discussion of the various factors that influence the acquisition of morphemes, and how they interact.

Furthermore, it is not the case that the equivalent of these morphemes are acquired in the same order across languages. So it is not necessarily the case that possessive morphology is acquired after plural morphology, nor that plural is acquired before past tense. This means that semantic content or function is not sufficient to predict the order of acquisition of morphemes. It seems that all these factors combine, on a language-by-language basis, to produce the order of acquisition that any language exhibits.

## Acquisition of a rule, or memorized chunk: Jean Berko Gleason

One question that arises from Brown's method is whether what is being documented is children's actual acquisition of rules of morphology, or whether children are simply acquiring memorized forms and very astutely recruiting them in the appropriate contexts. For example, does the child who correctly produces the word *cat* and its plural counterpart *cats* actually know that the rule of pluralization in English involves the suffixation of –*s* ([s], [z] or [əz], as the case may be), or have they simply memorized that when talking about one cat, "cat" is appropriate, and when talking about many cats, "cats" is appropriate? This latter possibility might result in correct production of plural morphology in obligatory context, but the nature of the knowledge being exhibited would be qualitatively different from adults'. It turns out this question had been answered more than a decade earlier in the seminal work of Jean Berko Gleason (1958). Berko Gleason is the innovator of the so-called *Wug Test*, a device whose basic principles are used by researchers in almost every modern experiment today.

The central idea behind the Wug test is to present children with a novel word that they have never heard before, and to ask what the inflected form of the word is. Because children have never heard these words before, they could not possibly produce the correct form based upon memorization of previously heard forms. In the experiment, children saw a picture of a stylized creature—something that was vaguely familiar but certainly nothing identifiable—and which was labeled as follows: "This is a wug." Children then saw a second picture with two of the novel creatures on it, and were prompted as follows: "Now there is another one. There are two of them. There are two______." Children were expected to complete the sentence with the correct noun form. If children were aware of the pluralization rule in English, this task should be easy: the correct answer is [wʌgz].

The findings showed that children generally did apply the rules of inflection from very early ages (as young as age four years). More recently, research has shown that knowledge of the rule of pluralization is in place (to some extent) at ages younger than even three years (Zapf & Smith, 2007; Lukyanenko & Fisher, 2014). Berko Gleason (and others) have also tested children on their knowledge of a variety of other inflectional forms (progressive aspect, past tense, third person singular agreement, possessive, the agentive derivational morpheme *–er*, and compounding). Thus while there is evidence that children employ memorized chunks in speech (e.g., Peters, 1983), there is strong evidence that, from very early on, children do indeed acquire and employ rules of grammar in the acquisition of morphology.

## Generalizations on the acquisition of inflection

We turn now to some generalizations about the acquisition of (inflectional) morphology that have been observed in a variety of different languages. All else being equal, we expect these generalizations to hold in the investigation of the acquisition of a previously un-studied language. While there are known exceptions to each of these generalizations, such exceptions can be explained through some language-specific factor. As such, when investigating the acquisition of any new language, divergence from one of these generalizations might be taken as evidence that there is something worthy of further investigation.

### *Rapidity*

Children learn inflection very quickly. Generally speaking, children acquire their inflectional system before age four years, with some errors related to the rare and non-systematic aspects of inflection remaining for another year or so. There is some debate in the literature on the underlying nature of the learning that accounts for this rapid acquisition. Some researchers (e.g., Wexler, 2004) claim that the acquisition of inflection is very quick, very easy, and maximally rule-governed because children are endowed with a disposition for rule learning, knowledge of the kinds of semantic categories typically encoded by inflection, and the kinds of morphological processes that languages typically exhibit. Others (e.g., Rubino & Pine, 1998) claim that the acquisition of inflection is actually an error-riddled process that varies from child to child, and that while inflection is acquired by age four years, the preceding few years is nothing like the clean, predictable process that rule-based systems might predict. Whatever the case, the fact remains that by age four years, children have acquired the bulk of morphology in their language, and that all that lags behind are the irregular and relatively rare forms.

## *Accuracy*

As a whole, children are very accurate in their use of inflectional morphology. This is true from very early stages, although the exact reason for such accuracy is the subject of great debate (see, for example, Lieven, Salomo & Tomasello, 2009; Ambridge & Lieven, 2011, and references therein). Nonetheless, the fact remains that children are very good learners of morphology (Wexler, 2004 refers to children, perhaps slightly hyperbolically, as little inflection learning machines), and as we shall see below, once different errors are categorized properly, the acquisition of morphology can be seen as a systematic, orderly process far from the messy process described by some.

## *Prefixation versus suffixation*

The two most common kinds of bound inflectional morphology found in the languages of the world are prefixes and suffixes. Children appear to find suffixation easier than prefixation, an observation first noted in Slobin's (1973) Operating Principle in which he says children "pay attention to the ends of words." This has been shown in many different kinds of languages, even those that have rich sets of both prefixes and suffixes. For example, Deen (2005) shows that children acquiring Swahili omit the obligatory prefixes in as much as 80% of obligatory contexts at certain stages of acquisition, while the obligatory suffixes are omitted in less than 1% of cases during the same time period. In fact, no stage was ever detected in which obligatory suffixes are omitted at significant rates, suggesting that inflectional suffixation is acquired (in Swahili at least) before the onset of multi-word utterances. See also Demuth (1992, Sesotho), Kunene (1979, ), Suzman (1996, Zulu), Courtney (2003, Quechua), and Imedadze and Tuite (1992, Georgian).

As for other kinds of morphology (infixes and circumfixes), there exists relatively little research. Segalowitz and Galang (1978) tested three age groups of children on their acquisition of focus morphology (also sometimes referred to as voice) in Tagalog: three year olds (3;1-3;11, mean 3;6), five year olds (5;1-5;9, mean 5;6) and seven year olds (7;1-7;5, mean 7;4). Verb morphology in Tagalog indicates the thematic role of the argument bearing the focus marker *ang*.When the agent of a transitive (2) clause is focus-marked with *ang*, the verb takes an agent-focus infix<um>; in contrast, when the theme of a transitive clause is focus-marked with *ang* (3), the verb takes a theme-focus infix<in>.

| | | | | | |
|---|---|---|---|---|---|
| (2) | K<um>a~kain | ng | mansanas | ang | lalake. |
| | <AF>IPFV~eat | NFOC | apple | FOC | boy |
| | "The boy is eating an/the apple." | | | | |

| | | | | | |
|---|---|---|---|---|---|
| (3) | K<in>a~kain | ng | lalake | ang | mansanas. |
| | <TF>IPFV~eat | NFOC | boy | FOC | apple |
| | "A/the boy is eating the apple." | | | | |

Segalowitz and Galang find that in an elicited production task, even their youngest age group exhibited knowledge of this focus morphology, correctly producing both focus types in more than 70% of responses. Tanaka, O'Grady, Deen, Kim, Hattori, Bondoc, and Soriano (2014) investigated the acquisition of relative clauses in Tagalog, but in so doing, also investigated the acquisition of Tagalog voice morphology. They too found that children were able to produce the voice morphology (along with the respectively correct word order) as young as age 3;10. This suggests that infixes are not as problematic as prefixes, but this is data from just one language.

## *Rich versus impoverished morphology*

Intuitively, one might think that a language with less morphology would be easier to acquire than one with a rich set of morphemes. English, for example, has a relatively meager inventory of bound inflectional morphology, while Italian has a much richer inventory. As adults, we find such complex inflectional systems daunting and challenging, but study after study has shown that children acquiring morphologically rich languages acquire adult-like proficiency at much earlier ages than children acquiring more impoverished languages like English. Some English-acquiring children continue to struggle with past tense, third-person singular, plural *-s*, and other morphemes as late as age 4 years, while Italian speaking children have been shown to master Italian morphology significantly earlier (Hyams, 1986; Valian, 1991; Guasti, 1993/4). A quick glance at the literature reveals that this is not limited to Italian—children acquiring languages from a variety of language families with rich morphology exhibit control over their morphology at very young ages: Turkish (Aksu-Koc & Slobin, 1985), Greek (Tsimpli, 1992; Stephany, 1995), Arabic (Aljenaie, 2010), Hungarian (Gábor & Lukács, 2012), Sesotho (Demuth, 1992), Zulu (Suzman, 1996), Hebrew (Schaeffer & Ben-Shalom, 2004), Malagasy (Hyams *et al.*, 2006), Swahili (Deen, 2005), etc.

One reason for this may be the role such rich morphology plays in the respective languages. As discussed by Dressler (2007), morphology in so-called rich morphological languages often does much of the work done by syntax in morphologically impoverished languages like English. Thus morphology is a more important cue in languages like Italian, Swahili, and Russian than in languages like English and French, which may very well be what leads to the difference in speed of acquisition amongst the two kinds of languages.

A second reason (thanks to a reviewer for pointing this out) is that in some languages, the root of the verb simply never occurs as a stand-alone form in the language. Thus in Italian, the root of the verb *mangiare* is *mang-*, but this is not a pronounceable form in the language. Thus were morphology to be omitted by the child, the resulting form would be completely ill-formed. In English, on the other hand, the root of the verb *to eat* is *eat*, which actually is a perfectly acceptable stand-alone form (e.g., as an imperative: *eat!*). Thus children, by avoiding morphologically ill-formed roots in languages like Italian, are exhibiting yet another form of knowledge of their language that they have acquired early in development.

## *Kinds of errors*

Saying that children are by-and-large accurate with their morphology (as discussed earlier) does not mean that their speech is error-free. In fact, this chapter began with the claim that inflectional errors are amongst the cutest errors children produce, so children obviously do make errors. But interestingly, the kinds of errors children make can be seen as evidence for knowledge of the inflectional system of their language, and not for ignorance of those properties.

There are two kinds of errors that children make with inflection: errors of omission and errors of commission. The former are errors in which children omit the required morpheme. For example, if a child says "Mommy love cookies," this is an error of omission in that the child omitted a verbal inflection (either third, singular *–s* or past tense *–ed*, depending on the context). Such errors are very common, and have been documented in a variety of languages (some examples are provided in (4); in (4b), IND = indicative mood).

| | | | |
|---|---|---|---|
| (4) | a. | He bite me | Agreement Omission (Brown, 1973) |
| | b. | Ø – qet – il – e | Agreement Omission, Sesotho (Demuth, 1992) |
| | target: | ke – qet – il – e<br>1sg–finish–PAST–IND<br>'I finished' | |
| | c. | alafu a – Ø – rud – i | Tense Omission, Swahili (Deen, 2005) |
| | target: | alafu a – li – rud – i<br>then 3sg–PAST–return–IND<br>'Then he returned.' | |
| | d. | I in the kitchen | Copula Omission (Becker, 2000) |
| | e. | baby talking | Auxiliary Omission (Radford, 1990) |
| | f. | Paula play ball | Determiner Omission (Radford, 1990) |
| | target: | Paula plays with the ball | |

Errors of omission seem to be a cross-linguistic phenomenon, being documented in a host of languages far too numerous to list here. Furthermore, errors of omission are not only widely attested across languages, but within any single language, errors of omission occur at very high rates. Sano and Hyams (1994) reported that in the speech of the children Eve (age 1;6-1;10), Adam (age 2;3-3;0), and Nina (2;4-2;5, all data available on CHILDES, MacWhinney, 2000), the rate of omission of *-s* in third person singular contexts was 78%, 81% and 75%, respectively. Deen (2005) reports similar percentages for errors of omission in Swahili, where the verb is minimally inflected for subject agreement (SA), tense (T) and mood, as shown in (5). In the speech of two children (Haw, aged 2;2-2;6, and Mus aged 2;0-2;3) subject agreement was omitted 72% and 54% of the time, respectively, and tense was omitted 70% and 40% of the time respectively. Thus the omission of inflection is both widely attested cross-linguistically, and very frequent within each language.

(5) Swahili minimal verbal complex: SA –T – V – Mood

| | | | |
|---|---|---|---|
| Example: | ni – | li – | anguk– a |
| | 1sg –PAST– | fall – | ind |
| | 'I fell.' | | |

Errors of commission (also known as errors of substitution), on the other hand, occur when a child produces morphology that is incorrect and contextually inappropriate. For example, if a child says "I loves cookies," this constitutes an error of commission (an agreement error). In this example, assuming the subject pronoun indicates that the context of the sentence involved a first person referent, the child's use of the third person singular suffix on the verb is incorrect. Such errors of commission have been found to be remarkably rare in child language.

It is important to distinguish two kinds of errors here, one of which is properly referred to as an error of commission (as in the example above, "I loves cookies"), the other being an overregularization error (as in a child saying "I finded the ball" instead of "I found the ball"). The former indicate that the child genuinely has not acquired the inflectional properties of the target language. The latter, however, is not an indication of ignorance of the correct usage of a particular morpheme, but more indicative of the particular idiosyncratic, unexpected properties of a particular language. Let us consider both types of errors in turn below.

Errors of commission like "I loves cookies" are extremely rare in child speech, as has been shown by numerous researchers over the years. Harris & Wexler investigated the speech of ten English speaking children (age range 1;6–4;1), and identified 1,724 verbs that occurred in the first person singular context, of which only 3 occurred with the incorrect third person singular *-s* suffix—a remarkably low error rate of 0.17%. Contrast this to the more than 70% error rate of error of omission described above.

Similarly, in the speech of two Swahili-acquiring children (aged 2;10-3;0 and 1;8-2;1), Deen (2005) found low rates of errors of commission. Of the 224 verbal utterances produced by the older child, only three agreement errors were found (an error rate of 1.3%), and of the 197 verbal utterances produced by the younger child, only one error was found (an error rate of 0.5%). Table 25.2 (adapted from Sano & Hyams, 1994) shows the rate of errors in agreement in a number of children acquiring various languages.

Moreover, in languages with agglutinating affixes, it is often the case that children must produce multiple affixes in any single verbal utterance. For example, in Swahili, the minimal verbal complex consists of three inflectional affixes (subject agreement, tense and mood) plus the verb root. A further three inflectional affixes may occur (e.g., object agreement in some transitive clauses), and a further half dozen or so derivational suffixes. In Deen's data, not a single error of affix ordering was observed—children omitted morphemes, sometimes to large extents, but they never placed them in the wrong order. The same has been found in numerous other languages with similar systems (Clancy, 1986 for Japanese; Kim, 1997 for

**Table 25.2** Rate of agreement errors in a range of languages

| *Child* | *Language* | *Age* | *Utterances* | *Percentage error* |
|---|---|---|---|---|
| Simone (Clahsen & Penke, 1992) | German | 1;7-2;8 | 1732 | 1.0 |
| Martina (Guasti, 1993/94) | Italian | 1;8-2;7 | 478 | 1.6 |
| Diana (Guasti, 1993/94) | Italian | 1;10-2;6 | 610 | 1.5 |
| Guglielmo (Guasti, 1993/94) | Italian | 2;2-2;7 | 201 | 3.3 |
| Claudia (Pizzuto & Caselli, 1992) | Italian | 1;4-2;4 | 1,410 | 3.0 |
| Francesco (Pizzuto & Caselli, 1992) | Italian | 1;5-2;10 | 1,264 | 2.0 |
| Marco (Pizzuto & Caselli, 1992) | Italian | 1;5-3;0 | 415 | 4.0 |
| Gisela (Torrens, 1995) | Catalan | 1;10-2;6 | 81 | 1.2 |
| Guillem (Torrens, 1995) | Catalan | 1;9-2;6 | 129 | 2.3 |

Korean; Aksu-Koc & Slobin, 1985 for Turkish; Courtney & Saville-Troike, 2002 for both Navajo and Quechua, amongst many others).

So errors of commission are rare in child language, especially when compared to the rate of errors of omission. However, when one inspects the transcripts of children in this age range (roughly age 1;10 to 3;6), one finds numerous morphological errors that look like errors of commission, such as the use of *mans* to mean *men*, or *mouses* to mean *mice*, or *foots* to mean *feet*. Such errors are referred to as errors of overregularization. There is a sense in which such errors cannot be characterized as errors of ignorance (as is the case with the errors of commission).

## *Overregularization and U-shaped development*

An interesting fact about such morphological errors is that they don't occur when children are very young, but (perhaps counterintuitively) such errors become frequent at a later point in development (Cazden, 1968; Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992). So a child at an early age might initially use the correct morphological forms, like *men, went, feet, ran*, and so on, only to later stop using these correct forms and start producing overregularized errors, like *mans, goed, foots*, and *runned*.[2] At a subsequent stage, children eliminate such errors from their speech and go back to the correct forms. This developmental

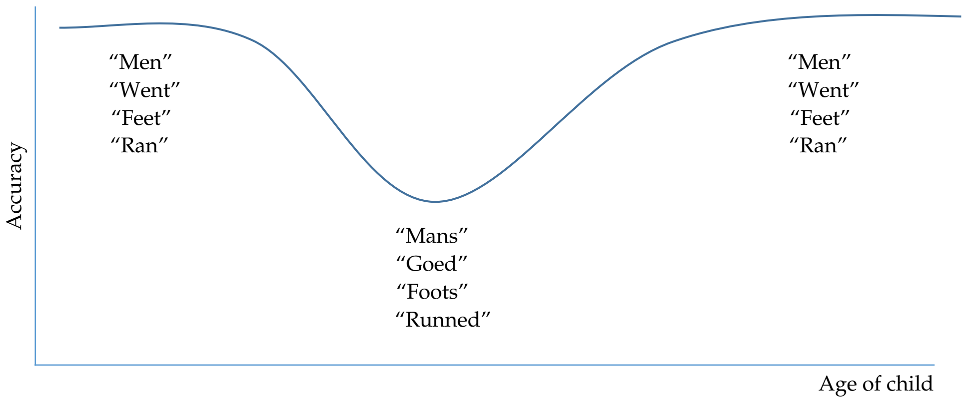



**Figure 25.1** Hypothetical U-shaped development.

path is referred to as U-shaped development since high accuracy is followed by a period of low accuracy, which leads to a return to high accuracy (see the hypothetical Figure 25.1).

Errors of overregularization are seen as evidence that children have acquired the underlying system in the language, but have not acquired the exceptions yet. A child who says *foots* knows that *–s* indicates plurality in English, but the child has not learned that *foot* is an exception to the plural rule. Strictly speaking, these forms are not accurate, and so the rate of accuracy drops. However, when accuracy begins to fall because of overregularization, this should be taken as evidence of a child having acquired the rules governing a particular inflectional paradigm, and not for an absence of knowledge. Thus errors of overregularization, while perhaps similar to errors of commission in that both involve the use of non-target morphology, are fundamentally errors that involve the over-application of a rule rather than errors of ignorance.

There is considerable debate on how children retreat from such overregularization errors. Space precludes a full discussion of this, but see Marcus *et al.*, 1992; Maratsos, 2000; Ambridge *et al.*, 2013 for discussion.

## *Root infinitives*

A Root Infinitive (RI) is a verb that is marked with overt non-finite morphology and that occurs in a root (main) clause, shown in (6).

(6) a. Thorsten das hab-en — German (Andreas, 2;1, Wagner, 1985)
Thorsten that have-INF
"Thorsten has that."

b. Ferm-er yeux — French (Daniel, 1;11, Lightbrown, 1977)
close-INF eyes
"(I have) closed (my) eyes"

The verbs here (*haben* and *fermer* in German and French respectively) are not just missing inflection, but are overtly marked as infinitives. In adult language, this is generally ungrammatical (although non-finite verbs do occur in certain root clauses). For example, the appropriate form of the verb in (6a) in adult German would be the finite *hat*, not *haben*. This could be construed as an error of commission since infinitival morphology is supplied by the child in a context that does not call for infinitival morphology. This is not how RIs are viewed, however.

One of the most striking facts about RIs is that the occurrence of the morphological infinitive is not a morphological error. Rather, children position the infinitival verb in a position reserved for non-finite verbs, thereby exhibiting knowledge that the form they are using is indeed a non-finite form. Take German as an example. The underlying word order of German is Subject-Object-Verb. In main clauses, however, finite (inflected) verbs move from this sentence-final position into the second position (7a). German has another rule which requires something else (anything else) to move into the position immediately preceding the verb, resulting in the phenomenon widely known as V(erb)-2: where the German finite verb occurs in second position of main clauses.

| | | | | | | |
|---|---|---|---|---|---|---|
| (7) | a. Ich | sehe | viele | Leute | | Finite German Verb |
| | I | see.1sg | many | people | | Verb in second position (V2) |
| | "I see many people" | | | | | |
| | b. Ich | mochte [viele Leute seh-en] | | | | Non-finite German Verb |
| | I want | [many people see-INF] | | | | Verb in final position |
| | "I want to see many people." | | | | | |

Infinitives, on the other hand, do not move anywhere because there is no need for the verb to obtain finite features. Infinitives occur in their original position at the end of the sentence, as seen in (7b). Thus the descriptive property relevant for our purposes is that in adult German, finite verbs occur in second position (7a) while infinitives occur at the end of the clause (7b). This property is mirrored in child root infinitives in that when RIs occur, they occur in the correct sentence-final position (as seen in 6a). In an analysis of main clause verbs in child German, Poeppel and Wexler (1993) found that, with few exceptions, inflected verbs occurred in the (correct) second position (197/208) while uninflected verbs (RIs) occurred in the (correct) sentence-final position (37/45). Similar results have been reported for other languages (e.g., Wijnen, 1997 for , Pierce, 1989 for French; see Hoekstra & Hyams 1998 for an overview).

There are other empirical reasons why researchers do not consider RIs to be errors of commission, many of which are discussed in Hyams and Orfitelli (this volume). But suffice to say that RIs are not considered errors of commission, but are generally seen as errors that reflect some other syntactic or semantic property of child grammar.

## The role of input

There has been a notable amount of research on the role input frequency plays in the arena of overregularization. There have been numerous analyses of the input to children to ascertain the number of correct irregular forms (e.g., went, feet, mice) a child must hear before the overregularized form (e.g., *goed, foots, mouses*) is eliminated from the child's language. The answer to this question depends largely on which theoretical approach one adopts. The two most prominent approaches in the literature may be dubbed the Blocking Approach (Marcus *et al*., 1992) and the Competition Approach. The former contends that once the child hears the irregular form in the input, the overregularized form is blocked, and is therefore immediately removed from the child's grammar. Input, therefore, will quickly purge the system of the overregularization.

Competition models (e.g., Rumelhart & McLelland, 1985) predict that children initially postulate the overregularized form, and when they hear the irregular form for the first time, they entertain both options for a while. As children hear the irregular time and time again, and fail to hear the overregularized form, the irregular gains strength and the overregularization is slowly purged from the system. As Maratsos (2000) points out, this is precisely why the more common irregulars (e.g., *went* and *feet*) are acquired earlier than the less frequent irregulars. This approach assumes that a far larger amount of input is required in order to cleanse the system of overregularization errors.

More generally though, how much input is required for an inflectional morpheme to be acquired? Are hundreds of tokens required, or just a handful? One answer to this question was provided by Kim *et al*. (2014), who investigated how much input (and what kind of input) is required in order for a child to acquire a very rare morpheme. Korean has a plural marker, *-tul*, that occurs to the right of the root to which it attaches, but to the left of the case marker: [root-*tul*-case]. The first noun in example (8) exemplifies this plural marker, referred to as the Intrinsic Plural Marker (IMP). This plural marker is very frequent in Korean. However, Korean also has a far rarer "plural" form, referred to as the Extrinsic Plural Marker (EPM), also pronounced *–tul*, exemplified on the second nominal in (8). Note that the EPM is homophonous with the IPM, but it occurs to the right of the case marker: [root-case-*tul*]. The EPM is not in fact a marker of plurality *per se*, but a marker of distributivity—with the EPM, the sentence must have a distributive meaning, such as the one provided in the gloss. In the absence of the EPM, the sentence could have a collective meaning, that is, "the students all (as a group) gave the children money." The questions Kim *et al*. ask is when do Korean children acquire the EPM, and what kind of input is required.

| | | | | |
|---|---|---|---|---|
| (8) | Haksayng-tul-i | ai-eykey-tul | ton-ul | cwu-ess-ta. |
| | student-IPM-NOM | child-DAT-EPM | money-ACC | give-PST-DECL |
| | "Students each gave the children money." | | | |

Kim *et al*. find that in a corpus of speech, 93 instances of EPM occur out of 800,000 words, none of which are in the pattern exemplified in (8) above (i.e., EPM in a ditransitive clause on the dative-marked nominal). They then test 20 Korean children aged 5;3 – 6;9 (mean 6;1) using a Truth Vale Judgment Task (Crain & Thornton, 1998) and find that all the children failed to show knowledge of the distributive requirement of EPM. This is not surprising, given the paucity of tokens in the input. But this raises an important question: if the EPM is so rare that children aged 6 years have not yet acquired it, how could they ever acquire this morpheme? How many tokens do they need to acquire EPM?

Kim *et al*. exposed children to two scenarios involving interaction between a mother and child in which the distributive meaning was exemplified. The first scenario provided the child with positive evidence of the distributive meaning of EPM (the child watches a movie clip of a model child who listens to her mother correctly use EPM to describe a situation, and the model child agrees that her mother had described the scene correctly) while the other provided negative (corrective) feedback on the meaning of EPM (the model child listens to her mother incorrectly use EPM, and the model child incorrectly agrees with the mother, but the mother then informs the child that the description of the scene was incorrect). The children were counterbalanced with respect to which scenario they saw first, and all children were tested between scenarios to isolate the relative effect of each scenario.

It was found that the negative feedback scenario had no effect on children, while the positive evidence scenario alone resulted in 14 of the 20 children acquiring the distributive meaning of EPM. A subsequent testing two weeks later showed that all 14 of the children retained knowledge that EPM carries a distributive function.

This shows that the amount of exposure required to acquire some properties of morphosyntax need not be large—as little as a single, meaningful exposure is enough for children to acquire some aspects of morphosyntax. Kim *et al*. refer to this as syntactic fast mapping, akin to lexical fast mapping, since just one exposure is sufficient to produce long-lasting knowledge of the properties of EPM in Korean.

## Other factors that affect the acquisition of morphology

There are two broad classes of theories that address why errors in morphology might occur. The first is that the child's underlying grammar is fundamentally different from that of the adult, and that the error is a faithful reflection of this non-adult-like grammar. For example, a child that omits tense morphology may have a grammar in which the T(ense) projection is either not projected, or underspecified in some manner. Thus the absence of tense morphology is an accurate reflection of the child's unadult-like grammar. This line of inquiry has been pursued extensively for the famous Root Infinitive phenomenon discussed earlier—where children produce nonfinite verb forms in root (finite) contexts. Such forms, which at first blush appear to be errors of commission, might more profitably be analyzed as principled errors—that is, errors that arise not out of ignorance or a breakdown in any production system, but rather as a reflection of a deeper linguistic principle that is exhibited in child language, and perhaps some adult languages. See Hyams and Orfitelli (this volume) for more on the RI phenomenon.

The second class of approaches holds that errors of commission arise from some other domain of language. For example, Song, Sundara and Demuth (2009), in an analysis of the speech of six English speaking children's (age 1;3-3;6, mean = 2;2) spontaneous speech, find that children produce third person singular *–s* in obligatory contexts more often when the *–s* forms a simple coda (e.g., *sees*) as opposed to a complex coda (e.g., *needs*). Furthermore, Mealings and Demuth (2014) find that third person *–s* is produced more accurately in medial position within the sentence. Moreover, accuracy was higher in shorter sentences (three words) than in longer sentences (five words). These findings (and others) suggest that there are multiple sources for morphological error, including syntactic, phonological (e.g., coda complexity), and processing (e.g., sentence length). Some have claimed that nominal morphology is acquired before verbal morphology (e.g., Dressler *et al*., 2007), but the evidence is not clear on this question. Much depends on what the properties of the particular languages are. As discussed above, there are numerous factors involved in the acquisition of morphology in general, and these may apply differently in the verbal versus nominal domain in any particular language. Thus any generalization regarding a difference between verbal and nominal morphology is likely to be difficult to substantiate.

## Conclusion and future research

In this chapter, we reviewed several clear generalizations in the acquisition of morphology. Children acquire morphology quickly and accurately; they find prefixes more challenging than suffixes; they omit morphology very often, and occasionally overregularize, but rarely commit errors of commission. These generalizations have emerged from decades of research on well-known languages like English, German, Italian, Spanish, and so on. Despite this, cross-linguistic research has grown in recent years, and we continue to learn more about the acquisition of morphology.

There remain many unanswered questions, however, which perhaps will be addressed by further cross-linguistic research. What is the relative role of frequency, phonology, saliency and function in the acquisition of morphology? One thing that would help answer this question would be a wealth of data from a variety of languages that show the relative effect of each of these factors in each language. Additionally, remarkably little is known about the acquisition of infixes and circumfixes. While these kinds of affixes are indeed rare in the languages of the world, one wonders whether they pose a difficulty for children to acquire, more so than even prefixes. The little research that there is (on Tagalog) suggests that children do not have difficulty with infixes, but this is just one language. Moreover, the vast majority of research on child morphology has focused on languages that are isolating, weakly inflected, or agglutinating. There has been relatively little research on the acquisition of synthetic or polysynthetic languages (although see Kelly, Wigglesworth, Nordlinger, & Blythe, 2014, for an excellent overview). This represents, in my view, an important avenue of future research.

Finally, below, as Table 25.3, is a non-exhaustive list of languages for which we have some information on how children acquire morphology (references are selected—other references may exist for each language). What is evident is that while our breadth of data is impressive, this is dwarfed by the vast number of languages for which we have little or no information. For example, the Austronesian language family consists of more than 1,200 individual languages, and yet it appears that we have acquisition data from less than half a dozen Austronesian languages. With roughly 6,000 languages on the planet, our sample of languages for the acquisition of morphology is less than 1% of that. And so the generalizations

**Table 25.3** Review of research on the acquisition of morphology across languages

| *Classification* | | *Language* | *Sample Reference* |
|---|---|---|---|
| Indo-European | Germanic | English | Berko, 1958; Brown, 1973 |
| | | German | Poeppel & Wexler, 1991; Clahsen, 1986 |
| | | Dutch | Wijnen & Verrips, 1998; Wijnen, Kempen & Gillis, 2001; Blom, 2007 |
| | | Icelandic | Ragnarsdóttir, Simonsen & Plunkett, 1999 |
| | | Swedish | Josefsson, Platzack & Håkansson, 2004 |
| | | Norwegian | Westergaard, 2008; Henden, 2013 |
| | | Danish | Hamann & Plunkett, 1998 |
| | Romance | Italian | Guasti, 1993/94 Valian, 1991; Hyams, 1986 |
| | | Spanish | Grinstead, 2000; 2009 |
| | | French | Pierce, 1992; Prévost, 2009 |
| | | Catalan | Grinstead, 2000; 2009 |
| | | Portuguese | Pires & Rothman, 2009 |
| | | Romanian | Babyonyshev & Marin, 2006 |
| | Slavic | Russian | Slobin, 1965; Gordishevsky & Schaeffer, 2008; Gagarina & Voeikova, 2009 |
| | | Croatian | Kovačević, Palmović & Hržica 2009 |
| | | Serbian | Ilic & Deen, 2003 |
| | | Polish | Smoczynska, 1985 |
| | | Ukranian | Olbishevska, 2004 |
| | | Czech | Lukavský & Smolík, 2009 |
| | | Slovenian | Rus, 2008 |
| | Hellenic | Greek | Stephany, 1986; Stephany & Christofidou, 2007 |
| | Indo-Aryan | Hindi | Narasimhan, 2005 |
| | Iranian | Persian | Narafshan, Sadighi, Bagheri & Shokrpour, 2014 |

**Table 25.3** (Continued)

| *Classification* | | *Language* | *Sample Reference* |
|---|---|---|---|
| African | Bantu | Sesotho | Demuth, 1992 |
| | | Swahili | Deen, 2005 |
| | | Zulu | Suzman, 1991; 1996 |
| | | Xhosa | Gxilishe, de Villiers & de Villiers, 2007 |
| | | Siswati | Kunene, 1977 |
| | Nilo-Saharan | Luo | Blount, 1988 |
| | Niger-Congo | Yoruba | Sanusi & Arokoyo, 2011 |
| East/ Southeast Asian | Altaic | Japanese | Clancy, 1986; Otsu, 1994; Sano, 1995; Murasugi & Sugisaki, 2008 |
| | Isolate | Korean | Kim, 1997; Choi & Gopnik, 1995 |
| | Sino-Tibetan | Mandarin | Erbaugh, 1992 |
| | | Chintang | Stoll et al., 2012 |
| | Austroasiatic | Vietnamese | Tran, 2011 |
| Other | Finnic | Estonian | Argus, 2009 |
| | | Finnish | Laalo, 2009 |
| | Finno-Ugric | Hungarian | Dasinger, 1997; Gábor & Lukács, 2012 |
| | Semitic | Arabic | Aljenaie, 2010 |
| | | Hebrew | Berman, 1981; Schaeffer & Ben-Shalom, 2004 |
| | Eskimo-Aleut | Inuktitut | Allen & Crago, 1992 |
| | | West Greenlandic | Fortesque & Olsen, 1992 |
| | Algonquian | Cree | Rose & Brittain, 2011; The CCLAS Project |
| | Athabaskan | Navajo | Courtney & Saville-Troike, 2002 |
| | Quechuan | Quechua | Courtney, 2003 |
| | Mayan | K'iche' | Pye, 1992 |
| | Trans-New Guinea | Kaluli | Schieffelin, 1985 |
| | Altaic | Turkish | Aksu-Koc & Slobin, 1985; Ketrez & Aksu-Koc, 2009 |
| | Pama-Nyungan | Warlpiri | Bavin & Shopen, 1985 |
| | Kartvelian | Georgian | Imedadze & Tuite, 1997 |
| | Austronesian | Tagalog | Segalowitz & Galang, 1978; Galang, 1982; Tanaka et al. 2014 |
| | | Samoan | Ochs, 1982 |
| | | Malagasy | Hyams, Ntelitheos & Manorohanta, 2006 |
| | Dravidian | Tamil | Sarma, 2000; Lakshmanan, 2006; Raghavendra & Leonard, 1989 |
| | | Telugu | Nirmala, 1983 |
| | | Tulu | Somashekar, 1999 |
| | Isolate | Basque | Austin, 2012 |

described in this chapter, while informed by more than 50 languages, may in the end require modification and revision. It is my hope that research spurred by this chapter will do just that.

## NOTES

1 An infix is a morpheme that is inserted into the root of a word, as opposed to attached to the beginning (prefix) or end (suffix). Examples 2–3 below show Tagalog infixes. A circumfix is a morpheme that includes two parts: one that attaches to the beginning of a word and another that attaches to the end of a word.
2 This same phenomenon is found in other learning populations, such as L2 and bilinguals, and is referred to in that literature as "backsliding" (Ellis, 1994).

## REFERENCES


Aksu-Koc, A. & Slobin, D. (1985). The acquisition of Turkish. In: D. I. Slobin (Ed.), *The Cross-linguistic Study of Language Acquisition, Volume 1: The Data*. Hillsdale, N.J.: Erlbaum, (pp. 839–878).

Aljenaie, K. (2010). Verbal inflection in the acquisition of Kuwaiti Arabic. *Journal of Child Language*, 37(4), 841–863. doi: 10.1017/S0305000909990031

Allen, S. & Crago, M. (1992). First language acquisition of Inuktitut. In M.-J. Dufour & F. Thérien (Eds.), *Inuit studies occasional papers 4: proceedings of the seventh Inuit studies conference*, (pp. 273–281). Quebec: Association Inuksiutiit Katimajit.

Ambridge, B. & Lieven, E. (2011). *Child Language Acquisition: Contrasting theoretical approaches*. Cambridge, UK: Cambridge University Press. ISBN: 9780521745239

Ambridge, B., Pine, J. M., Rowland, C. F., Chang, F. & Bidgood, A. (2013). The retreat from overgeneralization in child language acquisition: Word learning, morphology and verb argument structure. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4, 47–62. doi: 10.1002/wcs.1207

Argus, R. (2009). The early development of case and number in Estonian. In: U. Stephany & M. Voeikova (Eds.), *Development of Nominal Inflection in First Language Acquisition A Cross-Linguistic Perspective,* (pp. 111–152). New York: Mouton de Gruyter.

Austin, J. (2012). The case-agreement hierarchy in acquisition: Evidence from children learning Basque. *Lingua*, 122, 289–302. doi: 10.1016/j.lingua.2011.10.013

Babyonyshev, M. & Marin, S. (2006). Acquisition of Pronominal Clitics in Romanian. *Catalan Journal of Linguistics*, 5, 17–44.

Bavin, E. & Shopen, T. (1985). Children's acquisition of Warlpiri: Comprehension of transitive sentences. *Journal of Child Language*, 12(3), 597–610. doi: 10.1017/S030500090000667X.

Becker, M. (2000). *The Development of the Copula in Child English: The Lightness of Be*. Doctoral dissertation, UCLA.

Berko-Gleason, J. (1958). The Child's Learning of English Morphology. *Word*, 150–177.

Berman, R. (1981). Language development and language knowledge: Evidence from the acquisition of Hebrew morphophonology. *Journal of Child Language*, 8(3), 609–626. doi: 10.1017/S0305000900003457.

Blom, E. (2007). Modality, infinitives and finite bare verbs in Dutch and English child language. *Language Acquisition*, 14(1), 75–113. http://www.jstor.org/stable/20462480

Blount, B. (1988). Cognition and Phonology in Acquisition of Plurals and Possessives by Luo Children. *Language Sciences*, 10(1), 225–240. doi: 10.1016/0388-0001(88)90013-7

Brown, R. (1973). *A First Language*. Cambridge, MA: Harvard University Press.

Brown, R. & Bellugi, U. (1964). Three processes in the child's acquisition of syntax. *Harvard Educational Review*, 34, 133–151.

Brown, R., Cazden, C., & Bellugi-Klima, U. (1968). The Child's Grammar from I do III. In: J. P. Hill (Ed.), *Minnesota Symposia on Child Psychology*, (pp. 28–73). Minneapolis.

Cazden, C. (1968). The acquisition of noun and verb inflections. *Child Development*, 39(2), 433–448. doi: 10.2307/1126956.

Choi, S. & Gopnik, A. (1995). Early acquisition of verbs in Korean: a cross-linguistic study. *Journal of Child Language*, 22(3), 497–529. doi: 10.1017/S0305000900009934.

Clahsen, H. (1986). Verb inflections in German child language: acquisition of agreement markings and the functions they encode. *Linguistics*, 24(1), 79–121. doi: 10.1515/ling.1986.24.1.79

Clahsen, H. & Penke, M. (1992). The acquisition of agreement morphology and its syntactic consequences: new evidence on German child language from the Simone-corpus. In: J. Meisel (Ed.), *The acquisition of verb placement: functional categories and v2 phenomena in language acquisition,* (pp. 181–223). Dordrecht: Kluwer.

Clancy, P. (1986). *The Acquisition of Japanese*. In D. I. Slobin (Ed.), *The Cross-linguistic Study of Language Acquisition, Vol. 1* (pp. 373–524). Hillsdale, NJ. Lawrence Erlbaum Associates.

Courtney, E. (2003). Child acquisition of Quechua morphosyntax. Doctoral dissertation, UMI, Ann Arbor.

Courtney, E., & Saville-Troike, M. (2002). Learning to construct verbs in Navajo and Quechua. *Journal of Child Language,* 29(3), 623–654. doi: 10.1017/S0305000902005160.

Crain, S. & Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. MIT Press.

Dasinger, L. (1997). Issues in the acquisition of Estonian, Finnish and Hungarian: A cross-linguistic comparison. In D. Slobin (Ed.), *Cross-linguistic study of language acquisition*, volume 4, (pp. 1–86).

Deen, K. U. (2005). *The Acquisition of Swahili*. Amsterdam, Netherlands: John Benjamins Publishing.

Demuth, K. (1992). The acquisition of Sesotho. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition volume 3*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

de Villiers, J. G. & de Villiers, P. A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, 2(3), 267–278.

Dressler, W. (2007). Introduction. In S. Laaha & S. Gillis (Eds.), *Typological perspectives on the acquisition of noun and verb morphology, Antwerp Papers In Linguistics, Vol. 112*. (pp. 3–9). Antwerp: University of Antwerp.

Dressler, W., Stephany, U., Aksu-Koc, A., & Gillis, S. (2007). Discussion and

Conclusion. In S. Laaha & S. Gillis (Eds.), *Typological perspectives on the acquisition of noun and verb morphology, Antwerp Papers In Linguistics, Vol. 112* (pp 67–73). Antwerp: University of Antwerp.

Ellis, R. (1994). *The Study of Second Language Acquisition.* Oxford: Oxford University Press.

Erbaugh, M. (1992). The acquisition of Mandarin. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition volume 3* (pp. 373–455). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Fortescue, M. & Olsen, L. (1992). The acquisition of West Greenlandic. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition*, vol. 3, (pp. 111–220). Hillsdale, NJ: Lawrence Erlbaum.

Gábor, B. & Lukács, A. (2012). Early morphological productivity in Hungarian: evidence from sentence repetition and elicited production. *Journal of Child Language*, 39(2), 411–442. doi: 10.1017/S0305000911000110.

Gagarina, N. & Voeikova, M. D. (2009). Acquisition of case and number in Russian. In U. Stephany & M. Voeikova (Eds.), *Development of Nominal Inflection in First Language Acquisition A Cross-Linguistic Perspective,* (pp. 179–216). New York: Mouton de Gruyter.

Galang, R. (1982). Acquisition of Tagalog verb morphology: Linguistic and cognitive factors. *Philippine Journal of Linguistics*, 13(2), 1–15.

Gordishevsky, G. & Schaeffer, J. C. (2008). The development and interaction of Case and Number in early Russian. In P. Guijarro-Fuentes, M. P. Larrañaga, & J. Clibbens (Eds.), *First Language Acquisition of Morphology and Syntax: Perspectives Across Languages and Learners*, Amsterdam: John Benjamins, pp. 31-59.

Grinstead, J. (2000). Case, inflection and subject licensing in child Catalan and Spanish. *Journal of Child Language,* 27(1), 119–155.

Grinstead, J. (2009). *Hispanic Child Languages: Typical and Impaired Development*. Amsterdam: John Benjamins.

Guasti, M. T. (1993/94). Verb syntax in Italian child grammar: Finite and nonfinite verbs. *Language Acquisition,* 3(1), 1–40. doi: 10.1207/s15327817la0301_1

Gxilishe, S., de Villiers, P., & de Villiers, J. (2007). The Acquisition of Subject Agreement in Xhosa. In A. Belikova *et al.* (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 114–123). Somerville, MA: Cascadilla Proceedings Project.

Hamann, C. & Plunkett, K. (1998). Subjectless sentences in child Danish. *Cognition,* 69, 35–72. doi: 10.1016/S0010-0277(98)00059-6

Henden, Å. K. (2013). Early lexical and grammar development in Norwegian language Acquisition. Doctoral dissertation, Norwegian Institute of Science and Technology.

Hoekstra, T., & Hyams, N. (1998). Aspects of root infinitives. *Lingua* 106, 91–112.

Hyams, N. (1986). Language acquisition and the theory of parameters. Dordrecht: Reidel.

Hyams, N., Ntelitheos, D., & Manorohanta, C. (2006). Acquisition of the Malagasy voicing system: Implications for the adult grammar. *Natural Language and Linguistic Theory,* 24(4), 1049–1092. doi: 10.1007/s11049-006-9009-2

Ilic, T. & Deen, K. U. (2003). Object raising and Cliticization in Serbo-Croatian child language. In A. van Kampen & S. Baauw (Eds.), *The Proceedings to the GALA Conference* (pp. 235–243). LOT: Utrecht.

Imedadze, N. & Tuite, K. (1997). The acquisition of Georgian. In D. Slobin (Ed.), *The Cross-linguistic Study of Language Acquisition, Volume 4*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Josefsson, G., Platzack, C., & Håkansson, G. (2004). *The Acquisition of Swedish Grammar.* Amsterdam: Netherlands: John Benjamins.

Kelly, B., Wigglesworth, G., Nordlinger, R., & Blythe, J. (2013). The acquisition of polysynthetic languages. *Language and Linguistics Compass*, 8, 51–64. doi: 10.1111/lnc3.12062

Ketrez, F. N. & Aksu-Koc, A. (2009). Early nominal morphology in Turkish: Emergence of case and number. In M. Voeikova and U. Stephany (Eds.), *The Development of Number and Case in the First Language Acquisition: A Cross-Linguistic Perspective* (pp. 15–48). Berlin: Mouton de Gruyter.

Kim, C. E., O'Grady, W., & Deen, K. U. (2014). The Korean extrinsic plural marker. *Korean Linguistics*, 16(1), 1–17. doi: 10.1075/kl.16.1.01kim

Kovačević, M., Palmović, M., Hržica, G. (2009). The acquisition of case, number and gender in Croatian. In U. Stephany & M. Voeikova (Eds.), *Development of Nominal Inflection in First Language Acquisition: A Cross-Linguistic Perspective* (pp. 153–178). New York: Mouton de Gruyter.

Kunene, E. (1979). The acquisition of SiSwati as a first language: A morphological study with special reference to noun prefixes, noun classes, and some agreement markers. Doctoral dissertation, UCLA.

Laalo, K. (2009). Acquisition of case and plural in Finnish. In U. Stephany & M. Voeikova (Eds.), *Development of Nominal Inflection in First Language Acquisition A Cross-Linguistic Perspective* (pp. 49–90). New York: Mouton de Gruyter.

Lakshmanan, U. (2006). Assessing linguistic competence: Verbal inflection in child tamil. *Language Assessment Quarterly*, 3(2), 171–205. doi: 10.1207/s15434311laq0302_5

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, 20, 481–507. doi: 10.1515/COGL.2009.022

Lightbrown, P. (1977). Consistency and variation in the acquisition of French. Doctoral dissertation, Columbia University.

Lukavský, J. & Smolík, F. (2009). Word order and case inflection in Czech: On-line sentence comprehension in children and adults. *Proceedings of Cogsci*, 1363–1358.

Lukyanenko, C. & Fisher, C. (2014). 30-month-olds use verb agreement features in online sentence processing. In W. Orman & M.J. Valleau (Eds.), *Proceedings of the 38th annual Boston University Conference on Language Development*, (pp. 292–305). Boston, MA: Cascadilla Press.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. Third edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Maratsos, M. (2000). More overregularizations after all: New data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen & Xu. *Journal of Child Language*, 27(1), 183–212.

Marcus, G., Pinker, S., Ullman, M., Hollander, M., Rosen, J., & Xu, F. (1992). *Overregularization in language acquisition*. Chicago: University of Chicago Press.

Mealings, K. & Demuth, K. (2014). The role of utterance length and position in 3-year-olds' production of third person singular *–s*. *Journal of Speech, Language, and Hearing Research*, 57, 484–494. doi: 10.1044/2013_JSLHR-L-12-0354

Murasugi, K. & Sugisaki, K. (2008). The acquisition of Japanese syntax. In M. Shigeru (Ed.), *The Oxford Handbook of Japanese Linguistics*. doi: 10.1093/oxfordhb/9780195307344.013.0010

Narafshan, M. H., Sadighi, F., Sadegh Bagheri, M., Shokrpour, N. (2014). The role of input in first language acquisition. *International Journal of Applied Linguistics & English Literature*, 3(1), 86–91.

Nirmala, C. (1983). Development of plural in Telugu children. *Osmania Papers in Linguistics*, 9–10, 1–20.

Ochs, E. (1982). Ergativity and word order in Samoan child language, *Language*, 58(3), 646–671. doi: 10.2307/413852

Olbishevska, O. (2004). Acquisition of aspect (lexical vs. grammatical) by Ukrainian children. *Cahiers linguistiques d'Ottawa*, 32, 66–86.

Otsu, Y. (1994). Case marking particles and phrase structure in early Japanese. In B. Lust, M. Suner, and J. Whitman (Eds.), *Syntactic Theory and First Language Acquisition: Cross-linguistic Perspectives* (pp. 159–169), Hillsdale, NJ: Lawrence Erlbaum Associates.

Penke, M. (2012). The acquisition of inflectional morphology. In A. Spencer & A. M. Zwicky (Eds.), *The Handbook of Morphology*. Oxford: Blackwell.

Peters, A. (1983). *The Units of Language Acquisition*. Cambridge: Cambridge University Press.

Peters, A. (1997). Language typology, prosody and the acquisition of grammatical morphemes. In D. Slobin (Ed.), *The Cross-linguistic Study of Language Acquisition, vol.* 5, (pp. 136–197). Hillsdale NJ: Lawrence Erlbaum Associates.

Pierce, A. (1989). On the emergence of syntax: A cross-linguistic study. Doctoral dissertation, MIT.

Pires, A. & Rothman, J. (2009). Minimalist Inquiries into Child and Adult Language Acquisition. Berlin/New York: Mouton de Gruyter.

Pizutto, E. & Casello, C. (1992). The acquisition of Italian morphology: Implications for models of language development. *Journal of Child Language*, 19(3), 491–557. doi: 10.1017/S0305000900011557.

Poeppel, D. & Wexler, K. (1993). The full competence hypothesis of clause structure in Early German. *Language*, 69(1), 1–33, http://www.jstor.org/stable/416414.

Prévost, P. (2009). The acquisition of french: The development of inflectional morphology and syntax in L1 acquisition, bilingualism, and L2 acquisition. Amsterdam, Netherlands: John Benjamins.

Pye, C. (1992). The acquisition of K'iche' Maya. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition*, vol. 3, (pp. 221–308). Hillsdale, NJ: Lawrence Erlbaum.

Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammars of English*. Basil Blackwell: Oxford

Raghavendra, P., & Leonard, L. (1989). The acquisition of agglutinating languages: Converging evidence from Tamil. *Journal of Child Language*, 16(2), 313–322. doi: 10.1017/S0305000900010436.

Ragnarsdóttir, H., Simonsen, H., & Plunkett, K. (1999). The acquisition of past tense morphology in Icelandic and Norwegian children: An experimental study. *Journal of Child Language*, 26(3), 577–618.

Rose, Y. & Brittain, J. (2011). Grammar matters: Evidence from phonological and morphological development in Northern East Cree. In M. Pirvulescu (Ed.), *Proceedings of the 4th Conference on Generative Approaches to Language Acquisition in North America (GALANA)*, (pp. 193–208). Somerville, MA: Cascadilla Press.

Rubino, R. & Pine, J. (1998). Subject agreement in Brazilian Portuguese: What low error rates hide. *Journal of Child Language*, 25(1), 35–59.

Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tense of English verbs. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 2: Psychological and Biological Models*, (pp. 216–271). Cambridge, MA: MIT Press.

Rus, D. (2008). The acquisition of verbal inflection in child grammars in a variability model of early morphosyntactic development: A biolinguistic perspective. Doctoral dissertation, Georgetown University.

Sano, T. (1995). *Roots in Language Acquisition: A Comparative Study of Japanese and European Languages*, Doctoral Dissertation, University of California, Los Angeles.

Sano, T. & Hyams, N. (1994). Agreement, finiteness, and development of null arguments. In M. Gonzalez (Ed.), *The Proceedings to NELS 24*, (pp. 543–558). Amherst, MA: GLSA.

Sanusi, I. & Arokoyo, B. (2011). Acquisition of early verbs by Yoruba children: An analysis of a longitudinal study. *Ilorin Journal of Linguistics, Literature and Culture,* 2, 1–26.

Sarma, V. (2000). Case, agreement and word order: Issues in the syntax and acquisition of Tamil. Doctoral dissertation, MIT.

Schaeffer, J. & Ben-Shalom, D. (2004). On root infinitives in child Hebrew. *Language Acquisition,* 12(1), 83–96. doi: 10.1207/s15327817la1201_4

Schieffelin, B. (1985). The acquisition of Kaluli. In D. I. Slobin (Ed.), *The Cross-linguistic Study of Language Acquisition, Volume 1: The Data.* Hillsdale, NJ: Erlbaum.

Segalowitz, N. & Galang, R. (1978). Agent-patient word-order preference in the acquisition of Tagalog. *Journal of Child Language*, 5(1), 47–64. doi: 10.1017/S0305000900001938.

Slobin, D. (1966). The acquisition of Russian as a native language. In F. Smith & G. A. Miller (Eds.), *The Genesis of Language: A Psycholinguistic Approach*, (pp. 129–148). Cambridge University Press: Cambridge MA.

Slobin, D. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of Child Language Development,* (pp. 175–208). New York: Holt, Rinehart, & Winston.

Smoczynska, M. (1985). The acquisition of Polish. In D. I. Slobin (Ed.), *The Cross-linguistic Study of Language Acquisition, Volume 1: The Data*. Hillsdale, NJ: Erlbaum.

Somashekar, S. (1999). *Developmental trends in the acquisition of relative clauses: Cross-linguistic experimental study of Tulu*. Doctoral dissertation, Cornell University.

Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular -s. *Journal of Speech, Language, and Hearing Research*, 52(3), 623–642. doi: 10.1044/1092-4388(2008/07-0258)

Stephany, U. (1995). The acquisition of Greek. In D. I. Slobin (Ed.), *The Cross-linguistic Study of Language Acquisition, Volume 4* (pp. 183–333). Hillsdale, NJ: Lawrence Erlbaum.

Stephany, U. & Christofidou, A. (2007). The emergence of nominal inflection in Greek. In U. Stephany & M. Voeikova (Eds.), *Development of Nominal Inflection in First Language Acquisition: A Cross-Linguistic Perspective,* (pp. 217–264). New York: Mouton de Gruyter.

Stoll, S, Bickel, B., Lieven, E., Paudyal, N., Bandaje, G., Bhatta, T., Gaenszle, M., Pettigrew, J., Purna Rai, I., Rai, M., & Rai, N. A. (2012). Nouns and verbs in Chintang: Children's usage and surrounding adult speech. *Journal of Child Language*, 39(2), 284–321. doi: 10.1017/S0305000911000080.

Suzman, S. (1991). Language acquisition in Zulu. Doctoral dissertation, University of the Witwatersrand.

Suzman, S. (1996). Acquisition of noun class systems in related Bantu languages. In C. E. Johnson & J. H. V. Gilbert (Eds.), *Children's Language, Volume 9,* (pp. 87–104). Mahwah, NJ, Lawrence Erlbaum Associates.

Tanaka, N., O'Grady, W., Deen, K. U., Eun Kim, C., Hattori, R., Bondoc, I., & Soriano, J. (2014). Acquisition of Tagalog relative clauses. In W. Orman & M. J. Valleau (Eds.), *Proceedings of the 38th annual Boston University Conference on Language Development,* (pp. 463–470). Somerville, MA: Cascadilla Press.

Torrens, V. (1995). L'adquisició de la sintaxi en català i castellà: La categoria functional de flexió. Doctoral dissertation, Universitat de Barcelona.
Tran, J. (2011). The acquisition of Vietnamese classifiers. Doctoral dissertation, University of Hawaii at Manoa.
Tsimpli, I. (1992). Functional categories and maturation: The prefunctional stage of language acquisition. Doctoral dissertation. London: University of London.
Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, 40(1–2), 21–81. doi: 10.1016/0010-0277(91)90046-7.
Wagner, K. R. (1985). How much do children say in a day? *Journal of Child Language*, 12, 475–487.
Westergaard, M. (2008). Verb movement and subject placement in the acquisition of word order: Pragmatics or structural economy? In P. Guijarro-Fuentes, P. Larranaga, & J. Clibbens (Eds.), *First Language Acquisition of Morphology and Syntax: Perspectives Across Languages and Learners*. [Language Acquisition and Language Disorders 45], (pp. 61–86). Amsterdam: John Benjamins.
Wexler, K. (2004). Theory of phasal development: Perfection in child grammar. *MIT Working Papers in Linguistics*, 48, 159–209.
Wijnen, F. (1997). Temporal reference and eventivity in root infinitivals. In J. C. Schaeffer (Ed.), *The Interpretation of Root Infinitives and Bare Nouns in Child Language*. *MIT Occasional Papers in Linguistics*, 12, 1–25.
Wijnen, F. & Verrips, M. (1998). The acquisition of Dutch syntax. In S. Gillis & A. De Houwer (Eds.), *The Acquisition of Dutch*. Amsterdam: Benjamins.
Zapf, J. & Smith, L. (2007). When do children generalize the plural to novel nouns? *First Language*, 27(1), 53–73. doi: 10.1177/0142723707070286

# 26 The Acquisition of Syntax

NINA HYAMS[1] AND ROBYN ORFITELLI[2]

[1] UCLA
[2] The University of Sheffield

## Introduction

How children acquire their native language(s) is a central concern of linguistic theory and cognitive science more generally. There is a clear consensus that humans are specifically adapted to the task of language acquisition and that language development is driven by innate capacities and shaped by the environment. Though there are many different views on the nature of the innate mechanisms, as well as the degree of environmental influence, any explanation must reckon with the fact that the language/grammar that children acquire is vastly underdetermined by the linguistic input they receive. Children are exposed to adult utterances, not abstract grammatical structures, yet acquire a hierarchically organized system of categories and rules. Children hear only a finite number of adult utterances, yet have the capacity to produce and understand sentences way beyond this restricted set—including many that are not grammatical in the adult "input" language. Lastly, children are provided with "positive" exemplars (possible sentences in his language), yet somehow come to also know what is *not* possible in their language. In short, every typically developing child develops a rich and highly abstract system of rules that constitutes the grammar of their language. They do this in a very short period of time, based on rather limited evidence, without the benefit of explicit correction, instruction, or information about ungrammaticality, and much of what they come to know is not transparently exemplified in the language of the adults around them. The acquisition problem couched in these terms is often referred to as the *logical problem of language acquisition* (LPLA) (Chomsky, 1965; Baker & McCarthy, 1981).

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

The LPLA can be illustrated using the following sentences:

(1) a. Ernie looks like he rode his bike to school today.
    b. It looks like Ernie rode his bike to school today.

The sentences in (1) seem to be simple paraphrases of each other, and indeed, in certain situations either sentence could be felicitously uttered. For example, imagine Ernie is standing next to his bicycle in front of a school building, wearing a bicycle helmet. In that case we might equally well utter sentence (1a) or (1b). However, now imagine we're looking at a classroom. In it we see Ernie's desk and his bicycle helmet is under the desk, but the classroom is empty because all the children are at recess. In this case sentence (1b) is still a perfectly appropriate utterance, but (1a) is decidedly odd, because we have no visible evidence of Ernie himself. The contrast is subtle. Both sentences are fully grammatical, but they are not equivalent. They have different evidential requirements. Clearly, this is not something children are instructed on. Moreover, it is unlikely that they have any kind of (negative) evidence, which would lead them to know that (1a) is *not* felicitous in certain contexts. Yet, children as young as two years old use both kinds of sentences in the appropriate situations (Rett & Hyams, 2014; Rett, Hyams, & Winans, 2012).

There are many such cases in natural language, sentences that should be possible but are not because they are blocked by some grammatical constraint. In the examples in (1) the constraint is semantic in nature—the determiner phrase (DP) subject of the matrix clause must be the perceptual source for the assertion (Asudeh & Toivonen, 2012). In other cases, such as those discussed below, the constraint is syntactic. In all cases, the linguistic evidence for the constraint that is available to the child is slim or non-existent.

Most linguists assume that the solution to the LPLA lies in the theory of Universal Grammar (UG), a set of innately specified grammatical principles that provides a blueprint for human languages, and that "guides" the children's language development by restricting their grammar-forming options. While the exact form of UG is open to empirical investigation, it will ideally explain those properties that are invariant across languages, as well as cross-linguistic differences that are easily describable within a restricted parameter space. Most crucially, UG must be sufficiently articulated so that—together with the available input evidence—it provides an acquisition path to any (and all) target grammars.

But the path to the target grammar is not without some curves and potholes, and it is not instantaneous. Children go through fairly well-defined stages, including the well-studied null subject and root infinitive stages discussed later in this chapter, in which their productions deviate from the adult's in systematic ways. Similarly, they do not necessarily have an adult understanding of all sentence structures, among these certain types of passives, raising and control structures. We will discuss each of these phenomena below. A comprehensive picture of the child's grammatical development means understanding the initial state—the principles that constrain development—and also the nature of their linguistic "errors" and stages. We can refer this as the *developmental problem of language acquisition*. An important aspect of the developmental problem is determining which properties of early language are due to representational differences between the child and the adult and which properties result from the child's more limited language processing resources. In the latter case the child may have the adult grammatical representation but be unable to produce or understand a construction due to limitations of working memory, sentence planning, articulatory control, and so on. The interacting effects of linguistic competence and performance (Chomsky, 1965) are nowhere better illustrated than in child language.

In the following sections we briefly describe several aspects of children's syntactic development. Given the limitations of space we do not intend this to be a comprehensive review. Rather, we have chosen facets of grammar acquisition that have been especially well studied and which represent different developmental stages and processes. Additionally, we focus on areas that allow us to illustrate the different kinds of data that inform theories of children's grammars—naturalistic production data as well as experimental results. Finally, in line with the considerations just discussed, we chose topics that illustrate the challenges in deciding between competence-based or performance-based accounts of particular developmental phenomena.

## Early multi-word utterances: The root infinitive stage

At around two years of age children begin to produce multi-word utterances. This might properly be considered the first stage of syntax acquisition in production, although comprehension of syntax begins earlier (see, e.g., Golinkoff & Hirsch-Pasek, 1996). Children's earliest utterances are short (typically two to three words) and consist largely of open class lexical elements (e.g., nouns, verbs, adjectives), with relatively few closed class functional elements (e.g., articles, auxiliaries, prepositions). This early language is traditionally referred to as "telegraphic speech" (Brown, 1973). Morphological development in the telegraphic stage is described in detail by Deen (this volume); here we will limit our discussion to one of its most notable morphosyntactic features: *root infinitives* (RIs; Rizzi, 1993/4), non-adultlike sentences in the main verb occurs in its infinitival form. (Because children produce RIs alongside finite forms, the stage is also referred to as the *optional infinitive stage* (Wexler, 1994)). Rates of RI production vary by child and language, ranging from 26% to 61% of all (finite and non-finite) verbal utterances (Hoekstra & Hyams, 1998).

In (2) we provide examples of RIs and finite verbs from several of the languages that show an RI stage. We focus on languages other than English because, strictly speaking, English has no morphological infinitive. What passes for an infinitive, a bare verb, as in (2f), can also result from the simple dropping of the finite ending (see Song, Sundara, & Demuth, 2009 for evidence that phonological factors are at play in the production of English bare verbs).

(2) a. Hun sove. (Danish: Jens 2;0, Hamann & Plunkett, 1998)
She sleep-INF
"She sleeps."
b. Dormir petit bébé. (French: Daniel 1;11, Pierce, 1989)
sleep-INF little baby
"Little baby sleep."
c. Earst kleine boekje lezen. (Dutch: Hein 2;6, Haegeman, 1995)
First little book read-INF
"First (I/we) read little book."
d. S[ch]okolade holen. (German: Andreas 2;1, Krämer, 1993)
chocolate get-INF
"I got chocolate (?)"
e. Ty mama pomogat' (Russian: Vavara 2;0, Brun & Babyonyshev, 2006 )
You.nom mama.dat help-INF
"You (=I) help mommy"
f. Cromer wear glasses (English: Eve 2;0, Brown, 1973)

Additionally, the co-occurrence of non-finite and finite forms during the same period strongly suggests that RIs are not simply a "default" form that children use when they do not know inflected forms (cf. Blom, 2003). Indeed, there are many properties that distinguish the two clause types. The various ways in which RIs differ from finite clauses are outlined in (3):

(3) RIs typically:
(i) have infinitival morphology in languages with an infinitival form (e.g., 2a-e)
(ii) do not occur with subject clitics in French.
(iii) do not occur as copula or auxiliaries; the copula and auxiliaries are always finite.
(iv) appear in sentence final position in OV/V2 languages such as Dutch and German (cf. 2c-d), whereas finite verbs occur in second position.
(v) do not occur in *wh*-questions; over 97% of *wh*-questions are finite in Dutch, German Swedish, French.
(vi) do not occur with unambiguous (i.e., object, adverb, prepositional phrase) topics in languages with topicalization such as Dutch, German, Icelandic.
(vii) occur only with eventive verbs, not with statives.
(viii) occur overwhelmingly with null subjects (NS) (e.g., 2c-d); the NS rate in RIs is 75% to 90% compared to 8 to 32% NS rate in finite clauses.
(ix) do not occur (or occur at much lower rates) in null subject languages such as Italian and Spanish; in these languages fewer than 15% of verbal utterances are RIs and typically the rate is in the single digits.

The properties associated with RIs are reviewed in detail by Hoekstra and Hyams (1998). English bare verbs (cf. 2f) fail to show some of the properties in (3) that

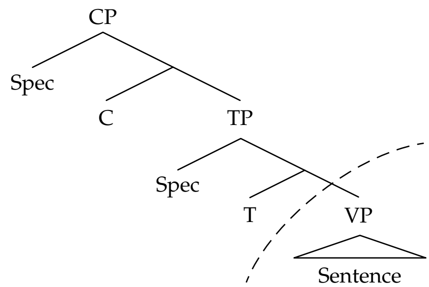



**Figure 26.1** VP truncation model of root Infinitives: Full CP projection vs. VP truncation.

would be expected if they are truly non-finite. For example, bare verbs occur in *wh*-questions (cf. 3v), with stative and eventive verbs (cf. 3vii), and the rate of null subjects is roughly the same for bare and finite verbs (cf. 3viii).

Various grammatical explanations for the RI stage have been proposed. Of particular interest is Rizzi's (1994) hypothesis that the grammar of young children need not project a full complementizer phrase (CP) structure, but instead can generate structures in which any maximal projection (XP) constitutes a licit root. RIs result when the structure is truncated below the tense phrase (TP), arguably at the verb phrase (VP); finite clauses result from a full CP projection, or given assumptions of an expanded left periphery (e.g., Rizzi, 1997), whatever projection is taken to be the highest. The *truncation hypothesis*, schematized in Figure 26.1, successfully accounts for many, if not all of the properties in (3).

Without an inflectional system (including tense, agreement, and auxiliaries) there is no finite morphology, subject clitics (as in French), or auxiliary verbs; without a CP layer there can be no *wh*-questions, topicalizations or movement of the verb to C (as in German and Dutch). Although Rizzi does not directly address the eventivity effect in (3vii), it is easily accountable under a theory in which eventive and stative verbs have a distinct syntax, with the latter involving higher functional structure, as proposed by Borer (2005) and elsewhere under different assumptions. The NS facts in (3viii) and a lack (or greatly reduced) RI stage in null subject languages such as Italian and Spanish (3ix) require some additional assumptions, but can also be handled by Rizzi's hypothesis.

Rizzi assumes that truncation is a UG-compliant option that also occurs in particular registers of adult grammars (4). For adults, special pragmatic or semantic factors are often required, but it would seem that children can avail themselves of the truncation option without such pragmatic licensing.

(4) a. John marry my sister? Never.
    b. Hier geen fietsen plaatsen! (Dutch)
       here no bicycles place-inf
       "Don't put bicycles here"

The majority of accounts of the RI stage are competence-based; that is to say, they propose that the child's grammar differs from the adult grammar in licensing RIs. Alternatives to the truncation theory include, for example, the Null Aux hypothesis (Boser *et al.*, 1992) and the Agreement-Tense Omission Model (ATOM, Schutze & Wexler, 1996), both of which propose a difference between the children's grammar and the target. In recent years, however, an increasing number of researchers have appealed to performance limitations as a possible explanation for the RI stage. According to Dye (2011), French-speaking children in the RI stage are attempting to produce a periphrastic structure (auxiliary/modal+non-finite verb) as in (5a), but often fall short of the target and produce a phonologically reduced form of the auxiliary, as in (5b) or delete it altogether, as in (5c), giving the appearance of a non-finite root clause.

(5) a. Je veux jouer avec le château fort moi
I want.1st p. play.inf. with the fortress me
"I want to play with the fortress."
b. J'v'ouer avec le tast fort moi
c. Rentre na maison
return the home
(= il est rentrè/va rentrer à la maison – "he entered/is going to enter the house")

Dye presents interesting arguments in support of the phonological reduction hypothesis including acoustic evidence of "covert" auxiliaries in some RIs. What is left largely unexplained by this kind of account are several of the findings in (3). For example, why is there not a robust RI stage in Italian and Spanish, as these languages are closely related to French? Why do we fail to find RIs in topicalized sentences and *wh*-questions in V2 languages? Finally, why are RIs restricted to eventive predicates? Although a processing-based account might in theory be able to explain these patterns, further work remains to be done.

Rather than an exclusively grammatical or performance account, it is possible that RIs result from an interaction of grammatical and performance factors. For example, Rizzi's grammatical truncation analysis is also consistent with a hypothesis in which children produce reduced (RI) structures under the pressure of more limited production or processing resources. In effect, they take advantage of a grammatical option that imposes a reduced computational load. Indeed, the notion that certain child language phenomena arise from a complex interaction of grammatical and processing factors has also been proposed by Rizzi (2005) as an explanation for another salient property of early language – the null subject stage. We turn to children's null subjects in the next section.

## Null subjects in early language

Children in the telegraphic stage also omit subjects, even where it is not a grammatical option in the adult language. This behavior not only occurs during a similar time period to the RI stage, but also in an overlapping set of languages (6).

(6) a. Falled in the briefcase. (English: Eve 1;10, Brown, 1973)
b. Ikke tøre traktor. (Danish: Jens 2;0, Hamann & Plunkett, 1998)
not drive tractor
c. Eerst kleine boekje lezen. (Dutch: Hein 2;6, Haegeman, 1995)
first little book read
d. Va sous la tabouret. (French: Philippe 2;2, Suppes *et al.*, 1973)
goes under the stool
e. Hubsauber putzn. (German: Andreas 2;1, Kramer, 1993)
helicopter clean

Like RIs, subject omission is optional; during the same period in which children drop subjects in some sentences, they also produce them in many others (Hyams, 1986). The frequency of subject omission varies across individual children, across languages, and within the same child across time. (Similar questions arise vis á vis the "null object" stage, in which children acquiring Romance languages omit objects in environments where adults would use a clitic pronoun. (See Mateu, 2014 for discussion).

Explanations of the null subject (NS) stage fall into two broad categories. Competence-based accounts take children's erroneous productions to reflect a difference between the child and adult grammars (Hyams, 1986, 1992; Hyams & Wexler, 1993; Rizzi, 1994, 2000; Yang, 2002, 2004, among others), although these accounts differ in various ways (e.g., Rizzi's account implicates truncation and Yang's contains an important statistical component). Conversely, performance accounts hold that subject omission is due to extra-syntactic factors (L. Bloom, 1970; P. Bloom, 1990; Gerken, 1991; Valian, 1991). On this view, children's syntactic representations include an overt subject, just as in the adult representation, but the subject is dropped in production because of these other influences. Finally, as noted in the previous section, Rizzi (2005) has proposed a hybrid model in which both competence and performance factors conspire to explain children's subject omissions. Given the parametric option of null versus overt subjects young children will choose the computationally least expensive option.

While competence- and performance-based accounts of the NS stage both predict that children will drop subjects in production (for a discussion of other issues related to comprehension versus production, see Hyams, 2011 and McKee, McDaniel, & Garrett, this volume), they categorically differ in their predictions of how children will *comprehend* NS sentences. Competence accounts predict that a child who produces NS sentences allows them to have a declarative interpretation. Performance accounts, however, predict that children should interpret NS utterances in an adult manner, which in English means only as imperatives.

Orfitelli and Hyams (2012) tested children's comprehension of NS sentences with a modified version of the Truth-Value Judgment (TVJ) experiment (Crain & McKee, 1985; Crain & Fodor, 1993). The task tested whether they would permit an NS sentence to have a declarative interpretation in addition to (or instead of) an adult-like imperative one.

Participants saw pairs of pictures depicting two sets of children. The first picture showed two older children engaged in a particular activity, such as drawing a picture or playing with blocks, while the second picture always showed two younger children in close proximity to the relevant items (e.g., paper and crayons or blocks) but not interacting with them. Participants were told that the four children have the same babysitter, and that while the older children are old enough to choose their own activities without permission, the younger children had to wait for the babysitter to tell them to engage in the activity in question. This set up a mood-based dichotomy: it is appropriate to use a declarative sentence in reference to the picture of the older children, but not the younger children; conversely, it is appropriate to use an imperative sentence in reference to the picture of the younger children, but not the older children. Strictly speaking, judging the appropriateness of imperative sentences does not involve computing a "truth-value," however, these items nonetheless test participants' *comprehension* of the experimental scenarios, which sharply distinguishes this task from metalinguistic receptive tasks such as the grammaticality/acceptability judgment task.

After each story, one of the two pictures was removed, and an observing puppet made a comment about the remaining picture. Some of these comments took the form of NS sentences (7). The participant was asked to tell the puppet if his comment was correct or incorrect.

(7) Play with blocks.

If children accept NS sentences when paired with the "older-children" scenario, it indicates that they permit a declarative syntactic structure for these sentences, which is fundamentally different from the target adult grammar. This is exactly what Orfitelli and Hyams found. Children in the NS stage accept both declarative and imperative interpretations for NS sentences, but stopped allowing the non-adult declarative interpretation at approximately 3½ years. This shift mirrors the one that has been seen in production, and indeed, we also found a near-perfect within-subject correspondence between NS production and non-adult comprehension.

This connection is not expected on a performance account, but is explicitly predicted by a grammatical account, as both behaviors would arise from the same root cause: the child's different grammar. A performance account that attributes NSs solely to planning or production limitation (e.g., Bloom, 1991; Valian, 1991) cannot explain why children accept NSs in comprehension, and in particular why they do so under a non-adult (=declarative) interpretation. However, the pattern of results in not entirely straightforward. If children's grammar licenses an ambiguity between imperative and declarative readings for NS sentences, we might expect them to answer *true* to all NS items, regardless of scenario. Instead, the children appeared to resolve the ambiguity in one direction or the other, and only then consider whether or not this mood matched the scenario that had just been presented. Approximately 50% of their answers seemed to indicate an adult, imperative interpretation for the NS items, while the other 50% of their answers seemed to indicate a non-adult, declarative interpretation.

The participants' justification responses support this conjecture. On the imperative and declarative control conditions, children provided justifications that unambiguously indicated the correct mood; for example, justifications for imperative controls used phrases such as "*supposed to*" or "*because he is telling them to do X*," but the justifications for the declarative control items were always simple declaratives. On the NS condition, participants' responses matched these two patterns. When answering correctly, their justifications always resembled their responses to the imperative controls (8a), which is expected if the children were assigning imperative mood to the NS sentence. Conversely, when answering incorrectly, their justifications matched those for the declarative controls, as if they were assigning declarative mood to the sentence (8b).

(8) a. He needs to tell the younger kids to put on socks. –S., 3;3
Child correctly answers true to "Put on socks."
b. Because they aren't eating a cookie –E., 3;10
Child incorrectly answers false to "Eat a cookie."

For any given NS test item, why were children allowing only a declarative *or* imperative interpretation, given the evidence that their grammar licenses both? This may be where processing factors may come into play. Young children are noted to have difficulty using context to disambiguate ambiguity in language comprehension tasks. Not only does this behavior show up in lexical ambiguity resolution (Swinney & Prather, 1989), but also in sentence-level attachment ambiguity (Trueswell *et al.*, 1999) and scopal ambiguity (Syrett & Lidz, 2005). The consequence of this is that in contrast to adults, who are able to revise incorrect parses to reflect additional contextual and other information, children appear to be bound by their original parse.

If NS sentences present an ambiguity for children, then when attempting to assign a meaning to a NS utterance, the children have to decide between a declarative and imperative representation for the sentence, evaluate the representation relative to the context of the experimental situation, and revise their representation when necessary. This revision process is precisely the kind that children have been shown to have difficulty with in comprehension. The key difference is that the ambiguity inherent in the NS stage is one that exists only for children up to the age of 3½ years, unlike attachment or scopal ambiguities, which are also ambiguous for adults. When children exit the NS stage, their grammar changes, and the processor is no longer faced with the problem of ambiguity resolution.

Note that these issues do not seem to impact language production in the same way, and indeed, whether a child in the NS stage produces an overt versus null subject seems to be sensitive to pragmatic-discourse factors, including givenness, person, and animacy (Hughes & Allen, 2008).

Overall, the experimental findings demonstrate that children in the NS stage have a non-target grammar. Additionally, while performance factors are not the underlying cause of the NS stage, they appear to play an important role in the resolution of the mood ambiguity associated with NS utterances. Children's processing resources do not allow them to integrate all sources of information needed to resolve the ambiguity and this leads them to assign one single parse and then stick with it.

## Later language development: A-movement

Once children have moved beyond telegraphic speech (including the RI and NS stages), their language development is incredibly rapid, and by the time they are five years old, they make very few errors in production. However, comprehension data show that children still fail to assign a target representation to a small set of sentence structures. Do these comprehension errors reflect a non-target grammar, and if so, is it consistent with UG? Conversely, are the errors due to extra syntactic limitations, despite the children having acquired the adult grammar?

The development of sentences involving A(rgument)-movement, such as verbal passives (9), is arguably the most widely studied instance of this type of late acquisition. One prominent hypothesis— the A-Chain Deficit Hypothesis (ACDH) (Borer & Wexler, 1987)—holds that the grammatical ability to form A-chains emerges maturationally in the child at around age six or seven. In A-movement, a DP is displaced from one argument position of a sentence to another (e.g. object to subject position in passives). English-acquiring children rarely produce the *be*-passive, in either naturalistic (e.g., Harwood, 1959) or elicited (e.g., Hayhurst, 1967) speech, although in some experimental conditions, children can be *primed* to produce verbal passives (Messenger *et al.*, 2012; Bencini & Valian, 2008).[1] English-acquiring children are also delayed in showing adult comprehension (Slobin, 1966) of both "short" passives, where the *by*-phrase is missing (9a), and "long" passives, where it is pronounced (9b) (Hirsch & Wexler, 2006, but see O'Brien, Grolla, & Lillo-Martin, 2006).[2] This non-adult comprehension is attested in many (though not all) languages, including Afro-Asiatic, Sino-Tibetan and Altaic languages as well as Indo-European ones. The most notable exception to the generalization is Sesotho, in which the verbal passive seems to be acquired early (Demuth, 1989; Demuth *et al.*, 2010, but see Crawford, 2009 for results suggesting a delay for a subset of passives in Sesotho. See also note 4).

(9) a. Ernie was seen. *Short passive*
b. Ernie was seen by Bert. *Long passive*

However, it is clearly not the case that children are delayed in acquiring all A-movement, contra the predictions of the ACDH. For example, in the active voice, subjects undergo A-movement out of the verbal domain (Koopman & Sportiche, 1991), yet children have no difficulties correctly placing the subject outside the VP (Stromswold, 1996). Very young Italian and French-speaking children have acquired various kinds of reflexive structures (impersonals and medio-passives), which involve A-movement (Snyder & Hyams, 2015). Recent experimental work shows that children have Subject-to-Object Raising (10) from as young as three years old (e.g., Kirby, 2011).

(10) Ernie wants Bert [~~Bert~~ to win the race].

Given this range of results, the acquisition of A-movement must be approached with the goal of understanding why some structures are acquired late and others early. Insight into this question is provided by the acquisition of another A-movement structure: Subject-to-Subject Raising (StSR). StSR sentences involving the verbs *seem* and *appear* are noted to be delayed in English and Dutch-acquiring children (Hirsch, Orfitelli, & Wexler, 2008).

In a series of seven experimental studies, Orfitelli (2012) compared the acquisition of StSR with *seem* (11) to those of other English StSR predicates, illustrated in (12).

(11) Ernie seems (to Bert) [to dance].
(12) Ernie is about/is going/tends (*to Bert) [to dance]

As in previous studies, children demonstrated non-adult comprehension of StSR with the verb *seem* (11) until approximately six years old. Importantly, however, these same children exhibited adult comprehension of StSR sentences with *be about*, *be going*, and *tend* (12) from as young as 4 years old. Further, a within-subjects comparison of 30 children found an over 96% correspondence between the comprehension of verbal passives and "*seem*-type" StSR (Table 26.1). These results show that children have no difficulty with the process of StSR itself; rather, it is a specific property of StSR with *seem* and *appear* that leads to their delay. Second, whatever this property is, it is shared with the verbal passive.

As with the previously discussed phenomena, accounts of A-movement acquisition can be divided into competence-based theories and performance-based theories. Although the major competence-based theories (Borer & Wexler, 1987; Babyonyshev *et al*., 2001; Hyams *et al*., 2006; Snyder & Hyams, 2015) differ in their exact implementation and predictions, they share the assumption that children's

**Table 26.1** Number of children in Orfitelli 2012b showing above chance performance for StSR sentences with *seem* and *(be) about* as compared to non-actional passives

| | *StSR* (seem) | *StSR* ((be) about) | *Short Non-Actional Passive* | *Long Non-Actional Passive* |
|---|---|---|---|---|
| 4 years (N = 10) | 0 | 10 | 0 | 0 |
| 5 years (N = 10) | 1 | 10 | 1 | 1 |
| 6 years (N = 10) | 7 | 10 | 6 | 7 |

grammar is not target-like. Here, we will focus on one grammatical account, the *Argument Intervention Hypothesis* (AIH), proposed by Orfitelli (2012):

(13) Children are delayed in acquiring exactly those structures which require A-movement across an intervening argument.

The AIH appeals to a difference between the two types of StSR predicates: *seem* and *appear* (11) optionally permit an additional experiencer argument that structurally intervenes between the base and final positions of the argument that undergoes raising (Collins, 2005a). This is in contrast to other raising predicates (12), which never permit an experiencer. Movement of the embedded subject argument to the matrix subject position therefore violates universal conditions on syntactic locality such as Relativized Minimality (Rizzi, 1990).

One of the benefits of such an analysis is that it straightforwardly predicts the developmental relationship between *seem*-type StSR and the verbal passive. Under analyses in which *by*-phrase of the English passive is initially merged in the same position as the subject of active sentences (Collins, 2005b), it intervenes between the base and final position of the promoted object.

These acquisition data are also in line with the idea that there is a "covert" *by*-phrase in the passive that is syntactically and/or semantically active even when unpronounced (Baker, Johnson, & Roberts, 1989). As noted above, children are equally delayed in comprehending verbal passives whether the *by*-phrase is explicit or implicit. Children are similarly delayed on *seem*-type StSR whether or not the experiencer is pronounced, suggesting that it is also covertly present.[3]

Beyond verbal passives and StSR, the AIH makes a clear theoretical prediction regarding the acquisition of A-movement cross-linguistically: children will be delayed in all structures that involve movement across an intervening argument, and not delayed in those which do not. Experimental data supports this conclusion. As noted, VP-internal subject movement, Romance impersonals and medio-passives, and Subject-to-Object Raising are mastered early, as predicted, because these structures do not involve movement over an intervening argument. Conversely, the inverse copula construction (14), which under Moro's (1997) analysis involves promotion of a predicational DP over an intervening subject DP, is not comprehended by English-acquiring children until six to seven years old, the age at which they begin to understand StSR and verbal passives (Hirsch & Wexler, 2007).[4]

(14) The cause of the fight is the bully.

To conclude, there appears to be a growing body of evidence in favor of the hypothesis that intervention/minimality is in some sense an inviolable constraint for young children. This is distinct from adults, who have mechanisms to circumvent it in some cases, allowing them to interpret verbal passives and *seem*-type StSR (see Collins, 2005a, for one suggestion and also Snyder and Hyams, 2015, for another). Indeed, the AIH and similar proposals could serve as a tool to adjudicate between multiple possible syntactic structures. For example, in the case of the middle voice (15), it is unclear whether the understood external argument—which cannot be overtly expressed—is syntactically present (Ackema & Schoorlemmer, 2007, and references therein). If indeed it is present, we predict a cross-linguistic delay in the acquisition of the middle voice, with similar timing as the delay in verbal passives, StSR, and inverse copulas.

(15) The toys sell easily.

Intervention effects in acquisition have been noted for A′-movement as well, although they seem to disappear around four to five years old (Friedmann *et al.*, 2009; Belletti *et al.*, 2012). We have restricted our discussion to intervention in A-movement, although a link may exist between the two phenomena.

## Control

The recursive embedding of one constituent inside another is the feature of human language that accounts for its infinite expressive power. For this reason the child's acquisition of sentential embedding represents a giant developmental leap. In the previous section we discussed the development of one kind of embedded structure, involving raising. In this section we discuss another type of complex sentence–control structures, which are superficially identical to raising structures, but with very different syntactic and semantic properties.

In general, embedded sentences can be tensed (16a), or infinitival (16b). In a tensed clause all the grammatical functions associated with the verb are expressed; for example, *Ernie* is the subject and *the piano* is the object of the embedded sentence in (16a). Most often, however, the subject of an embedded infinitive is not overtly expressed (16b). We refer to this "silent" subject as PRO.

(16) a. Bert thinks that [Ernie plays the piano].
b. $\text{Ernie}_i$ likes/wants [$\text{PRO}_i$ to play the piano].

Sentences like (16b) are referred to as control structures because the identity or reference of the embedded subject (PRO) is determined by the matrix (higher) subject (Ernie is the liker *and* the player). Because control sentences involve this silent material, they pose a potential learning challenge for the child. The problem is made apparent when we also consider the sentences in (17).

(17) a. Bert told/persuaded $\text{Ernie}_i$ [$\text{PRO}_i$ to play the piano].
b. $\text{Bert}_i$ promised/threatened Ernie [$\text{PRO}_i$ to play the piano].

In (16b) the matrix subject controls PRO; in (17a) the object, *Ernie*, is the controller (i.e., Ernie is the piano player). A possible generalization that emerges is: the structurally closest nominal controls PRO (object if there is one, otherwise the subject), but (17b) is a clear exception—*Bert* is the controller of PRO (i.e., Bert is the piano-player).

As in (16b), the embedded sentence in (17) is a complement to the matrix verb, that is, the higher verb requires a direct object and a sentential/propositional object. Control also occurs in temporal adjunct clauses, as in (18). In temporal (e.g., *before, after, while*) clauses PRO is always controlled by the matrix subject. This is assumed to be for structural reasons; the position at which the adjunct attaches to the matrix clause precludes the object from controlling PRO.

(18) $\text{Bert}_i$ hugged Ernie before/after $\text{PRO}_i$ playing the piano.

Not surprisingly, children take time to fully acquire an adult system of control. Beginning with the seminal work of C. Chomsky (1969), various experimental studies (Hsu, Cairns, & Fiengo, 1985; Cairns, McDaniel, Hsu, & Rapp, 1994; Goodluck, 1981; McDaniel, Cairns, & Hsu, 1991; Wexler, 1992, Brohier & Wexler, 1995, among others) have shown that many children ages 4 to 5½ (though not all) go through similar stages in the development of control (19):

(19) (i) Children initially take PRO to have free reference (i.e. to refer to an independent figure in the experimental setting), in both complement (e.g., 17a) and adjunct clauses (18).
(ii) Children persist in assigning free reference in adjunct clauses (18) after they have sorted out object control in complement structures (17a).
(iii) Difficulties with *promise* type verbs persist for a longer time than *like* or *tell* type verbs. As an example, for the 4 year old, the player in (17b) will generally be the object, Ernie.
(iv) Other children (or the same children at a later stage) seem to require that PRO in adverbials (18) be controlled, either by the subject of the matrix clause (adultlike) or the object (non-adultlike).
(v) These children also seem to extend the control relation to overt pronouns in sentences like (20). This has been referred to as the *adverbial coreference requirement* (McDaniel et al., 1991).

(20) $\text{Bert}_i$ washed Ernie after $\text{he}_i$ swam in the pool.

So the child who requires PRO to have Bert as controller in (18) will also interpret *he* to refer uniquely to Bert in (20). This is a possible, though not required, meaning for adults, for whom *he* can refer to any male individual given an appropriate context.

The non-adult interpretations noted in (19) have been found in various experiments using both act-out and/or judgment tasks. These results led McDaniel and colleagues to propose a stage model of control, according to which a child might construct a series of grammars (based on different attachment sites for the complement or adjunct clause) before arriving at the adult system.

Two sorts of analyses exist for why children have free reference of PRO. According to the stage model, children initially have a "flat" structure for control sentences, that is, the complement or adjunct clause is not subordinated to the matrix clause, but is more like a coordinate structure (Cairns *et al.*, 1994; Goodluck, 1981; McDaniel *et al.*, 1991). As a result PRO is not in the right structural configuration to get an adult control reading. McDaniel and Cairns (1990) suggest that children may resort to the flatter structure for processing reasons. Basically, the child treats the two "conjuncts" as independent clauses, an analysis that requires fewer computational resources because the child need not keep the matrix clause in working memory while analyzing the subordinate clause. Another prominent hypothesis is that children have not yet acquired PRO or the syntactic operation necessary for linking the event time of the adjunct to the event time of the matrix. Lacking the adult syntax they treat the embedded clause as more of a nominal, roughly as in (21) (Wexler, 1992; Carlson, 1990). In (21) anybody can be the piano player. This "nominalization hypothesis" (originally suggested by Carlson, 1990) would account for the findings in (19i-ii).

(21) Bert hugged Ernie before (the) playing (of) the piano.

Brohier and Wexler (1995) question the assumptions and empirical basis of the stage model. Based on a reexamination of previous results and their own experiment using a truth value judgment task, they find very few children with non-adult (i.e., obligatory object) control in adjuncts (18). Rather, they show that children either have adult (i.e., subject) control or they allow "free reference" of PRO, which on their analysis arises because of the nominalization strategy.

While the nominalization hypothesis is consistent with the free interpretation of PRO, there is thus far little independent evidence for this analysis (but see Goodluck, 1991 for some weak support). For example, children are unlikely to spontaneously produce nominalizations, and experimental work suggests that children between the ages of 3½ and 5 (the same ages of the children in the control studies) do not have an adult interpretation of nominals like *the kicking of him*, choosing "him" as the subject rather than the object of the action (Roeper, 1978; de Villiers *et al.*, 1995). Also, it is unclear what would force children to give up the nominalization structure in favor of an adult control analysis, given that the nominalization is consistent with any possible controller, viz. (21) is true whether Bert or Ernie or anyone else is playing the piano.

An alternative to the stage and nominalization hypotheses might relate children's non-adult performance with control to their behavior with raising structures (discussed in the previous section). Perhaps the presence of an intervening argument (*Ernie*) in sentences like (17a) and (17b) blocks the connection between PRO and the matrix subject. As a result PRO remains "unlinked," that is, free in reference (property 19ii), unless it associates with the intervener itself, that is, is object controlled. This kind of analysis, based on the argument intervention hypothesis (AIH) would also account for why *promise* type verbs are especially difficult (see also Snyder & Hyams, 2015; Mateu, 2016). In this case, the child must also skip over an intervening argument to establish the adult (subject) control relation.

Lacking the ability to do so, he associates PRO with the intervener itself.[5] The intervention explanation is particularly natural under a newer 'movement theory of control' (MTC) (Hornstein, 1999, 2003), according to which "control" is derived by movement of the embedded subject to a matrix argument position (e.g., Ernie in (22a) below, similar (though not identical) to raising (22a) below. According the MTC, control into adverbial clauses involves "sideways" movement out of the adjunct. The marked status of this operation might explain why children show a more delayed acquisition of control into adjuncts relative to complement clauses. Given the MTC and other recent advances in the syntactic analysis of control (e.g., Landau, 2000), the acquisition issues surrounding control seem ripe for further study.

## Control versus raising

As just noted, control structures are superficially similar to raising structures. The two sentence types are repeated in (22).

(22) a. Ernie seems [__ to play the piano]. *Raising*
b. Ernie likes [__ to play the piano]. *Control*

It is reasonable to ask if the surface similarity of these two sentence types poses a learning problem for the child, a question that Becker (2006, 2009) explores in detail. How does the child know if he is dealing with a raising or control structure given that the embedded subject position is phonologically empty in both cases, an even more perspicuous problem if control reduces to movement (Hornstein, 1999, 2003). In principle, expletive *it* could tell the children that they are faced with a raising verb and not a control verb, given the well-known contrast in (23) (Postal, 1978). However, as noticed by Becker, a more child-friendly difference between the two verb types concerns the animacy of the subject. Verbs like *seem* can occur with inanimate subjects, while *like*-type verbs generally require animate subjects, as shown in (24).

(23) a. It seems that Ernie plays the piano. (Ernie seems to play the piano)
b. *It likes that Ernie plays the piano. (cf. Ernie likes to play the piano)
(24) a. The rock seems to be heavy/The flower seems to open at dawn.
b. *The rock likes to be heavy/*The flower likes to open at dawn.
(cf. The dog likes to chase the ball)

Thus, probabilistic cues in the input such as (in)animacy (and others discussed by Becker) would allow the child to separate control and raising verbs despite their superficially similar structures. And given that the two structures have very different developmental profiles, it seems likely that they are indeed distinguished by children at a young age.

## General conclusions

In this chapter, we have discussed four notable stages of language acquisition. Early in their sentence production, children optionally omit both verbal inflection (RI stage) and subjects (NS stage), and there is evidence from both production and comprehension that these omissions are due to a difference between the child and adult grammars. Later in development, children whose productive language is largely adult-like nonetheless continue to show difficulties in comprehension of certain A-movement and control structures. Once again, these errors seem to reflect the child's developing grammar. In addition, each of these areas also reveals a role for extra-syntactic—processing and/or pragmatic—factors in language development.

Universal Grammar is a theory about the form of grammar the human mind can acquire under normal conditions of language exposure and cognitive growth, and in this sense it also defines the grammars a child can entertain in the course of development. And while linguistic theory does not tell us directly why the two year old opts for a NS or RI grammar, or why five year olds adhere more strictly to intervention than their parents, it does provide a framework for understanding these stages. Where children's output does not reflect their input, as in the cases discussed here, we might assume that there is some cognitive pay-off in their choice, either on a domain general or domain specific level. Perhaps they have settled (temporarily) on a system that is computationally simpler, either in terms of linguistic representation or processing, or both. The goal of a comprehensive theory of language development is to understand not only the various stages and their underlying grammars, but also to explain why the child takes a particular path to the target grammar. In large measure this will require a deeper understanding of what is "complex" for the child—whether defined by a theory of linguistic performance or by grammatical theory more directly.

## NOTES

1 Children's primed passive responses included many *get*-passives, which are known to be acquired earlier, and which are argued to be structurally distinct from 'canonical' verbal *be* passives. Children's responses also included many "reversed passives," that is, passive morphology but with the thematic mappings of the active version. In these cases children have a passive "syntactic frame" but fail to execute the movement and/or theta-role transfer operations that are argued to be the locus of children's delays.

2 For various reasons, in some languages (including English) there is an effect of verb-type. Passives with "actional" participles are acquired much earlier, around three to four years. Those with "non-actional" participles are delayed in comprehension until later (Maratsos *et al.*, 1985, among others). See Orfitelli, 2012 for a discussion of these facts and Snyder and Hyams (2015) for a possible explanation.

3 A related alternative is proposed by Choe (2013), who suggests that children's intervention difficulty is a performance effect driven by limited processing resources which prevent them from constructing a dependency over an overt experiencer (e.g., Bert in (11)). Such an analysis cannot explain children's equal delay on "short" raised sentences where the experience is unpronounced, as well as their commensurate delay on short passives where the *by*-phrase is silent.

4 Snyder and Hyams (2015) offer an explanation, also in terms of minimality, for the apparent earlier acquisition of passives in specific discourse contexts such as those given in the O'Brien *et al*., 2006 experiments and also in topic oriented languages such as Sesotho (Demuth *et al*., 2010).

5 Free reference in complement clauses such as (17a) is more difficult to explain (property 19i) via the AIH. However, this stage is attested in far fewer subjects (e.g., 2 out of 20 children in McDaniel *et al*., (1991) and 9 observations out of 45 in Cairns *et al*.'s (1994) longitudinal study. It should be noted as well that in act-out tasks children allow free reference for PRO in adjuncts but not in complement clauses (Goodluck, 1981; Hsu *et al*., 1985).

# REFERENCES

Ackema, P., & Schoorlemmer, M. (2007). Middles. In M. Everaert & H. van Riemsdijk (Eds.), *The Blackwell Companion to Syntax*. Blackwell Publishing Ltd, 131–203.

Asudeh, A., & Toivonen, I. (2012). Copy raising and perception. *Natural Language & Linguistic Theory*, 30(2), 321–380. doi: 10.1007/s11049-012-9168-2

Babyonyshev, M., Ganger, J., Pesetsky, D., & Wexler, K. (2001). The maturation of grammatical principles: evidence from Russian unaccusatives. *Linguistic Inquiry* 32, 1–44.

Baker, C. L. & McCarthy, J. (Eds.). (1981). *The logical problem of language acquisition*. Cambridge, MA: MIT Press.

Baker, C. C. L., & MacCarthy, J. J. (Eds.). (1981). *The logical problem of language acquisition*. MIT Press (MA).

Baker, M., Johnson, K., & Roberts, I. (1989). Passive arguments raised. *Linguistic Inquiry* 20, 219–251.

Becker, M. (2006). There began to be a learnability puzzle. *Linguistic Inquiry*, 37, 441–456. doi: 10.1162/ling.2006.37.3.441

Becker, M. (2009). The role of NP animacy and expletives in verb learning. *Language Acquisition*, 16, 283–296. doi: 10.1080/10489220903178997

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua*, 122(10), 1053–1069. doi: 10.1016/j.lingua.2012.02.007

Bencini, G. M. L & Valian, V. (2008). Abstract sentence representation in 3-year old: evidence from language production and comprehension. *Journal of Memory and Language*, 59, 97–113. doi: 10.1016/j.jml.2007.12.007

Blom, E. (2003). *From root infinitive to finite sentence*. Doctoral dissertation, University of Utrecht/LOT.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry* 21, 491–504.

Borer, H. (2005). *The normal course of events* (Vol. 2). Oxford University Press.

Boser, K., Lust, B., Santelmann, L., & Whitman, J. (1992). The syntax of CP and V-2 in early child German: The strong continuity hypothesis. In *Proceedings of NELS* (Vol. 23, pp. 51–65). Amherst, MA: GLSA.

Brohier, K. & Wexler, K. (1995). Children's acquisition of control in temporal adjuncts. In C. Schütze, J. Ganger, & K. Brohier (Eds.), *Papers on Language Processing and Acquisition*. MITWPL 26.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Brun, D. & Babyonyshev, M. (2006). Aspectual properties of root infinitive verbs in child Russian. *Proceedings of the 30th Annual Boston University Conference on Language Development*.

Cairns, H. McDaniel, D. Hsu, J.R., & Rapp, M. (1994). A longitudinal study of principles of control and pronominal reference in child English. *Language* 70, 2, 60–288. doi: 10.2307/415829

Carlson, G. (1990). Intuition, categories & structure. In L. Frazier & J. de Villiers (Eds.), *Language Processing and Language Acquisition* (pp. 327–333). Dordrecht: Kluwer Academic Publishers.

Chomsky, C. (1969). *The Acquisition of Syntax in Children from 5 to 10*. Cambridge, MA: MIT Press

Chomsky, N. (1965). *Aspects of the Theory of Syntax* (No. 11). MIT Press.

Collins, C. (2005a). A smuggling approach to raising in English. *Linguistic Inquiry* 36, 289–298.

Collins, C. (2005b). A smuggling approach to the passive in English. *Syntax* 8, 81–120.

Crain, S., & Fodor, J. (1993). Competence and performance in child language. In E. Dromi (Ed.), *Language and cognition: A developmental perspective* (pp. 141–171). Norwood, NJ: Ablex.

Crain, S. & McKee, C. (1985). The acquisition of structural restrictions on anaphora. S. Berman, J. -W. Choe, & J. McDonough (Eds.), *NELS 16* (pp. 94–110). Amherst: University of Massachusetts, Graduate Linguistic Student Association.

Crawford, J. (2009). Sesotho passives: The long and short of it. In *Proceedings of BUCLD* (Vol. 33, pp. 109–120).

de Villiers, J., Roeper, T., Moore, D., Johnson, K., & Bateman, S. (1995). On the acquisition of word order in nominals. Unpublished manuscript.

Demuth, K. (1989). Maturation and the acquisition of the Sesotho passive. *Language* 65, 56–80. doi: 10.2307/414842

Demuth, K., Moloi, F., & Machobane, M. (2010). 3-Year-olds' comprehension, production, and generalization of Sesotho passives. *Cognition*, 115(2), 238–251. doi: 10.1016/j.cognition. 2009.12.015

Dye, C. 2011. Reduced auxiliaries in early child language: Converging observational and experimental evidence from French. *Journal of Linguistics*, 47, 301–339. doi: 10.1017/s0022226710 00037x

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, 119(1), 67–88. doi: 10.1016/j. lingua.2008.09.002

Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language* 30, 431–451.

Golinkoff, K., & Hirsch-Pasek, R.M. (1996). *The Origins of Grammar*. MIT Press.

Goodluck, H. (2001). The nominal analysis of children's interpretations of adjunct PRO clauses. *Language*, 77(3), 494–509. doi: 10.1353/lan.2001.0159

Goodluck, H. (1981). Children's grammar of complement-subject interpretation. In Tavakolian (Ed.), *Language acquisition and Linguistic Theory*. Cambridge, MA: MIT Press, pp. 139–166.

Haegeman, L. (1995). Root infinitives and root null subjects in early Dutch. In C. Koster & F. Wijnen (Eds.), *Groningen Assembly on Language Acquisition*, 205–255. Groningen: Center for Language and Cognition.

Hamann, C., & Plunkett, K. (1998). Subjectless sentences in child Danish. *Cognition* 69, 35–72.
Harwood, F. W. (1959). Quantitative study of the speech of Australian children. *Language and Speech*, 2(4), 236–270.
Hayhurst, H. (1967). Some errors of young children in producing passive sentences. *Journal of Verbal Learning and Verbal Behavior*, 6, 634–639. doi: 10.1016/s0022-5371(67)80028-8
Hirsch, C., Orfitelli, R. & Wexler, K. (2008). The acquisition of raising reconsidered. In A. Gavarró & M. J. Freitas (Eds.), *Language Acquisition and Development: Proceedings of GALA 2007*, (pp. 252–261). Cambridge, UK: Cambridge Scholars Press.
Hirsch, C., & Wexler, K. (2006). Children's passives and their resulting interpretation. In K. ud Deen, J. Nomura, B. Schulz & B. D. Schwartz (Eds.), *The Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition North America*, vol. 4, (pp. 125–136). Honolulu, HI: University of Connecticut Occasional Papers in Linguistics.
Hirsch, C., & Wexler, K. (2007). The development of inverse copulas (and clefts). Talk presented at Western Conference on Linguistics (WECOL) 2007.
Hoekstra, T. & Hyams, N. (1998). Aspects of root Infinitives. *Lingua*, 106, 81–112. doi: 10.1016/s0024-3841(98)00030-8
Hornstein, N. (1999). Movement and control. *Linguistic Inquiry*, 30, 69–96. doi: 10.1162/002438999553968
Hornstein, N. (2003). On control. In R. Hendrick (Ed.), *Minimalist Syntax*, Oxford: Blackwell Publishing, pp. 6–81.
Hsu, J. R., Cairns, H. S., & Fiengo, R. W. (1985). The development of grammars underlying children's interpretation of complex sentences. *Cognition*, 20(1), 25–48. doi: 10.1016/0010-0277(85)90003-4
Hughes, M., & Allen, S. (2006). A discourse-pragmatic analysis of subject omission in child English. In *Proceedings of the 30th Annual Boston University Conference on Language Development* (Vol. 1, pp. 293–304).
Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.
Hyams, N. (1992). A reanalysis of null subjects in child language. In J. Weissenborn, H. Goodluck, and T. Roeper (Eds.) *Theoretical issues in language acquisition: Continuity and change in development* (pp. 249–267). Hillsdale, NJ: Lawrence Erlbaum.
Hyams, N. (2011). Missing Subjects in Early child Language. In J. De Villiers and T. Roeper (Eds.), *Handbook of Generative Approaches to Language Acquisition*, Springer, Dordrecht.
Hyams, N., Ntelitheos, D., & Manorohanta, C. (2006). Acquisition of Malagasy voicing system: Implications for the adult grammar. *Natural Language & Linguistic Theory* 24, 1049–1092.
Hyams, N., & Wexler, K. (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry* 24, 421–459.
Kirby, S. 2011. Move over, control freaks: Syntactic raising as a cognitive default. In 35th Annual Boston University Conference on Language Development 35 Proceedings Supplement.
Koopman, H., & Sportiche, D. (1991). The position of subjects. *Lingua* 85, 211–258.
Krämer, I. (1993). The licensing of subjects in early child language. In C. Phillips (Ed.) *Papers on case and agreement II*, (pp. 197–212). MIT Working Papers in Linguistics 19. Cambridge, MA: MIT, MIT Working Papers in Linguistics.
Landau, Idan. (2000). *Elements of Control*. Dordrecht: Kluwer.
Maratsos, M., Fox, D. E., Becker, J. A., & Chalkley, M. A. (1985). Semantic restrictions on children's passives. *Cognition*, 19(2), 167–191. doi: 10.1016/0010-0277(85)90017-4
Mateu, V. E. (2014). Clitic omission in Spanish-speaking children: Evaluating the roles of competence and performance. *Proceedings of the 38th Boston University Conference on Language Development*

[BUCLD 38]. Somerville, MA: Cascadilla Press.

Mateu, V. E. (2016). Intervention effects in the acquisition of raising and control: Evidence from Spanish and English. Doctoral dissertation, UCLA.

McDaniel, D. & Cairns, H. (1990). The processing and acquisition of control structures by young children. In L. Frazier & J. de Villiers (Eds.), *Language Processing and Language Acquisition* (pp. 313–325). Dordrecht: Kluwer Academic.

McDaniel D., Cairns, H., & Hsu, J. (1991). Control principles in the grammars of young children. *Language Acquisition*, 1, 297–336. doi: 10.1207/s15327817la0104_1

Messenger, K., Branigan, H. P., McLean, J. F., Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language*, 66, 568–587. doi: 10.1016/j.jml.2012.03.008

Moro, A. (1997). *The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure*. Cambridge: Cambridge University Press.

O'Brien, K., Grolla, E. & Lillo-Martin, D. (2006). Long passives are understood by young children. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *BUCLD 30: Proceedings of the 30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla.

Orfitelli, R. (2012a). 'Experiencing' a slight delay: Intervening arguments and the acquisition of subject-to-subject raising. In A. K. Biller, E. Y. Chung & A. E. Kimball (Eds.), *BUCLD 36: Proceedings of the 36th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, pp. 374–386.

Orfitelli, R. (2012b). *Argument Intervention in the Acquisition of A-movement*. Doctoral dissertation, UCLA.

Orfitelli, R., & Hyams, N. (2012). Children's grammar of null subjects: evidence from comprehension. *Linguistic Inquiry*, 43(4), 563–590. doi: 10.1162/ling_a_00106

Postal, P. (1978). *On Raising. An Inquiry Into One Rule of English Grammar and Its Theoretical Implications*. Cambridge, MA: MIT Press.

Rett, J., Hyams, N., & Winans, L. (2013). The effects of syntax on the acquisition of evidentiality. *Proceedings of the 38th Annual Boston University Conference on Language Development*.

Rett, J., & Hyams, N. (2014). The acquisition of syntactically encoded evidentiality. *Language Acquisition*, 21(2), 173–198. doi: 10.1080/10489223.2014.884572

Rizzi, L., (1994). Some notes on linguistic theory and language development: the case of root infinitives. *Language Acquisition*, 3, 371–393. doi: 10.1207/s15327817la0304_2

Rizzi, L. (1994). Early null subjects and root null subjects. In T. Hoesktra & B. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar* (pp. 151–176). Amsterdam: John Benjamins. doi: 10.1075/lald.8.09riz

Rizzi, L. (1997). The fine structure of the left periphery. In *Elements of grammar* (pp. 281–337). Springer Netherlands. doi: 10.1007/978-94-011-5420-8_7

Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, L. (2000). Remarks on early null subjects. In M-A. Friedemann and L. Rizzi (Eds.), *The Acquisition of Syntax*, (pp. 269–292). Harlow, England: Longman.

Rizzi, L. (2005). On the grammatical basis of language development: A case study. In G. Cinque & R. Kayne (Eds.), *The Oxford Handbook of Comparative Syntax*, (pp. 70–109). New York: Oxford University Press.

Roeper, T. (1978). Linguistic universals and the acquisition of gerunds. In H. Goodluck & L. Solan (Eds.), *Papers in the Structure and Development of Child Language*. UMass Occasional Papers in Linguistics, vol. 4.

Schütze, C., & Wexler, K. (1996). Subject case licensing and English root infinitives. In *Proceedings of the 20th annual Boston University conference on language development* (Vol. 2, pp. 670–681). Cascadilla Somerville, MA.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of verbal learning and verbal behavior*, 5(3), 219–227. doi: 10.1016/s0022-5371(66)80023-3

Snyder, W. & Hyams, N. (2015). Minimality effects in children's passives. In E. Di Domenico, C. Hamann, & S. Matteini (Eds.), *Structures, Strategies and Beyond*. (Linguistik Aktuell). Amsterdam/ Philadelphia: John Benjamins.

Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on production of English 3rd person –s. *Journal of Speech, Language, and Hearing Research*, 52(3), 623–642. doi: 10.1044/1092-4388(2008/07-0258)

Swinney, D. & Prather, P. (1989). On the comprehension of lexical ambiguity by young children: Investigations into the development of mental modularity. In D. S. Gorfein (Ed.), *Resolving semantic ambiguity* (pp. 225–238). New York: Springer Verlag. doi: 10.1007/978-1-4612-3596-5_12

Syrett, K., & Lidz, J. (2005). Children want to access every interpretation adults do. In L. Bateman & C. Ussery (Eds.), *NELS 35* (pp. 591–602). Amherst: University of Massachusetts, Graduate Linguistic Student Association.

Stromswold, K. (1996). Does the VP-internal subject stage really exist? Paper presented at the 21st Annual Boston University Conference on Language Development.

Suppes, P., Smith, R., & Leveillé, M. (1973). The French syntax of a child's noun phrases. *Archives de Psychologie*, 42, 207–269.

Trueswell, J. C., Sekerina, I. A., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73, 89–134. doi: 10.1016/s0010-0277(99)0 0032-3

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition* 40, 21–81.

Wexler, K. (1992). Some issues in the growth of control. In R. K. Larson, S. Iatridou, U. Lahiri, & J. Higginbotham (Eds.), *Control and Grammar* (pp. 253–295). Dordrecht: Kluwer Academic Publishers.

Wexler, K. (1994). Optional infinitives, verb movement and the economy of derivation in child grammar. In: D. Lightfoot & N. Hornstein, (Eds.), *Verb Movement*. Cambridge University Press.

Yang, C. (2002). *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

Yang, C. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Science* 8, 451–456.

# 27 Social Interaction and Language Acquisition: Toward a Neurobiological View

SARAH ROSEBERRY LYTLE AND
PATRICIA K. KUHL

University of Washington

## Introduction

Social interaction is critical for children's language acquisition (Adamson, 1995; Bloom, 2000; Bruner, 1981; Hollich *et al*., 2000; Nelson, 2007; Tomasello, 1992). From the moment they are born, children engage in social interaction (Meltzoff & Moore, 1977), and a child's language development is dependent on the social environment. As Nelson (1985) suggests, "language learning takes place within the framework of social interaction" (p. 109). In fact, children are exposed to language in social settings with parents and caregivers from the beginning (Bloom & Tinker, 2001; Clark, 2003). Children learn more than new words by engaging with social partners. As parents and children interact, parents demonstrate the rules of social interactions by engaging infants in a give-and-take format (Hirsh-Pasek & Golinkoff, 1996; Snow, 1997). These "proto-conversations" (Snow, 1997) begin to model conversational structure for the child (Clark, 2003). Eventually, children's developing language abilities allow them to become true conversational partners.

In this chapter, we review the brain and behavioral data on the effects of social interaction on language acquisition in children, discuss related work on children with autism that demonstrates dual impairments in social and linguistic processing, and relate these findings to the acquisition of communicative repertoires in non-human animals. Using this evidence, we advance the hypothesis that social interaction "gates" language learning (Kuhl, 2007; 2011). We then review candidate brain systems that could explain the existing results. Finally, the chapter discusses new approaches to the question, including neuroscience studies conducted in our laboratory, which may provide breakthrough data about the role of social factors in language acquisition.

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

## Social interaction matters: Humans vs. machines

The role of social cues in language learning is grounded in a rich theoretical literature and researchers have begun to test this claim in the laboratory. In each case, children's ability to learn language from a live social interaction is compared to learning from an equivalent non-social source, machines. These machines—videos, audio recordings, or robots—do not allow children to engage in the back-and-forth exchanges that are characteristic of social interactions with live humans and therefore provide a useful non-social medium to test the importance of social interactions.

In one study from our laboratory, Kuhl and colleagues (2003) investigated infants' ability to learn foreign language phonemes through social and non-social contexts. Given that children who are exposed to a second language early in a natural setting discriminate sounds in both of their languages (e.g., Garcia-Sierra *et al.*, 2011), the researchers asked whether monolingual children would learn from foreign language exposure under both the social and non-social conditions. Nine month olds were exposed to Mandarin Chinese in twelve 25-minute laboratory visits. Each infant experienced one of three exposure styles: a speaker on video, an audio recording of the same speaker, or a live social interaction. Four native Mandarin speakers served as the tutors. Phonemic learning was assessed in a Conditioned Head Turn procedure in which infants were trained and then tested on their ability to turn their head toward a loudspeaker when they detect a target phoneme interspersed among the background sounds, called standard sounds (Werker, Polka, & Pegg, 1997). Learning was also assessed using Event Related Potentials, or ERPs (Kuhl, 2011). Results of both the behavioral and the brain measures demonstrated that phonetic learning was not supported by video displays or by audio recordings, but that children exposed to live Mandarin speakers discriminated the foreign phonemes as well as native Mandarin speakers.

A related line of research used Spanish to replicate nine month olds' phonetic learning (Conboy & Kuhl, 2011). Infants participated in exposure sessions with live Spanish tutors who played with toys and read books to the infants for a total of five hours of exposure over four weeks. Using ERP to detect voltage fluctuations in neural activity, researchers employed a "double oddball" ERP paradigm to test Spanish phonetic learning both prior to social exposure to Spanish (at 9 months of age) and after social exposure to Spanish (at 11 months of age). In the test, the "standard" phonetic unit, an unaspirated [ta], was common to both Spanish and English, although perceived as/da/in English and/ta/in Spanish. Two "deviant" sounds, [$t^h$a] used only in English, and [da] used only in Spanish, examined the change in the brain measures for both English and Spanish as a result of exposure to Spanish. Infants' ERP responses to English revealed that their brain responses showed evidence of learning between 9 and 11 months: Infants' responses to Spanish revealed no discrimination at 9 months, and discrimination at 11 months. In other words, infants demonstrated phonetic learning for Spanish as a function of exposure.

The results of Kuhl and colleagues' social exposure studies, which show a lack of learning in the absence of a socially responsive person, appear inconsistent with studies on "statistical learning" that demonstrate phonetic (e.g., Maye *et al.*, 2008; Yoshida *et al.*, 2010; Teinonen, 2009, Bosseler *et al.*, 2016) and word (Saffran, Newport, & Aslin, 1996) learning without any social cues. However, these studies differ in the type of language material to be learned: in Kuhl and colleagues' social exposure studies, infants hear large amounts of natural complex language in a socially interactive setting, over 30,000 syllables during exposure. In contrast, the typical statistical learning experiment presents infants with a set of 8-10 syllables or 8 pseudo-words from a loudspeaker without any social cues or interaction available. One hypothesis that can be tested in future research is that natural language, with all its attendant complexity and variation, may require social interaction, at least in the early learning phases, whereas simple syllable and word learning, when isolated from natural language, does not. This is an empirical question that can be addressed and answered in future studies.

The influence of social context on language learning is not limited to phonemes. Conboy and Kuhl (2010) also used an ERP word paradigm to assess the brain's response to Spanish words infants had heard during the Spanish exposure sessions (described above) versus a set of control Spanish words they had not heard during Spanish exposure. As these children were from English-speaking households, the only Spanish exposure they had was through the exposure session. Analyses revealed the role of social interactions during social exposure to Spanish. During the exposure sessions, measures of infants' social behaviors, in this case, eye gaze shifts between the foreign language speaker and the toys that were the referent of her speech, predicted both phonetic, as well as word learning (Conboy, Brooks, Meltzoff, & Kuhl, 2015). The more adept infants were at enlisting their social skills during a social interaction, the better their language learning. Interestingly, other researchers have investigated the reverse relationship, that hearing words might influence infants' phonetic learning. Yeung and Nazzi (2014) demonstrated that after hearing object labels highlighting a stress contrast, infants showed evidence of phonemic discrimination. Together, these results link infant phonetic learning and word learning, while also highlighting the importance of social cues for language acquisition.

Another example leads to the same conclusion, in this case investigating children's ability to learn action words, or verbs, which can be difficult for young children to master (Gentner, 1982; Gentner & Boroditsky, 2001). Video clips from *Sesame Beginnings* presented two novel verbs (e.g., *blicking*) to children aged 30 to 42 months (Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009). These video clips were perceptually salient to children and the verbs were presented in full grammatical context (e.g., "Look at Dad *blicking* Elmo!"). These videos also allowed for the manipulation of social cues: Whereas some of the children saw the novel verbs presented entirely on video, others saw half of the presentations on video and half delivered by a live social partner. Children were tested on their ability to extend the novel verb to a new actor performing the same action (e.g., if children learned that a bouncing action was *blicking* from Elmo and his dad, children were tested on their ability to recognize real people *blicking*). Results indicated that toddlers who interacted with an adult while watching a video were able to learn the novel verbs at a younger age than children who passively viewed the video. Although this research found evidence of word learning from video, whereas Kuhl and colleagues (2003) did not, the results may differ due to the age of the participants—Kuhl's work was done with infants at 9 months, whereas the action word studies were done with children between 30 and 42 months of age. Together, the findings represent a developmental trajectory in children's ability to learn language from screens. As others have reported, infants do not show language learning from video (Robb, Richert, & Wartella, 2009) whereas older toddlers and preschoolers show increasing language acquisition from screens (Krcmar, Grela, & Lin, 2007; Sachs, Bard, & Johnson, 1981). In line with previous research (Krcmar *et al.*, 2007; Reiser, Tessmer, & Phelps, 1984), Roseberry and colleagues (2009) demonstrated that even though children older than three years gained some information from video alone, this learning was not as robust as learning from live social interactions.

Recent evidence suggests that the relative advantage of learning language from social interactions cannot be attributed to some drawback of the machine presentation itself. One study used video chats to ask if 24 to 30 month olds can learn language in a video context that is social (Roseberry, Hirsh-Pasek, & Golinkoff, 2014). Video chats present a speaker via screen media, yet this particular technology differs from traditional video in several important ways. Video chats approximate live social interactions in that children and the speaker can participate in a two-way exchange. Adults can be responsive to children and ask questions that are relevant to them. Although the speaker's eye gaze is often distorted in video chats because of the placement of the camera relative to the screen, video chat preserves many of the qualities of social interactivity that help children learn (Csibra, 2010). In fact, when 24 to 30 month olds were exposed to novel verbs via video chat, children learned the new words just as well as from live social interactions. Toddlers showed no evidence of learning from non-interactive video. Similarly, research with robots has discovered that a robot's social behavior influences children's ability to learn from this machine. For example, when robots oriented their heads toward 18- to 24-month-old children and named a toy in Finnish, the English-speaking children begin to follow the robot's eye gaze and learn the Finnish names for common objects (Kuhl, 2011; Meltzoff, Kuhl, Movellan, & Sejnowski, 2009; Movellan, Eckhardt, Virnes, & Rodriguez, 2009).

Taken together, empirical evidence highlights the crucial contributions of social interactions to natural language learning. Importantly, this holds across different levels of linguistic analysis (phonemes and words), across social interactions of different durations (extended over many sessions or isolated in one laboratory visit) and across specific media (video, audio, and robots). These studies are among many to offer compelling evidence that social or pragmatic cues are related to language outcomes (Bloom, Lightbown, & Hood, 1975; Childers & Tomasello, 2002, 2006; Hoff, 2006; Naigles, Hoff & Vear, 2009; Nelson, 2007; Tomasello & Farrar, 1986).

## Social interaction matters: Children with autism

In typically developing children, social interactions appear to have a positive influence on language learning. Yet, the same does not hold for children with autism. Dual impairments in both the language and social domains are characteristic of children with autism. Although there is considerable variability in the language abilities of children with autism, 42% of the population is impaired in both receptive and expressive language (Chan, Cheung, Leung, Cheung, & Cheung, 2005) and there is some evidence that specific types of language, such as relational terms or mental state verbs, are particularly affected (Tager-Flusberg, 1992). Similarly, the effects of social deficits in children with autism range from decreased abilities to orient to social stimuli (Dawson, Meltzoff, Osterling, Rinaldo, & Brown, 1998) to difficulty tracking eye gaze (Grice, Halit, Farroni, Baron-Cohen, Bolton, & Johnson, 2005) and a reduced frequency in engaging in joint attention with social partners (Dube, MacDonald, & Mansfield, 2004).

Language and social deficits in autism should not be considered independently. Increasingly, evidence suggests that the social impairments in children with autism may, in fact, influence their ability to acquire language. One tool that typically developing children use to learn language is child-directed speech (CDS). Characterized by generally slower speech and word lengthening, longer vowel sounds, and greater variation in frequency (Brand & Tapscott, 2007; Fernald & Kuhl, 1987; Garnica, 1975; Golinkoff & Alioto, 1995; Grieser & Kuhl, 1988; Papousek, Bornstein, Nuzzo, & Papousek, 1990), typically developing infants prefer this type of speech as early as two days post birth (Cooper & Aslin, 1990; Fernald, 1985). Furthermore, this type of speech exaggerates the acoustic cues that distinguish phonemes and thus words (Kuhl *et al.*, 1997). Children who are exposed to a greater degree of acoustic exaggeration in CDS exhibit greater sensitivity to phonological contrasts when tested in the laboratory (Liu, Kuhl, & Tsao, 2003), and infants who hear more parentese at home, especially in one-on-one social interactions, have higher concurrent and future language abilities; this has been shown in both monolingual (Ramirez-Esparza, Garcia-Sierra, & Kuhl, 2014) and bilingual children (Ramirez-Esparza, Garcia-Sierra, & Kuhl, in press, 2017). CDS facilitates children's ability to segment words within a stream of speech (Golinkoff & Alioto, 1995). Both phonological sensitivity and word segmentation skills are known to facilitate language learning. In fact, recent research suggests that infants who are better at segmenting streams of speech at 7 months of age have larger vocabularies at 24 months of age (Singh, Reznick, & Xuehua, 2012).

Given that CDS is an inherently social cue and that children with autism do not tend to prefer social cues, researchers have asked whether children with autism have a preference for CDS. One investigation tested this directly (Kuhl, Coffey-Corina, Padden, & Dawson, 2005). To gauge auditory preference, children were given a choice of listening to eight-sec clips of CDS or a nonspeech signal that matched the spectrum and duration of the CDS speech samples. Children with autism indicated their choice by making slight head turns to the left or right (with location of CDS randomized across children). Interestingly, typically developing children attended to both sounds. Although younger infants typically show a preference for CDS, the children in this study were matched to the children with autism on mental age. This yielded a slightly older participant group (range = 13 months to 48 months, M = 27.78), which may explain their lack of preference. In contrast to the typically developing children, children with autism showed a strong preference for the non-speech signal. The non-speech preference of children with autism was significantly correlated to the severity of their autism symptoms, as well as to their ability to discriminate phonemes when tested neurally.

That children with autism show such a strong preference for a non-social (non-speech) signal over a social signal (CDS) has many implications for the relationship between language learning and social interaction. Children with autism fail to attend to the very signal that supports language learning. Children with autism are therefore limited in their ability to benefit from CDS in the same way as typically developing children. Without CDS to highlight phonological contrasts and facilitate speech segmentation, children with autism may not have sufficient language learning tools in their repertoire.

In fact, recent data from our laboratory suggests that the level of social functioning in two year olds with autism is related to the children's brain responses during word processing (Kuhl, Coffey-Corina, Padden, Munson, Estes, & Dawson, 2013). Using an ERP task, children with autism and typically developing children listened to a series of words that included known words, unknown words, and known words presented backwards. Children with autism who were categorized as having less severe social symptoms according to the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, & Le Couteur, 1994) showed differential processing of known and unknown words that was localized to the left temporal/parietal electrode sites. This activation was similar to the typically developing matched control group. In contrast, children with more severe social symptoms showed a more diffuse response across the right hemisphere. Using the strength of individual children with ASD's responses to known words at the parietal site at the age of two years as a predictor, our results showed that these children's linguistic, cognitive, and adaptive response abilities were strongly predicted at the age of four years and six years, with correlations improving over time. In fact, regression analysis showed that this early brain measure of responses to known words in children with autism at the age of two years was more highly correlated with later linguistic, cognitive and adaptive skills than early cognitive ability (at two years) which is a frequently reported predictor of functional outcome in children with ASD (Anderson, Lord, Risi, DiLavore, Shulman, *et al.*, 2007; Munson, Faja, Meltzoff, Abbott & Dawson, 2008). Moreover, the neural measure predicted future behavior regardless of the two types of treatment that the children with autism had received in the interim. Although causal relationships cannot be established with these correlational data, the findings strengthen the argument that social interaction is strongly linked to language learning, both at the individual and group levels.

## Social interaction matters: Non-human animals

The research linking social interactions to human communication is clear. Yet social interactions also appear to be critical for species other than humans. Communicative songbirds, for example, rely on several forms of social interactions to hone their song production. In a laboratory setting, zebra finches expect that their social environment will include visual cues to song learning from their tutor (Eales, 1989). When visual cues are available, they are so powerful that zebra finches are likely to learn enemy songs, like that of a Bengalese finch, if they are fed by the alien bird (Immelmann, 1969). White-crowned sparrows are similarly reliant on social information: although live tutors effectively teach their song, sparrows are unable to learn from equivalent audio taped information (Baptista & Petrinovich, 1986).

Like human infants, songbirds recruit social information from a variety of cues. Blindfolded zebra finches, for example, are able to learn their songs through non-visual interactions such as pecking and grooming, or through contingently responsive audio of their songs (Adret, 1993). Furthermore, female cowbirds do not sing, but are able to give social feedback to young male cowbirds through their wingstrokes (West & King, 1988).

In sum, evidence from human infants, children with autism and non-human animals suggests a very powerful role for social interactions. Yet, it is unclear why social interactions facilitate language acquisition as they do. In the following section, we explore the specific social cues that have been implicated and we discuss some possible underlying mechanisms.

## Mechanisms of social interactions

The importance of social interaction emerges clearly in research that has investigated human infants, children with autism, and even non-human animals. Kuhl (2007) has gone so far as to argue that social information "gates" natural language learning suggesting that social experience is important for language learning. The more children have access to social cues and the better children's ability to use social cues, the greater their ability to learn language. This aligns with the research reviewed above on typically developing children (e.g., Kuhl *et al.*, 2003; Roseberry *et al.*, 2014) as well as children with autism (e.g., Kuhl *et al.*, 2013). How social interaction "gates" language learning is less well understood, and investigations are now directed toward answering this question.

Researchers have identified many different social cues that may contribute to the "gating" process. Eye gaze, for example, has been shown to facilitate word learning in young children. Even infants show remarkable sophistication in their use of these social cues. Children distinguish between adults' open and closed eyes, and only follow an adult's gaze when their eyes are open (Brooks & Meltzoff, 2002, 2005). This suggests that infants are sensitive to a social partner's eyes and are not merely following head turns. Children also recruit an adult's eye gaze to help them narrow the possible referents for an unknown word (Baldwin, 1993; Dunham, Dunham & Curwin, 1993). Language learning is supported by children's ability to follow gaze direction. Novel labels typically refer to the referent in the speaker's visual field (Baldwin, 1993; Bloom, 2002; Tomasello, 1995) and in fact, when the referent of a novel word is ambiguous, children are more likely to check speaker gaze to determine the correct referent than when the referent of a novel word is not ambiguous (Baldwin, Bill, & Ontai, 1996). Older infants use eye gaze to label boring objects even when they would prefer to look at other interesting objects (Pruden *et al.*, 2006).

When adults are contingently responsive to children, their responses are reliable and timely (Beebe *et al.*, 2011; Catmur, 2011), appropriate in content (Bornstein, Tamis-LeMonda, Hahn, & Ha, 2008) and matched in intensity (Gergely & Watson, 1996). These responses establish the "conversational duet" which is characterized by back-and-forth turn taking (Hirsh-Pasek *et al.*, 2015). Contingency is a powerful social cue that has recently garnered attention for its role in the social interactions that facilitate language learning. Infants are drawn to contingent responses from others very early in life. As early as four months, for example, infants prefer an adult who responds contingently to their behaviors to an adult who does not (Bigelow, MacLean, & MacDonald, 1996; Hains & Muir, 1996). Children's preference for contingent interactions extends into toddlerhood (Bloom *et al.*, 1975; Brand & Tapscott, 2007; Goldstein, King, & West, 2003) and to word learning (e.g., Tamis-LeMonda, Bornstein, Baumwell, & Damast, 1996). Moreover, as we have reviewed above, contingency may be the critical social cue that transforms machines into social vehicles for word-learning children, as is the case in video chats (Roseberry *et al.*, 2014) and social robots (Movellan *et al.*, 2009).

One of the most studied social cues is joint attention, in which both partners focus their attention on a common object or event (Adamson, Bakeman, & Deckner, 2004; Baldwin, 1991; Moll & Tomasello, 2007; Mundy, Block, Delgado, Pomares, VanHecke, & Parlade, 2007; Tomasello & Farrar, 1986). Naturalistic observations find that both children and parents talk more during episodes of joint attention (Tomasello & Farrar, 1986), and that children show increased word learning during joint attention (Adamson *et al.*, 2004; Mundy & Gomes, 1998; Smith & Ulvund, 2003; Tomasello & Todd, 1983). Baldwin (1991), for example, asked adults to label an object either when the 16- to 19-month-old infant was attending to it or when the infant was attending to another object. Infants were much more successful at mapping labels onto objects when they were already attending to the labeled object (see also Dunham, Dunham, & Curwin, 1993; Pruden *et al.*, 2006; Tomasello & Farrar, 1986).

Although each of these social cues contributes to children's language learning, there have been relatively few efforts to specify the underlying mechanisms. Kuhl (2007) hypothesizes two broad mechanisms that would help explain why social interaction could support language learning. Social interaction increases motivation and information: social interactions between adults and infants increase infant attention and heighten social arousal, and moreover, social interaction increases the amount of information children have from which to learn.

With respect to the role of motivation, social interactions increase children's attention to the communicative learning situation. A live speaker may alert children to pay attention because the information being presented is directed to them. This possibility is supported by data from the Mandarin phonemic discrimination study described earlier (Kuhl *et al.*, 2003), in which children exposed to a live Mandarin speaker were more attentive and visibly excited than children in the non-social exposure conditions. In terms of social cues, both contingency and child directed speech captures and maintains young children's attention (Fernald & Kuhl, 1987; Landry, Smith & Swank, 2006; Ratner, 1984). Infants produce more vocalizations when their parent responds contingently to them, as compared to parents who are directed to respond only on a fixed schedule (Goldstein *et al.*, 2003; Ramirez-Esparza *et al.*, 2014). Also, focusing the child on important aspects of their environment through CDS may help children learn language by directing their attention to the referent. Indeed, children who hear more CDS at 12 months of age have larger receptive vocabularies at 24 months of age (Ramirez-Esparza *et al.*, 2014).

Social interaction may also prove motivational to children through the mere presence of a social partner, as some data suggests that even minimal social connections to another person increase young children's motivation to learn (Walton *et al.*, 2012). Recent evidence from our laboratory indicates that baby peers may increase motivation, or social arousal, in the context of social interactions (Lytle, Garcia-Sierra, & Kuhl, in preparation). Nine month olds' phoneme learning was tested after exposure to a foreign language via contingent touch screen video, in which infants controlled video presentations by touching the screen. Infants were either exposed individually or in pairs. Both groups learned when videos were contingent on infants' screen touches, but infants tested in pairs showed better learning and produced greater numbers of vocalizations. In this study, the only difference between the two groups was the presence of the second baby in the paired condition. Across groups, children demonstrated equal mobility, equal screen touches to activate the video, and equal amounts of joint attention. The mere presence of the social partner appeared to motivate infants' learning.

In addition to motivational cues, social interactions may also provide children with the precise information they need to learn language. Children may gain information about the referents of novel words through social interactions with adults. Eye gaze and joint attention align with this informational hypothesis, as children look to these social cues to gather relevant information. In a sense, social cues like eye gaze and joint attention serve as spotlights for children learning words, narrowing children's focus to a small subset of possible referents (Baldwin, 1993; Dunham *et al.*, 1993).

Occasionally, the same social cues provide motivation as well as additional information about language to children. For example, contingent interactions have been shown to increase children's attention to adults (Brand & Tapscott, 2007; Goldstein *et al.*, 2003), but one recent study suggests that another form of contingency, contingent touch, may provide critical information for children's word learning. Seidl and colleagues (2015) played a stream of artificial language for four month olds, similar to the classic design of infant statistical learning studies (e.g., Aslin, Saffran, & Newport, 1998). While infants listen to this artificial language, an experimenter touched either the elbow or the knee each time a particular target "word" appeared in the language. The experimenter also touched the infant on the other body part, either the knee or the elbow, once for every grouping of non-target "words" in the language. Thus, the infant experienced equal touches to the knee and to the elbow, always contingent to words in the artificial language, though only touches to one body part could be reliably associated with a particular word. At test, infants responded differently to target words, than they did to non-target words and non-words, or rearrangements of syllables into patterns that never appeared in the language. Interestingly, there was no differentiation between test items when an experimenter touched her own elbow or knee. The authors suggest that contingent touch may be a powerful mechanism that calls attention and provides information for word-learning children.

Taken together, there is evidence that social interactions with adults promotes language learning, perhaps because it increases motivation and confers more information to the child. Yet, the behavioral evidence only tells part of the story. There is a growing body of research on the candidate brain systems that might explain the existing results as well as new approaches to these questions. In the following section, we examine these systems as well as neuroscience studies using magnetoencephalography (MEG) brain imaging conducted in our laboratory, which may provide breakthrough data.

## Candidate brain mechanisms

As reviewed above, the increase in attention and the increase in information that is provided by interaction with another human may help explain social learning effects for language. However, it is also possible that social interaction is connected to language through even more fundamental mechanisms. Social interaction may activate brain mechanisms that invoke a sense of relationship between the self and other, as well as activating social understanding systems that link perception and action (Hari & Kujala, 2009). Neuroscience research focused on shared neural systems for perception and action have a long tradition in speech research (Liberman & Mattingly, 1985). Recent interest in "mirroring systems" specifically and the "social brain" more generally have re-invigorated this tradition (Kuhl & Meltzoff, 1996; Meltzoff & Decety, 2003; Pulvermuller, 2005; Rizzolatti, 2005; Rizzolatti & Craighero, 2004).

There is tantalizing evidence from the Spanish exposure experiment suggesting that exposure to Spanish not only changes speech perception but also changes speech production. The English-learning infants who were exposed to 12 sessions of Spanish (Conboy & Kuhl, 2011) showed subsequent changes in their patterns of babbling after experience with Spanish; interestingly, babbling was language-specific after exposure (Ward, Sundara, Conboy, & Kuhl, 2009). After the 12 exposure sessions were complete infants were brought back into the laboratory for play sessions with a Spanish speaker and with an English speaker. When the children interacted with a Spanish speaker, a new pattern of infant vocalizations occurred, one that reflected the prosodic patterns of Spanish, rather than English, with longer utterance duration and more multi-syllabic utterances occurring in response to Spanish as opposed to English, consistent with the characteristics of those languages. The fact that this new pattern of vocalization only occurred in response to Spanish speech, not English, suggests that the learning that occurs in the language exposure experiments not only involves perceptual learning, but also may affect motor systems in the brain. Future language exposure experiments are needed to examine how babbling changes as a function of perceptual experience in these studies.

Thus, social exposure to language may alter both the sensory mechanisms and the motor systems underlying speech—in essence, *hearing* speech creates *motor* learning. This kind of audio-motor coupling may be activated in social settings in which we listen to others and expect to talk back reciprocally. This speculation would be enhanced by brain studies (see below) showing that listening to speech activates not only auditory sensory areas but also the motor areas underlying speech.

## A neurobiological view

Recent advances in neuroscience allow us to test the hypothesis that the pure perception of speech activates motor brain systems. Two infant studies provide intriguing data. Imada and colleagues (2006) used MEG to study newborns, 6-month-old infants, and 12-month-old infants while they listened to nonspeech, harmonics, and syllables. Dehaene-Lambertz and colleagues (2006) used fMRI to scan three-month-old infants while they listened to sentences. Both studies show activation in brain areas responsible for speech production (the inferior frontal region, Broca's area) in response to auditorily presented speech. Imada and colleagues reported synchronized activation in response to speech in auditory and motor areas at 6 and 12 months, and Dehaene and colleagues reported activation in motor speech areas in response to sentences in 3 month olds. Newborns showed no activation in motor speech areas for any signals (Imada *et al.*, 2006), whereas auditory areas responded robustly to all signals, suggesting the possibility that perception-action linkages for speech develop by three months of age as infants begin to produce vowel-like sounds.

Previous studies demonstrated activation in motor brain areas in response to speech but did not explain the role played by these areas in perceptual processing. A new study goes further in that regard. In two experiments using MEG, we investigated motor and auditory brain activation during perceptual processing of native and non-native syllables in infants at two ages that straddle the developmental transition from language-universal to language-specific speech perception (Kuhl, Ramirez, Bosseler, Lin, & Imada, 2014). MEG data revealed that seven-month-old infants activate auditory (superior temporal) as well as motor brain areas (Broca's area, cerebellum) in response to speech, and equivalently for native and non-native syllables. However, in 11- and 12-month-old infants, *native* speech activated auditory brain areas to a greater degree than non-native, while *non-native* speech activated motor brain areas to a greater degree than native speech. This double dissociation in 12-month-old infants matched the pattern of results obtained in adult listeners. The data were interpreted in the context of two historical theories from the 1950s and 1960s that dealt with the nature of the interaction between perceptual and motor representations of speech, *The Motor Theory* (MT) (Liberman *et al.*, 1967) and *Analysis by Synthesis* (AxS) (Stevens & Halle, 1967), a framework derived from artificial intelligence. Both MT and AxS hold that speech perception involves access to motor representations of speech in adults, but differ with regard to the role of development. A tenet of MT is that knowledge of speech production is innate (Liberman & Mattingly, 1985), whereas AxS holds that perception involves access to stored representations that result from the learning of motor patterns—analysis of incoming speech uses a kind of synthesis of the motor patterns of speech to assist perception.

Kuhl *et al.*'s experiment posed the question using MEG technology in infants for the first time. At seven months of age, infants activate both auditory and motor brain areas equally to both native and non-native syllables, and 11 month olds show greater activation in motor areas to non-native syllables. These results were interpreted in the framework of *Analysis by Synthesis* arguing that infants coupled auditory *analysis* of speech with approximations of the motor plans necessary to produce the speech signal gleaned from their own nascent abilities to produce speech (a form of *synthesis*). In other words, Kuhl *et al.* argued that infants are engaged in a kind of crude motor rehearsal of the patterns needed for speech well before they can articulate the sounds they are listening to. This form of motor activation is not seen for non-speech (Imada *et al.*, 2006). We suggest that this brain activation may underpin infants' differential babbling to English and Spanish after exposure to Spanish (Ward *et al.*, 2009), and also that motor brain activation plays a role in the developmental transition in infant perception (Kuhl *et al.*, 2014). Further support for connections between sensory and motor interactions in speech perception derive from a recent study reporting that infants' perception of speech is disrupted by a prosthetic device inserted in the baby's mouth—a device disrupting the lips, for example, impeded discrimination of bilabial sounds (ba-pa) but not dental (da-ta) sounds (Bruderer *et al.*, 2015). Taken together, these studies prompt us to revisit the original theories of speech perception, which argued that deep connections exist between sensory and motor representations of speech.

Action-perception linkages early in development could play a role in supporting social reciprocity in humans. The fact that the infant's motor brain systems respond to the speech actions of others is a step toward social communication. What the existing data do not reveal is whether joint activation of the perception-action systems is evoked especially when language is presented socially, and not when language is presented through a disembodied source such as a television set. In the recent Kuhl *et al.* (2014) study, infants heard synthetic adult-directed speech, not motherese in face-to-face interchanges, and yet motor activation was observed in the brain. We expect that if infants viewed a social stimulus (e.g., a human face speaking parentese) motor brain activation would be enhanced. Experiments are currently underway with infants using MEG in our laboratory to address this question. These tests may provide tangible evidence that speech occurring in a social face-to-face setting is especially effective in activating motor brain systems.

## Conclusions

In both animals and humans, the idea that a social context is critical to communicative learning is gaining traction. In the case of humans, language learning has been suggested as grounded in a rich social setting. Researchers testing this claim in the laboratory have provided ample evidence showing that social contexts provide both *motivation* in the form of increased attention and social arousal, as well as *information*, such as eye-gaze following, that provides added information about speakers' intentions and goals. These features of social contexts are not present to the same degree in non-social contexts. New depth regarding theoretical explanations of the effects of social contexts on communicative learning are expected from studies now underway using the tools of modern neuroscience. These studies directly investigate how communicative signals and social settings alter the brain's responses. It is hoped that these studies will eventually provide a full neurobiological account that explains how and why human language learning is fundamentally imbedded in social interaction.

## Acknowledgments

The work described here was supported by a grant to the University of Washington LIFE Center from the NSF Science of Learning Center Program (SMA-0835854: PI, P. Kuhl), and from The Ready Mind Project.

## REFERENCES

Adamson, L. (1995). *Communication Development During Infancy* (pp. 86–120). Madison, WI: Brown and Benchmark.

Adamson, L. B., Bakeman, R., & Deckner, D. F. (2004). The development of symbol-infused joint engagement. *Child Development*, 75, 1171–1187. doi: 10.1111/j.1467-8624.2004.00732.x

Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour*, 46, 149–159. doi: 10.1006/anbe.1993.1170

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., *et al*. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*, 75, 594–604.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9, 321–324. doi: 10.1111/1467-9280.00063

Baldwin, D. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, 62, 875–890. doi: 10.1111/j.1467-8624.1991.tb01577.x

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20, 395–418. doi: 10.1017/S0305000900008345

Baldwin, D. A., Bill, B., & Ontai, L. L. (1996, May). *Infants' tendency to monitor others' gaze: Is it rooted in intentional understanding or a result of simple orienting?* Paper presented at the International Conference on Infant Studies, Providence, RI.

Baptista, L. F. & Petrinovich, L. (1986). Song development in the white-crowned sparrow: Social factors and sex differences. *Animal Behaviour*, 34, 1359–1371. doi: 10.1016/S0003-3472(86)80207-X

Beebe, B., Steele, M., Jaffe, J., Buck, K. A., Chen, H., Cohen, P., *et al*. (2011). Maternal anxiety symptoms and mother-infant self- and interactive contingency. *Infant Mental Health Journal*, 32, 174–206. doi: 10.1002/imhj.20274

Bigelow, A. E., MacLean, B. K., & MacDonald, D. (1996). Infants' response to live and replay interactions with self and mother. *Merrill-Palmer Quarterly*, 42, 596–611.

Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MA: The MIT Press.

Bloom, P. (2002). Mindreading, communication and the learning of names for things. *Mind & Language*, 17, 37–54. doi: 10.1111/1468-0017.00188

Bloom, L., Lightbown, P., & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development*, 40 (Serial No. 160). doi: 10.2307/1165986

Bloom, L. & Tinker, E. (2001). The intentionality model and language acquisition: Engagement, effort, and the essential tension in development. *Monographs of the Society for Research in Child Development*, 66(4). doi: 10.1111/1540-5834.00163

Bosseler, A. N., Teinonen, T., Tervaniemi, M., & Huotilainen, M. (2016). Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS ONE*, 11:e0162177. doi: 10.1371/journal.pone.0162177.

Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C., & Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, 44, 867–874. doi: 10.1037/0012-1649.44.3.867

Brand, R. J. & Tapscott, S. (2007). Acoustic packaging of action sequences by infants. *Infancy*, 11, 321–332. doi: 10.1080/15250000701310413

Brooks, R. & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, 38, 958–966. doi: 10.1037/0012-1649.38.6.958

Brooks, R. & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8, 535–543. doi: 10.1111/j.1467-7687.2005.00445.x

Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences, 44*, 13531–13536. doi: 10.1073/pnas.150863112

Bruner, J. S. (1981). The social context of language acquisition. *Language and Communication*, 1, 155–178.

Catmur, C. (2011). Contingency is crucial for creating imitative responses. *Frontiers in Human Neuroscience*, 5, 1–2. doi: 10.3389/fnhum.2011.00015

Chan, A. S., Cheung, J., Leung, W. M., Cheung, R., & Cheung, M. (2005). Verbal expression and comprehension deficits in young children with autism. *Focus on Autism & Other Developmental Disabilities*, 20, 117–124. doi: 10.1177/10883576050200020201

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, 38, 967–978. doi: 10.1037/0012-1649.38.6.967

Childers, J. B., & Tomasello, M. (2006). Are nouns easier to learn than verbs? Three experimental studies. In K. Hirsh-Pasek & R. Golinkoff (Eds.), Action meets word: How children learn verbs (pp. 311–335). Oxford: Oxford University Press.

Clark, E. V. (2003). *First language acquisition*. New York: Cambridge University Press.

Conboy, B. T., Brooks, R., Meltzoff, A. N., & Kuhl, P. K. (2015). Social interaction in infants' learning of second-language phonetics: An exploration of brain-behavior relations. *Developmental Neuropsychology*, 40, 216–219.

Conboy, B. T., & Kuhl, P. K. (2010). Brain responses to words in 11-month-old infants after exposure to a second language. Poster presented at the American Speech-Language-Hearing Association Conference, Philadelphia, PA.

Conboy, B. T., & Kuhl, P. K. (2011). Impact of second-language experience in infancy: Brain measures of first- and second-language speech perception. *Developmental Science*, 14, 242–248. doi: 10.1111/j.1467-7687.2010.00973.x

Cooper, R. P. & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61, 1584–1595. doi: 10.1111/j.1467-8624.1990.tb02885.x

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, 25, 141–168. doi: 10.1111/j.1468-0017.2009.01384.x

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldo, J. & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, 28, 479–485. doi: 10.1023/A:1026043926488

Dehaene-Lambertz, G., *et al*. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Sciences, 103*, 14240–14245. doi: 10.1073/pnas.0606302103

Dube, W. V., MacDonald, R. P. F., & Mansfield, R. C. (2004). Toward a behavioral analysis of joint attention. *Behavior Analyst*, 27, 197–207.

Dunham, P. J. Dunham, F., & Curwin, A. (1993). Joint-attentional states and lexical acquisition at 18 months. *Developmental Psychology*, 29, 827–831. doi: 10.1037/0012-1649.29.5.827

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Animal Behaviour*, 37, 507–508. doi: 10.1016/0003-3472(89)90097-3

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181–195. doi: 10.1016/S0163-6383(85)80005-9

Fernald, A. & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279–293. doi: 10.1016/0163-6383(87)90017-8

Garcia-Sierra, A., Rivera-Gaxiola, M., Conboy, B. T., Romo, H., Percaccio, C. R., Klarman, L., Ortiz, S., & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics*, 39, 546–557. doi: 10.1016/j.wocn.2011.07.002

Garnica, O. K. (1975). How children learn to talk. *Theory Into Practice*, 14, 299–305. doi: 10.1080/00405847509542591
Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Language, thought, and culture, Volume 2* (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum Associates.
Gentner, D. & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S.C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). New York, NY: Cambridge University Press.
Gergely, G. & Watson, J. S. (1996). The social biofeedback theory of parental affect-mirroring: The development of emotional self-awareness and self-control in infancy. *The International Journal of Psychoanalysis*, 77, 1181–1212.
Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, 100*, 8030–8035. doi: 10.1073/pnas.1332441100
Golinkoff, R. M. & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language*, 22, 703–726. doi: 10.1017/S0305000900010011
Grice, S. J., Halit, H., Farroni, T., Baron-Cohen, S., Bolton, P., & Johnson, M. H. (2005). Neural correlates of eye-gaze detection in young children with autism. *Cortex*, 41, 342–353. doi: 10.1016/S0010-9452(08)70271-5
Grieser & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, 24, 14–20. doi: 10.1037/0012-1649.24.1.14
Hains, S. M. J. & Muir, D. W. (1996). Infant sensitivity to adult eye direction. *Child Development*, 67, 1940–1951. doi: 10.1111/j.1467-8624.1996.tb01836.x
Hari, R. & Kujala, M. V. (2009). Brain basis of human social interaction: From concepts to brain imaging. *Physiological Review*, 89, 453–479. doi: 10.1152/physrev.00041.2007
Hirsh-Pasek, K., Adamson, L.B., Bakeman, R., Owen, M.T., Golinkoff, R.M., & Pace, A., *et al*. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, 26, 1071–1083. doi: 10.1177/0956797615581493
Hirsh-Pasek, K. & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.
Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26, 55–88. doi: 10.1016/j.dr.2005.11.002
Hollich, G. J., Hirsh-Pasek, K., & Golinkoff, R. M. With Hennon, E., Chung, H. L., Rocroi, C., Brand, R. J., & Brown, E. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65, (3, Serial No. 262). doi: 10.1111/1540-5834.00090
Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., & Kuhl, P. K. (2006). Infant speech perception activates Broca's Area: A developmental MEG study. *NeuroReport*, 17, 957–962. doi: 10.1097/01.wnr.0000223387.51704.89
Immelmann, K. (1969). On the effect of early experience upon sexual object fixation in estrildine finches. *Zeitschrift fur Tierpsychologie*, 26, 677–691.
Krcmar, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychology*, 10, 41–63. doi: 10.1080/15213260701300931
Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, 10, 110–120. doi: 10.1111/j.1467-7687.2007.00572.x

Kuhl, P. K. (2011). Early language learning and literacy: Neuroscience implications for education. *Mind, Brain, and Education*, 5, 128–142. doi: 10.1111/j.1751-228X.2011.01121.x

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., & Ryskina, V. L. (1997). Cross-linguistic analysis of phonetic units in language addressed to infants. *Science*, 277, 684–686. doi: 10.1126/science.277.5326.684

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science*, 8, 1–12. doi: 10.1111/j.1467-7687.2004.00384.x

Kuhl, P. K., Coffey-Corina, S., Padden, D., Munson, J., Estes, A., & Dawson, G. (2013). Brain responses to words in 2-year-olds with autism predict developmental outcomes at age 6. *PLoS ONE*, 8, e64967. doi: 10.1371/journal.pone.0064967

Kuhl, P. K. & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, 100, 2425–2438. doi: 10.1121/1.417951

Kuhl, P. K., Ramirez, R., Bosseler, A., Lin, J. -F. L., & Imada, T. (2014). Infants' brain responses to speech suggest Analysis by Synthesis. *Proceedings of the National Academe of Sciences, USA.* 111, 11238–11245. doi: 10.1073/pnas.1410963111

Kuhl, P. K., Tsao, F., & Liu, H. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *PNAS*, 100, 9096–9101. doi: 10.1073/pnas.1532872100

Landry, S. H., Smith, K. E., & Swank, P. R. (2006). Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills. *Developmental Psychology*, 42, 627–642. doi: 10.1037/0012-1649.42.4.627

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol Rev*, 74, 431–461.

Liberman, A. M. & Mattingly, I. G. (1985). The motor theory of speech perception revisited. *Cognition*, 21, 1–36. doi: 10.1016/0010-0277(85)90021-6

Liu, H., Kuhl, P. K., & Tsao, F. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, 6, 1–10. doi: 10.1111/1467-7687.00275

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule – WPS (ADOS-WPS)*. Los Angeles, CA: Western Psychological Services.

Lytle, S. R., Garcia-Sierra, A., & Kuhl, P. K. (in preparation). Two are better than one: Infant phoneme learning in the presence of peers.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.

Meltzoff, A. N. & Moore, K. M. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75–78. doi: 10.1126/science.198.4312.75

Meltzoff, A., Kuhl, P., Movellan, J. R., & Sejnowski, T. (2009). Foundations for a new science of learning. *Science*, 235, 284–288. doi: 10.1126/science.1175626

Meltzoff, A. N. & Decety, J. (2003). What imitation tells us about social cognition: A rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society: Biological Sciences*, 358, 491–500. doi: 10.1098/rstb.2002.1261

Moll, H. & Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, 43, 309–317. doi: 10.1037/0012-1649.43.2.309

Movellan, J. R., Eckhardt, M., Virnes, M., & Rodriguez, A. (2009). Sociable robot improves toddler vocabulary skills. *Proceedings of the International Conference* *on Human Robot Interaction*, HRI2009. doi: 10.1145/1514095.1514189

Mundy, P., Block, J., Delgado, C., Pomares, Y., VanHecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, 78, 938–954. doi: 10.1111/j.1467-8624.2007.01042.x

Mundy, P. & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, 21, 469–482. doi: 10.1016/S0163-6383(98)90020-0

Munson, J., Faja, S., Meltzoff, A., Abbott, R., & Dawson, G. (2008) Neurocognitive predictors of social and communicative developmental trajectories in preschoolers with autism spectrum disorders. *Journal of International Neuropsychological Society*, 14, 956–966.

Naigles, L., Hoff, E., & Vear, D. (2009). Flexibility in early verb use: Evidence from a multiple-n diary study. *Monographs of the Society for Research in Child Development*, 74 (2). doi: 10.1111/j.1540-5834.2009.00520.x

Nelson, K. (1985). *Making Sense: The Acquisition of Shared Meaning*. New York: Academic Press.

Nelson, K. (2007). *Young Minds in Social Worlds: Experience, Meaning, and Memory*. Cambridge, MA: Harvard University Press.

Papousek, M., Bornstein, M. H., Nuzzo, C., & Papousek, H. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior & Development*, 13, 539–545. doi: 10.1016/0163-6383(90)90022-Z

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R., & Hennon, E. (2006). The birth of words: Ten-month-olds learn words though perceptual salience. *Child Development*, 77, 266–280. doi: 10.1111/j.1467-8624.2006.00869.x

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6, 576–582. doi: 10.1038/nrn1706

Ramirez-Esparza, N., Garcia-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, 17, 880–891. doi: 10.1111/desc.12172

Ramírez-Esparza, N., Garcia-Sierra A., & Kuhl, K. P. (in press, 2017). The impact of early social interactions on later language development in Spanish-English bilingual infants. *Child Development*. doi: 10.1111/cdev.12648

Ratner, N. B. (1984). Patterns of vowel modification in mother-child speech. *Journal of Child Language*, 11, 557–578. doi: 10.1017/S030500090000595X

Reiser, R. A., Tessmer, M. A., & Phelps, P. C. (1984). Adult-child interaction in children's learning from "Sesame Street." *Educational Communication & Technology Journal*, 32, 217–223.

Rizzolatti, G. (2005). The mirror neuron system and imitation. In S. Hurley & N. Chater (Eds.), *From Neuroscience to Social Science: Vol. 1: Mechanisms of Imitation and Imitation in Animals*. Cambridge, MA: MIT Press.

Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Robb, M. B., Richert, R. A., & Wartella, E. (2009). Just a talking book? Children's word learning from watching baby videos. *British Journal of Developmental Psychology*, 27, 27–45. doi: 10.1348/026151008X320156

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, 85, 956–970. doi: 10.1111/cdev.12166

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development*, 80, 1360–1375. doi: 10.1111/j.1467-8624.2009.01338.x

Sachs, J., Bard, B., & Johnson, M. L. (1981). Language learning with restricted input: Case studies of two hearing children with deaf parents. *Applied*

*Psycholinguistics*, 2, 33–54. doi: 10.1017/S0142716400000643

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52. doi: 10.1016/S0010-0277(98)00075-4

Seidl, A., Tincoff, R., Baker, C., & Cristia, A. (2014). Why the body comes first: effects of experimenter touch in infants' word learning. *Developmental Science*, 18, 1–10. doi: 10.1111/desc.12182

Singh, L., Reznick, J. S., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science*, 15, 482–495.

Stevens, K. N., Halle, M. (1967). Models for the Perception of Speech and Visual Form: Proceedings of a Symposium., W. Waltham-Dunn (Ed.) (pp. 88–102). Cambridge, MA: MIT Press.

Smith, L. & Ulvund, S. E. (2003). The role of joint attention in later development among preterm children: Linkages between early and middle childhood. *Social Development*, 12, 222–234. doi: 10.1111/1467-9507.00230

Snow, C. E. (1997). Mothers' speech research: From input to interaction. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to Children: Language Input and Acquisition*. Cambridge: Cambridge University Press.

Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: Deficits in the early acquisition of a theory of mind. *Child Development*, 63, 161–172. doi: 10.1111/j.1467-8624.1992.tb03604.x

Tamis-LeMonda, C. S., Bornstein, M. H., Baumwell, L., & Damast, A. M. (1996). Responsive parenting in the second year: Specific influences on children's language and play. *Early Development & Parenting*, 5, 173–183. doi: 10.1002/(SICI)1099-0917(199612)5:4<173::AID-EDP131>3.0.CO;2-V

Teinonen T., Fellman V., Näätäänen R., Alku P., & Huotilainen M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, 10, 21. doi: 10.1186/1471-2202-10-21

Tomasello, M. (1992). *First Verbs: A Case Study of Early Lexical Development*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511527678

Tomasello, M. (1995). Pragmatic contexts for early verb learning. In M. Tomasello & W. E. Merriman (Eds.), *Beyond the Names For Things: Young Children's Acquisition of Verbs* (pp. 115–146). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tomasello, M., & Farrar, J. (1986). Joint attention and early language. *Child Development*, 57, 1454–1463. doi: 10.2307/1130423

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, 4, 197–212. doi: 10.1177/014272378300401202

Walton, G. M., Cohen, G. L., Cwir, D., & Spencer, S. J. (2012). Mere belonging: The power of social connections. *Journal of Personality and Social Psychology*, 102, 513–532. doi: 10.1037/a0025731

Ward, N., Sundara, M., Conboy, B., & Kuhl, P.K. (2009). Consequences of short-term language exposure in infancy on babbling. Poster presented at the Acoustical Society of America, San Antonio, TX.

Werker, J. F., Polka, L., & Pegg, J. E. (1997). The conditioned head turn procedure as a method for testing infant speech perception. *Early Development and Parenting*, 6, 171–178. doi: 10.1002/(SICI)1099-0917(199709/12)6:3/4<171::AID-EDP156>3.0.CO;2-H

West, M. J. & King, A. P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature*, 334, 244–246.

Yeung, H. H. & Nazzi, T. (2014). Object labeling influences infant phonemic learning and generalization. *Cognition*, 132, 151–163.

Yoshida, K. A., Pons, F., Maye, J., & Werker, J.F. (2010). Distributional phonetic learning at 10 months of age. *Infancy* 15, 420–433. doi: 10.1111/j.1532-7078.2009.00024.x

# FURTHER READING

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. M., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, 16, 3–34. doi: 10.1177/1529100615569721

This paper provides an excellent theoretical structure for how to think about children's media, specifically apps. The framework proposed by this paper can be used to guide decision-making about which types of screen media might be a useful tool for children's learning.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26, 55–88. doi: 10.1016/j.dr.2005.11.002

This is a fantastic review about how children learn language in their environments and the social cues in those environments that they recruit for language acquisition.

Kuhl, P. K. & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual Reviews of Neuroscience*, 31, 511–534.

This paper provides an excellent review of the neuroscience techniques used to study infant language learning and provides a thorough discussion of the questions researchers can ask with such techniques.

Institute for Learning & Brain Sciences (producer). (2014) Online Training Modules [online modules]. Available at: http://modules.ilabs.uw.edu/outreach-modules/

These free online resources offer the latest science of child development, presented in a way that is relevant, accessible, and useable. Each module takes 20-25 minutes to complete, and discusses a specific topic in child development.

# 28 Bilingual Acquisition: A Morphosyntactic Perspective on Simultaneous and Early Successive Language Development

JÜRGEN M. MEISEL

University of Hamburg and University of Calgary

## Language acquisition types: Monolingual and bilingual

It is a truth widely acknowledged that children exposed to a language from birth—more precisely prenatally—develop a full grammatical competence in this language. Exposure to the primary linguistic data in meaningful communicative interactions with native speakers is the only necessary requirement for this to happen. The fact that all children, except for pathological cases, succeed, in spite of particularities in acquisitional settings, individual properties like intelligence or personality, and different social backgrounds, can be accounted for by assuming that grammatical development is guided and constrained by the Language Making Capacity common to all humans. Yet is this also true for children acquiring more than one language, simultaneously from birth or successively in early childhood?

The goal of the present chapter is to explore this issue and to determine whether bilinguals represent a "special population," as compared to monolingual first language (L1) learners. This discussion is concerned with children acquiring two (or more) languages, either simultaneously or successively before the age of five years, focusing on the development of syntax and morphology; for discussions of phonetic and lexical development in young bilinguals see Chapter 25 (Byers-Heinlein & LewWilliams) of this volume. Note that comparing bilinguals to monolinguals does not imply that monolingualism represents the accepted standard that bilinguals must meet. Rather, by contrasting these acquisition types, we hope to discover particularities of each of them and to gain insights into the nature of

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

the Language Making Capacity, its possibilities and its limits. Note further that this discussion does not extend to adult second language (L2) acquisition, that is, successive acquisition starting at of age of onset 11 years or later. It does, however, embrace the questions of whether early successive acquisition exhibits L2 properties or whether it resembles L1 development. This comparison is motivated by the assumption that L2 and L1 grammatical development differ substantially; see Meisel (2011) for an in-depth treatment of this issue or chapter 29 (Martohardjono & Klein) of this volume for a different approach.

## Simultaneous bilingualism: Bilingual first language acquisition

Systematic investigations of child bilingualism based on carefully collected empirical data began 100 years ago with Ronjat (1913). Even this early study discussed the relationship between bilinguals' two languages. In fact, this is undoubtedly the most frequently studied issue in research on child bilingualism. The main controversy concerns the question whether the two languages can develop separately or whether they are necessarily merged—at least temporarily— into a single system. Leopold (1939–1949), for example, adopted the latter view in a data-rich study of his daughter's acquisition of English and German. Ronjat (1913), on the other hand, analyzing the development of French and German in his son, arrived at conclusions similar to the Dual System Hypothesis according to which the two grammars develop independently.

This hypothesis is widely accepted today, due largely to a number of case studies carried out in the late 1980s and 1990s; see earlier state-of-the-art summaries of research on bilingual acquisition by De Houwer (1995, 2005) or Meisel (2001, 2004). In fact, these findings as well as subsequent research have led to the conclusion that simultaneous acquisition can be qualified as development of two first languages (2L1), meaning that it does not exhibit qualitative differences as compared to monolingual L1 development. This amounts to saying that 2L1 exhibits the same characteristics as L1 during early developmental phases (differentiation of grammatical systems) and in the subsequent course of development (developmental sequences), and that bilinguals attain grammatical competences not distinct from those of monolinguals (ultimate attainment). In what follows, I will address each of these three claims in more detail.

In order to avoid misunderstandings, let me emphasize that qualifying 2L1 as an instance of dual first language acquisition refers to the acquired grammatical knowledge. This is not to say that bilinguals are in every respect "two monolinguals." Rather, as Grosjean (1989, 2008) pointed out, bilinguals rarely use their languages equally frequently in every domain of their social environment. Hence, their abilities and skills in using each of these languages reflect their communicative and social preferences and needs. These include the ability to switch between languages in interactions with other bilinguals, obviously not an option for monolinguals. However, the stipulation that "it may well be that [the bilingual's] competencies are in some ways different from those of the two corresponding monolinguals" (Grosjean, 2008, p. 15) requires empirical support that can only be obtained by contrasting bilingual and monolingual grammars.

The empirical findings on which this chapter reports consist mainly of longitudinal case studies that contrast the development of the two languages of bilingual individuals either with each other or with grammatical developments in the respective monolinguals. Proceeding in this fashion reveals whether the two languages follow distinct acquisition trajectories in cases where functionally equivalent linguistic expressions are encoded differently in the target grammars. At the same time, it allows us to decide whether the distinct acquisition paths are the same as the ones followed by monolingual children. This means that bilinguals are compared to monolingual peers rather than to a monolingual system using the yardstick of a standard norm, thus avoiding what Putnam and Sánchez (2013, p. 479) criticized as a "dubious practice."

## *Differentiation of grammars*

The first and perhaps most crucial task that the bilingual child faces is to initiate the development of two distinct mental grammars rather than to incorporate all the acquired knowledge into a single system. From a monolingual perspective, this must seem like a formidable challenge. It is therefore not surprising that grammatical differentiation was the most intensely debated issue for about 25 years, starting in the mid-1970s when linguistic and psycholinguistic research developed an increasing interest in bilingualism. Although empirical support for the Dual System Hypothesis was available from early on, see, for example, Bergman (1976), the Unitary Language System Hypothesis (Genesee, 1989) became the dominant view in psycholinguistics for some time. Its most influential version, the three-stage model proposed by Volterra and Taeschner (1978) and Taeschner (1983), claimed that bilingual acquisition is characterized by an initial period during which children develop only one system before they succeed in differentiating first the lexical and subsequently the grammatical systems of their languages. Hence, differentiation is normally achieved, but only after a phase during which the child applies "the same syntactic rules to both languages" (Volterra & Taeschner, 1978, p. 311).

However, empirically based studies analyzing a variety of linguistic phenomena in the speech of children acquiring different language pairs revealed that bilinguals do not have to proceed through such a unitary system phase; see De Houwer (1990), Genesee (1989), Meisel (1989). In fact, analyses of various longitudinal corpora demonstrated that early differentiation is not only possible but characteristic of the developmental pattern through which simultaneous bilinguals typically proceed. This is evidenced by De Houwer's (2005) review of research of simultaneous language acquisition published between 1985 and 2002. She reported that these studies investigated the linguistic development of a total of 29 children, acquiring 12 languages in 13 different combinations, and she concluded that the results of these studies support the Dual System Hypothesis. The significance of this finding derives from the fact that the reviewed publications analyzed 13 different morphosyntactic phenomena, including noun and verb morphology, markings of tense and aspect, pronouns and clitics, negation, and word order in main and subordinate clauses. In fact, the claim that bilingual children are able to differentiate grammatical systems from early on received ample support by studies published during the following 12 years, analyzing these and other grammatical domains, based on spontaneous as well as experimental data. The Dual System Hypothesis thus rests on a large database covering a variety of language combinations, including non-Indo-European languages like, for example, Basque, Hebrew, Inuktitut, Korean, Mandarin, or Quechua.

Most importantly, this research established beyond any reasonable doubt that morphosyntactic systems are differentiated during the earliest phases of productive use of particular constructions. In fact, differentiation may happen even earlier, but for obvious reasons analyses of production data can only provide empirical evidence for hypotheses once the relevant phenomena are attested in children's speech. To mention one example, distinct word order patterns in the two languages of bilinguals can only be detected once children use multi-word utterances. Empirical evidence shows that target word orders are indeed used as soon as multi-word utterances appear in the speech of bilinguals. Children acquiring an OV language like Basque or German and a VO language like Spanish or French, use OV order from early on in the OV but not in the VO language (see Mahlau, 1994, Meisel, 1986). Moreover, when acquiring V2 languages that require finite verbs to be placed in structural second position, this effect is attested in early utterances of the V2 but not in the non-V2 language. French-German bilinguals, for example, place German finite verbs in pre-subject clause-second position if an adverb or an object appears in clause-initial position, as required by the target V2 grammar. In French, however, the verb appears after the subject, in third position, in accordance with the non-V2 grammar; cf. Meisel (1989). This can be illustrated by examples from a French-German child, Pascal, uttered at age 2;2 (years, months) and 2;4: *da in tasche musst du das* "there in (the) bag must you (put) that," *un petit peu ça pique* "a little bit it's itchy." Importantly, different word orders of this sort begin to emerge when children's MLUs (mean length of utterance) attain values of approximately 1;75–2;0, typically at around age 2;0. During the same age period, bilingual children begin to productively use verb inflection in both languages, starting with subject-verb agreement (Meisel, 1994), even in languages like Basque, which also mark object-verb agreement, cf. Ezeizabarrena (1994). In sum, empirical evidence for grammatical differentiation in morphosyntax is available as of approximately age 2;0, that is, as soon as the relevant phenomena are present in child language.

However, there exist good reasons to assume that grammatical differentiation actually sets in earlier, for differentiation of phonological systems precedes the separation of syntactic systems. It is therefore plausible to assume that language differentiation is initiated and enhanced by prosodic bootstrapping. Research on phonological processing demonstrated that newborns can distinguish between languages exhibiting different prosodic or rhythmic properties. This suggests that prosodic information facilitates the discovery of two distinct systems in bilingual settings, certainly with respect to the phonological systems, but possibly also in the morphosyntactic domain. Sebastián-Gallés and Bosch (2005, p. 69) report that, even when exposed to languages of the same rhythmic group, for example, Catalan and Spanish, infants, as early as four to five months of age, were able to separate the two languages. They conclude that "attention to specific prosodic and distributional cues of syllabic or segmental units in the speech signal may help the infant reach an early differentiation between the languages" (Sebastián-Gallés & Bosch, 2005, p. 71). In view of these discriminating capacities, which bilinguals display even in the prelexical phase of linguistic development, it is much less surprising to find that they are able to distinguish the lexical (cf. Quay, 1995) and grammatical systems of the ambient languages from very early on, possibly even before they actively use the corresponding linguistic devices in their speech production.

To conclude this section, it is probably not exaggerated to state that there exists an almost unanimous consensus in current research that children exposed to two languages from birth normally succeed from very early on in separating their languages and in developing two distinct grammatical systems. I will return to the question of what constitutes sufficient exposure. What matters at this point is that early differentiation of grammars represents a necessary condition for the development of two native grammars, that is, competences that are qualitatively not distinct from those of the corresponding monolinguals. In other words, the findings summarized here show that simultaneous bilingualism indeed qualifies as an instance of dual first language acquisition during early phases of grammatical development. The question then is whether this is still the case during later developmental phases.

## *The course of grammatical development*

The issue at stake is whether the two grammars of bilinguals develop independently, following the same trajectories as in the corresponding monolinguals, or whether the course of development is altered as a result of interaction between the two systems. In other words, the primary concern is to determine whether cross-linguistic interaction affects the grammatical competence of bilinguals. Only if this is the case, do we have to conclude that simultaneous bilingualism is not an instance of dual first language acquisition but a different type of language acquisition. This is not to say that the study of cross-linguistic interaction in language use should be neglected, but rather that it can provide insights into different aspects of child bilingualism. If, namely, such effects appear in children's speech although the underlying knowledge systems are not affected, as predicted by the Autonomous Development Hypothesis (Meisel, 2001), findings of this sort can contribute to the unraveling of mechanisms of bilingual speech production.

The question then is how to detect empirically the possible instances of cross-linguistic interaction. Since, by definition, they are the result of a fusion of grammatical (sub)systems or of a failure to inhibit the activation of the respective other language when one language is being used, these phenomena should be confined to bilingual speech. In other words, the point of reference is once again the language of monolinguals. In view of the considerable amount of variation across individuals, among monolinguals as well as bilinguals, these comparisons must focus on inter-individually invariant aspects of language acquisition and distinguish between quantitative and qualitative differences. Varying acquisition rates, for example, count as quantitative effects, as do changing frequencies in the use of particular constructions, or the temporary preference for particular linguistic devices. In colloquial German, for example, demonstratives can replace subject pronouns: *sie/die arbeitet gerade im Garten* "she is working in the garden right now." Although their pragmatic functions differ in adult language, some children prefer one of these categories whereas others use them interchangeably during early phases of syntactic development. If bilinguals consistently use one of these options, this still constitutes a quantitative difference, resulting from the overuse of an option also offered by the monolingual grammar. Only if a particular construction is attested exclusively in bilingual speech, can this be argued to be a qualitative difference, assuming that this construction is not generated by the monolingual grammar. Note that, in principle, both quantitative and qualitative differences can be caused by cross-linguistic influence. Whether this is a plausible explanation of the observed facts needs to be demonstrated by empirically based analyses. As argued above, contrasting inter-individually invariant aspects of linguistic development is probably the most promising approach to this research task.

In order to decide whether simultaneous bilingualism qualifies as an instance of 2L1 acquisition, the crucial issue is to determine whether cross-linguistic interaction can affect bilinguals' grammatical knowledge in the course of acquisition. As Paradis and Genesee (1996) pointed out, grammatical interdependence in acquisition can take the form of acceleration or delay of development, or of transfer from one system to the other. Yet only the latter represents a qualitative change, assuming that the notion of transfer implies that a grammatical device proper to one grammar ($G_a$) is incorporated into the other one ($G_b$) where it never appears in the competence of monolinguals. However, if the course of development in one language were altered under the influence of the other language, this would also have to count as a qualitative change, even if no alien property is incorporated into $G_b$. A reliable criterion to detect this kind of interdependence in the development of grammars is derived from the fact that L1 acquisition is characterized by an ordered series of developmental milestones, cf. Guasti (2002). More specifically, children uniformly proceed through invariant developmental sequences defined in terms particular grammatical phenomena; see Meisel (2011). For example, as mentioned before, inflectional morphology on verbs encoding subject agreement emerges before object-verb agreement, and subject clitics are acquired before object clitics. The question then is whether language-specific orderings of sequences can be altered under the influence of the other language of bilinguals.

In principle, acceleration or delay could lead to such effects rather than merely causing the rate at which learners proceed through a sequence to be accelerated or protracted. A possible scenario of qualitative change is one where successful acquisition of a particular phenomenon in $G_a$ triggers the acquisition of the corresponding device in $G_b$ where it normally emerges only later. For example, subject-verb inversion in interrogatives is extremely rare in colloquial varieties of French and therefore also in child-directed speech. In fact, inversion of clitic subjects is virtually non-existent in colloquial speech. Hence, children only hear constructions like *où ils son*t? "where they are = where are they?" or *qui tu as vu hier*? "who you have seen yesterday = who did you see yesterday?" As a result, clitic inversion is not attested in the speech of French children before the age of 5;0 (Meisel, Elsig, & Bonnesen, 2011). In syntactic terms, subject-verb inversion is commonly analyzed as movement of finite verbs to the head of the CP (complementizer phrase), an operation which grammars of colloquial French have been argued to lack, although it is part of the grammar of more formal varieties of the language. In German, on the other hand, it is used frequently—not only in interrogatives but also in V2 constructions—and acquired early, as mentioned above. Consequently, if cross-linguistic interaction plays a significant role in bilingual acquisition, one might expect that French-German bilinguals acquire verb raising to CP earlier than monolinguals, triggered by their syntactic knowledge of inversion in German. After all, this operation is not alien to French since it is part of the grammars of formal varieties, where it is typically acquired at around age 5;0, and occasional examples of subject clitic inversion do occur in the primary linguistic data to which children are exposed even before that age. A similar effect has been predicted to happen in case one language develops at a much faster rate than the other; the former might then fulfill a "booster function" for the latter; cf. Gawlitzek-Maiwald and Tracy (1996).

The obvious question to ask at this point is whether qualitative alterations actually happen in the grammatical development of simultaneous bilinguals, and also whether cross-linguistic interaction can trigger such effects. Anticipating the result of the following short review of a long debate, the answers to both questions are negative. Admittedly, this is a more controversial issue than the one concerning early grammatical differentiation. Still, to my knowledge no compelling evidence has been presented to date that would invalidate the Autonomous Development Hypothesis. In what follows, I will briefly explain this claim; for a discussion of specific aspects, I must refer to the state-of-the-art discussions by Meisel (2007) or Serratrice (2013). They show that the issue of autonomous *versus* interdependent development was one of the most intensely debated problems in research on bilingual acquisition during the first decade of the new century.

The debate started during the second half of the 1990s when researchers, almost all of them in conformance with the claim that bilinguals differentiate grammatical systems early on, suggested that subsequent developments showed effects of cross-linguistic interaction. Earlier studies commonly explained similar observations in terms of transfer, and they usually assumed, explicitly or more often implicitly, that dominance of the source language caused this to happen. Unfortunately, the term "dominance" is not used uniformly, referring either to the majority language in cases of family bilingualism in monolingual societies, or to the language to which the child is most frequently exposed, the one preferred by the child, the one developing faster, and so on. More importantly, "transfer" is also used inconsistently, referring to interdependence between grammatical systems as well as to interferences in processes of language use. It is therefore impossible to assess the relevance of findings by these investigations for the Autonomous Development Hypothesis.

One study that avoids these ambiguities is the one by Yip and Matthews (2000), reporting on the acquisition of Cantonese and English by a bilingual boy. They define dominance primarily in terms of MLU values and transfer as incorporation of grammatical properties from another language. Their strongest case concerns word order in wh-constructions, suggesting that wh-in situ is transferred from Cantonese into English, precisely during those periods when MLU values are higher in Cantonese than in English. Yet although this child does use these constructions more often than monolingual English children, the conclusion that wh-in situ is a property of his English grammar is not warranted. Not only is the first occurrence of an English wh-expression one with the wh-word in initial position, constructions with clause-initial wh co-occur with wh-in situ at every moment of his linguistic development, as is documented here. In other words, the English option is not replaced by the Cantonese one; rather, the two construction types co-exist. If the claim is that transfer resulted in an English grammar offering both options, we are not told how this grammar works. Moreover, the authors of this study do not explain how the boy succeeded in subsequently unlearning the Cantonese parts of his English grammar, a serious acquisition puzzle, for there cannot possibly exist positive evidence in the primary linguistic data triggering this grammatical restructuring. In reality, the observed pattern of usage, with wh-initial constructions chronologically preceding and co-occurring with wh-in situ, suggests that these uses result from the activation of Cantonese grammatical principles while processing English, rather than from a restructuring of the grammar of English. In order to corroborate the claim that cross-linguistic influence can lead to the incorporation of grammatical properties into the other grammar, one needs to show that a property of $G_a$ replaces the equivalent one in $G_b$. One also needs to explain how it is subsequently replaced again by the one required by the target version of $G_b$, for, to my knowledge, alleged effects of interdependence have never been claimed to be irreparable.

In fact, two observations redirected the course of the debate on interdependence and stimulated it at the same time. Firstly, dominance does not reliably predict which of the two languages of a bilingual will be affected by cross-linguistic interaction. Secondly, not all aspects of grammars are equally concerned; rather, there seem to exist "vulnerable domains of grammar" (Meisel, 2001, p. 36). Hulk and Müller (2000) and Müller and Hulk (2001) were especially influential in triggering a debate on possible structural causes and constraints on cross-linguistic interaction. They argued that structural domains particularly prone to cross-linguistic influence are the ones in which the two languages overlap, provided they are located at interfaces at which grammars interact with other cognitive systems. Structural ambiguity in one of the languages is thus what makes such areas vulnerable, "ambiguity" meaning that the surface realization of a construction allows for more than one structural analysis by the learner. If the other language provides positive evidence in favor of one analysis, bilinguals are predicted to carry over the unambiguous solution to the language exhibiting ambiguity. An example discussed by Müller and Hulk (2001) concerns the omission of objects. Whereas objects must be lexically realized in Romance languages, they may be dropped in colloquial speech in languages like German if they are topicalized, that is, if they are placed in clause-initial position, as in *hab ich gesehen* "have I seen = that I have seen." According to these authors, French children encounter apparent evidence for object-drop in constructions with topicalized (*ça j'ai vu* "that I have seen") or cliticized (*Jean le voit* "John him sees" = "John sees him") objects, for the canonical (post-verbal) object position is lexically empty in these cases. This instance of structural ambiguity is located in the C-domain (the structural level of the CP), arguably an interface level where syntactic and pragmatic information are exchanged, and it might therefore lead children to assume that French is an object-drop language. French monolinguals indeed omit objects, but only occasionally and only for a brief period. Following Müller and Hulk (2001), they abandon this option as soon as they discover that the empty object position is licensed by the pre-verbal clitic. German-French bilinguals, on the other hand, use target deviant object omissions in French more frequently and for a longer period, influenced, according to these authors, by the input they receive from German. Müller and Hulk (2001) interpreted this as a case of unidirectional influence, whereas dominance can operate in both directions.

This research agenda, focusing on structural constraints on interdependent developments, has been successful in that it confirmed that interaction does not occur across the board. Limiting vulnerability to phenomena at the C-domain, however, turned out not to be the most adequate generalization. A considerable number of phenomena that are located at the C-domain have been demonstrated *not* to be affected by cross-linguistic influence, although they involve information exchange at interface levels and do exhibit ambiguity; see Meisel (2007) for a more detailed discussion. Still, the suggestion that interface phenomena are particularly vulnerable has proven to be a fruitful one and has been elaborated in more detail as the Interface Hypothesis. Sorace and Filiaci (2006) initially proposed this hypothesis in order to account for residual optionality in adult L2 learners, but it has subsequently been extended to 2L1 and other types of bilingualism. In its revised version, it refers not only to the syntax-pragmatics but also to the syntax-semantics interface and to grammar–internal interfaces at which syntax, morphology, phonology, and semantics interact. For an updated version of the Interface Hypothesis and its reception by acquisition researchers, see Sorace (2011) and the commentaries on this target paper.

Any attempt to summarize the results of this research can, of course, only be tentative, since this is a very prolific and still ongoing debate; see, for example, Thomas and Mennen (2014). It has undoubtedly been successful in that it discovered a number of phenomena that can be argued with reasonable certainty to result from cross-linguistic interaction. Yet cross-linguistic influence has also been shown to emerge with phenomena *not* related to interfaces; cf. Serratrice (2013, p. 13ff). In other words, neither structural ambiguity nor being located at an interface level are necessary conditions for cross-linguistic interaction to happen. Moreover, numerous studies have shown that effects of interaction do not appear in the speech of *all* children, but only in that of some individuals. This is to say that the predicted effects are not necessarily triggered when the structural conditions are met. They thus represent neither necessary nor sufficient conditions for cross-linguistic interaction. This disappointing conclusion suggests that structural properties represent only one of several interacting factors. Language dominance and, more generally, the learners' access to the target language, count among these factors, after all. However, in order to be able to weigh each of them and to understand how they interact, more research is needed.

Concerning the questions asked at the beginning of this section, it should have become apparent that effects of cross-linguistic influence do exist and that some grammatical domains are more likely to be affected than others. Linguistic interdependence is thus constrained by structural properties, even if these alone cannot fully account for the particularities of bilingual speech. Both delay and acceleration occur in their grammatical development. In fact, in their vast majority, effects of interaction are quantitative in nature. Nevertheless, as stated by Serratrice (2013, p. 5) qualitative differences have also been observed, for example, placement of wh-words in clause-final position (Yip & Matthews, 2000). However, the crucial question is whether attested instances of cross-linguistic interaction affect the grammatical knowledge of bilingual children. The reason why I think this question should be answered negatively is that no evidence has as yet been provided suggesting that the developmental trajectory of bilinguals differs from that of monolinguals or that elements from one grammar are incorporated into the other. If this were really the case, it should be possible to identify developmental phases during which the alien constructions are used exclusively. Yet the available empirical facts show that they occur simultaneously with the corresponding constructions of the target systems, either as the predominant or as the less favored choice. Most importantly, effects of interaction can disappear and reappear in the course of acquisition. If they were to be analyzed as instances of grammatical transfer, one would have to explain these repeated switches between grammatical systems or, much worse, from an acquisition perspective, their apparent unlearning and relearning.

In reality, these developmental patterns exhibit characteristics of variable use of two distinct grammatical systems due to changes in the on-line activation of the other language. In fact, we know that the two languages of bilinguals are always simultaneously active (Green, 1998); for further references cf. Chapter 14 (Kroll & Ma). Hence, one of them needs to be inhibited in communicative interaction. What seems to happen in cross-linguistic interaction is that inhibition of the respective other language is not entirely successful. This explanation accounts for the observed developmental patterns as well as for the fact that interdependency effects appear in only some children. It can also explain why external factors like dominance, or quantity and quality of input trigger the emergence of alien constructions, for example, Cantonese patterns in English recordings when Cantonese MLU values are higher than those in English (Yip & Matthews, 2000). In fact, experimental studies, for example, Nicoladis (2006), provide further support for the idea that simultaneous co-activation of both languages is likely to be responsible for instances of cross-linguistic interaction; see Serratrice (2013, p. 16 ff.) for a brief survey of this research.

## *Ultimate attainment*

The third assumption, implied by the hypothesis that simultaneous bilingualism is a case of dual first language acquisition, concerns the grammatical knowledge ultimately attained by bilingual children. The claim is that bilinguals attain grammatical competences not distinct from those of monolinguals. Serratrice (2013, p. 4) correctly observed that there is no "overwhelming support" for this view. This is not to say, however, that strong arguments or empirical evidence had been presented against this claim. Rather, it is an issue rarely addressed directly in research on child bilingualism. In fact, it is not possible to provide an exhaustive list of criteria that must be met in order for an individual to qualify as a native speaker of a language. The considerable extent of variation across individuals, mentioned at the beginning of the preceding section, led to the conclusion that comparisons between acquisition types should refer to inter-individually invariant aspects. This also applies to attempts to define what constitutes a native competence in a given language. Since grammars of individual members of a speech community never comprise the full set of properties attributable to all varieties of the language, competent native speakers of a specific language must be defined by a grammatical core common to the mental grammars of all speakers.

Proceeding in this way should help to avoid a dilemma which has haunted the language sciences since their early days in the nineteenth century. "Language" is not a well-defined term, as far as grammatical knowledge is concerned. From a cognitive perspective, only mental grammars of individual speakers are proper objects of investigation. Yet although there arguably exist no two fully identical ones, not even among speakers of one (regional, social, situational) variety of a language, contenting oneself with the description of idiolects, of which there exist as many as speakers, is not an adequate solution. It misses an important generalization, namely that these grammars are not equidistant within the variation space defined by the human language capacity. Put differently, although idiolectal grammars of 'Spanish' are not fully identical, they share many more properties among themselves than with any idiolectal grammar of "Basque." "Language" can thus be understood as referring to the unification of properties represented in the mental grammars of individuals.

What falls outside the domain of shared grammatical properties has to be determined case by case, and comparisons between competences of bilinguals and monolinguals can obviously not refer to the thus identified phenomena. Idiolectal variation across individuals does not, however, make such comparisons impossible or illicit. It goes without saying that the same is true for comparisons between monolinguals. Stating that mental grammars of native speakers of a language are not fully identical (Dąbrowska, 2012), is therefore neither a new nor an original insight. Children acquire the grammatical properties underlying the speech data to which they are exposed, and they are all exposed to numerous idiolects. Concluding from the resulting inter-individual grammatical variability that acquisition is not guided by universal principles is a *non sequitur*, and contrary to what Dąbrowska (2012) believes, guidance by Universal Grammar does not mean that inductive learning plays no role; cf. Meisel (2011).

In sum, if we do not find strong support for the claim that bilinguals are able to develop grammatical competences that are qualitatively equivalent to those of monolinguals, it does not follow that this assumption is incorrect, nor does it mean that evidence in its support could not be obtained. Rather, in view of early grammatical differentiation and autonomous development, it is plausible to assume that bilinguals are on their way to native competences in both languages. The null hypothesis thus is that they are able to attain this goal, and it should be regarded as valid until evidence to the contrary is presented. Such counterevidence must refer to grammatical properties shared by monolinguals but not by bilinguals.

To conclude this section, we may say that research on child bilingualism has established beyond reasonable doubt that children acquiring two languages simultaneously differentiate the grammars of these language very early (Dual System Hypothesis) and that they seem to be able to keep them apart during subsequent developmental phases (Autonomous Development Hypothesis), even if cross-linguistic interaction can affect their language use. Finally, currently available evidence does not question the claim that the grammatical knowledge ultimately attained by simultaneous bilinguals is equivalent to that of the corresponding monolinguals. Taken together, these findings corroborate the assumption that the human language capacity represents an endowment for bilingualism. Yet although exposure to the two languages at an early age may well be a necessary condition for the acquisition of native competences in both of them (see the next section), it is certainly not a sufficient one. Rather, insufficient exposure to the primary linguistic data has been argued to result in incomplete development of grammars, for example in the acquisition of heritage languages; cf. Montrul (2008).

The claim that drastically reduced exposure to a language severely reduces the learner's possibility of acquiring a native competence, is certainly not a controversial one. However, the assumption that heritage language speakers typically acquire an "incomplete" grammar is more controversial and difficult to assess. This is partly due to the fact that the term "heritage language learner," as it is used by different authors, oscillates considerably, referring to early as well as late bilinguals, including L2 learners. For the latter, the idea of incomplete acquisition is less controversial than for 2L1 children. Heritage language speakers are commonly defined as L1 learners of a minority language who end up speaking the majority language as their dominant language and who acquire a reduced variety of the heritage language, used primarily in family settings. The question is whether this constitutes a failure to acquire a native competence in spite of exposure to the language in early childhood, as is suggested by Montrul (2008, p. 19) who defines incompleteness as "non-native like attainment." If, namely, incomplete acquisition is a case where "an individual fails to learn the entire system of a given language" (Polinsky, 2006, p. 194), this includes instances where learners fail to acquire properties of the target system to which they are not exposed because their use of the language is confined to family contexts. Only the former case, the acquisition of a non-native competence, is relevant for the current discussion. More research is needed in order to determine to what extent this happens, if at all.

Moreover, it is crucial to establish what constitutes "insufficient" exposure to the PLD. So far, research on heritage language learners has not yet answered this question; cf. Putnam and Sánchez (2013, p. 482). Note that bilingual children are able to develop two native competences even though the relative time of exposure to each of the languages amounts at best to 50% of the amount of child-directed speech. In fact, even when the two languages are not balanced with respect to their availability for the child, bilinguals can attain two native competences (Montrul, 2008). Yet although the acquisition device is quite obviously a robust enough device to enable children to acquire native grammars in spite of reduced input, there can be no doubt that a minimum amount of exposure to the input data is necessary for them to be able to achieve this, for, trivially, zero input will result in acquisition failure. The challenge thus is to determine how decreasing amounts of input affect grammatical development and to quantify the minimum threshold for successful acquisition.

## Successive bilingualism: Child second language acquisition

The preceding summary of research on the simultaneous acquisition of two languages has led to the conclusion that the Language Making Capacity enables children to develop two native competences. Mere exposure to primary linguistic data from both languages in communicative interactions suffices for them to attain this goal, although further research is needed to determine what constitutes sufficient exposure for bilingual first language acquisition to be possible. The last issue to be addressed in this paper concerns a different type of possible limitation of the Language Making Capacity, age of onset of acquisition. The question is whether exposure from birth is a necessary condition for bilingual L1 acquisition to be possible or whether two native competences can also be attained in successive acquisition in early childhood.

This question is motivated by the observation that adult L2 differs from L1 acquisition in substantial ways. This claim in itself is not controversial, for it is generally assumed that L1 and L2 learners rely on substantively different kinds of knowledge at the initial state, and it is widely agreed that they subsequently proceed through distinct developmental sequences. However, controversy does exist concerning the nature of the differences, that is, whether they are superficial in nature or whether they reflect different kinds of grammatical knowledge. What is also controversial is whether it is possible, in principle, for adult learners to attain native competence in an L2. I will refrain from engaging in a discussion of the controversial issues. For the present purpose, it must suffice to acknowledge the existence of substantial differences between L1 and adult L2 acquisition. For example, children acquiring German as an L1 never place non-finite verbs incorrectly in the structural second (V2) position of clauses, whereas L2 learners do so frequently. Such differences can serve as points of reference in contrasting early successive bilinguals with L1, 2L1, and adult L2 speakers. If they use constructions not attested in the speech of another learner type, these can be considered as qualitative differences distinguishing acquisition types.

Linguistic differences between adult L2 and child L1 speakers have been argued to result from age-related changes of the innate acquisition device, caused primarily by brain maturation. The underlying hypothesis is that the ability to develop a native competence weakens in the course of development; cf. Hyltenstam and Abrahamsson (2003) for an insightful discussion of this topic. However, only some grammatical domains are affected by maturational changes—and not all of them simultaneously. Phonology, for example, does not follow the same developmental agenda as syntax or morphology, and asynchronous developments happen even within subcomponents of grammar. Consequently, the acquisition of grammar is characterized by a series of *sensitive phases*, defined as periods of heightened sensitivity for particular aspects of grammar. Sensitive phases may cluster chronologically during particular developmental periods. Concerning the age ranges during which this happens, linguistic as well as neuropsychological research suggests that the age at around seven is a critical one, but recent findings indicate that age-related changes affecting morphosyntax can already be observed earlier.

This insight is due to a shift of research interests to successive acquisition in childhood; see Meisel (2009). Until about 10 years ago, little was known about learners first exposed to a second language before age 10. In fact, successive acquisition beginning during the first three years of life is still a largely unexplored topic. However, there now exists a substantial amount of research dedicated to successive acquisition setting in between ages three and eight. The evidence accumulated by these studies led to the conclusion that, if first exposure happens during the second half of the fourth year or later, acquisition exhibits some L2 properties, although it still resembles L1 development in other respects. Successive language acquisition in early childhood (age of onset ranging from approximately 3;6 through 7 years of age) has therefore been qualified as child L2 acquisition (cL2), a distinct acquisition type, sharing characteristics with L1 and increasingly with L2, depending on the age of first exposure.

There is still uncertainty about which domains of grammar are subject to age-related changes during early developmental phases. Since neither linguistic nor neuropsychological theorizing provides a developmental agenda, it must be established inductively, scrutinizing phenomena known to represent acquisition difficulties for L2 learners. Proceeding accordingly, studies providing empirical evidence for L2 properties in cL2 speech have revealed that inflectional morphology is particularly vulnerable at this early age. The phenomena concerned include the expression of finiteness, tense and aspect, root infinitives, gender assignment and agreement, the clitic status of subject and object clitics as well as the placement of object clitics if the target requires placement in a position which differs from that of the respective DP. However, the acquisition of syntax is also concerned, although the empirical evidence is still somewhat scarce in this domain. Nevertheless, a study investigating the acquisition of German by Polish-speaking children (age of onset of acquisition 3;8–4;7) found L2 properties in the syntax of these children. Contrary to German L1 children, they did not use OV as their preferred order in German, they placed finite verbs frequently in target-deviant *V3 position during early phases, and they moved non-finite verbs to the V2 position, a feature unambiguously characterizing L2 learners; cf. Sopata (2011).

In sum, successive language acquisition differs in important ways from L1 development, varying according to age of onset of acquisition. cL2 shares crucial properties with adult L2 acquisition, and both differ from monolingual as well as bilingual first language development. The grammatical features in which early cL2 resembles adult L2 acquisition fall primarily into the domain of inflectional morphology, but some aspects of syntax are also concerned at an early age. These and similar research results corroborate the claim that exposure to two languages from birth is a necessary condition for acquiring native competences in both languages.

## Conclusions and open questions

Children exposed to two languages from birth are able to acquire native grammatical competences in both of them. Exposure to the ambient languages in communicative interactions suffices for this to become possible. They differentiate the grammatical systems from very early on, and they then proceed through the same developmental sequences as the corresponding monolinguals. Cross-linguistic influence during later developments is mostly quantitative in nature. Qualitative effects have been observed only rarely, and they have been argued not to affect children's grammatical knowledge; rather, they result from insufficient inhibition of the respective other language. However, this issue requires further research. The same is true of the claim that simultaneous bilinguals ultimately attain grammatical knowledge not distinct from that of monolinguals.

Whereas simultaneous acquisition of languages qualifies as bilingual *first* language development, successive acquisition differs in at least some aspects from L1 development if age of onset of acquisition happens at around age 3;6 or later. It exhibits properties characteristic of adult L2 acquisition but never attested in the speech of L1 children. Consequently, exposure to languages from birth is a necessary condition for the acquisition of native competences, although it is arguably not a sufficient condition, for sufficient exposure to the target language is also required. An adequate definition of what constitutes "sufficient" exposure is, however, still an urgent *desideratum* for research on early bilingualism.

# REFERENCES


Bergman, C. R. (1976). Interference versus independent development in infant bilingualism. In G. D. Keller, R. V. Teschner, & S. Viera (Eds.), *Bilingualism In the Bicentennial and Beyond* (pp. 86–96). New York: Bilingual Press/Editorial Bilingüe.

De Houwer, A. (1990). *The Acquisition of Two Languages From Birth: A Case Study*. Cambridge: Cambridge University Press.

De Houwer, A. (1995). Bilingual language acquisition. In P. Fletcher, & B. MacWhinney (Eds.), *The Handbook of Child Language* (pp. 219–250). Oxford: Blackwell.

De Houwer, A. (2005). Early bilingual acquisition. In J. F. Kroll, & A.M.B. de Groot (Eds.), *Handbook of Bilingualism* (pp. 30–48). Oxford: Oxford University Press. doi: 10.1017/CBO9780511519789

Ezeizabarrena, M. (1994). Primeras formas verbales de concordancia en euskera. In J. M. Meisel (Ed.), *La Adquisición Del Vasco y Del Castellano* (pp. 181–229). Frankfurt/Main: Vervuert.

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, 2, 219–253. doi: 10.1075/lab.2.3.01dab

Gawlitzek-Maiwald, I. & Rosemarie, T. (1996). Bilingual bootstrapping. *Linguistics*, 34, 901–926. doi: 10.1515/ling.1996.34.5.901

Genesee, F. (1989). Early bilingual development, one language or two? *Journal of Child Language*, 16, 161–179. doi: 10.1017/S0305000900013490

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81. doi: 10.1017/S1366728998000133

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3–15. doi: 10.1016/0093-934X(89)90048-5

Grosjean, F. (2008). *Studying Bilinguals*. Oxford: Oxford University Press.

Guasti, M. (2002). *Language Acquisition: The Growth of Grammar*. Cambridge, MA: The MIT Press.

Hulk, A. & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition*, 3, 227–244. doi: 10.1017/S1366728900000353

Hyltenstam, K. & Abrahamsson, N. (2003). Maturational constraints in second language acquisition. In C. Doughty, & M. H. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 539–588). Oxford: Blackwell.

Leopold, W. F. (1939–1949). *Speech Development of a Bilingual Child: A Linguist's Record*. Evanston, Illinois: Northwestern University Press, vols. 1–4; New York: AMS Press, 1970.

Mahlau, A. (1989). Orden de palabras y estructura oracional en los niños bilingües. In J. M. Meisel (Ed.), *La Adquisición Del Vasco y Del Castellano* (pp. 69–111). Frankfurt/Main: Vervuert.

Meisel, J. M. (1986). Word order and case marking in early child language: Evidence from simultaneous acquisition of two first languages: French and German. *Linguistics*, 24, 123–183. doi: 10.1515/ling.1986.24.1.123

Meisel, J. M. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity, and Loss* (pp. 13–40). Cambridge: Cambridge University Press.

Meisel, J. M. (1994). Getting FAT: Finiteness, Agreement and Tense in early grammars.

In J. M. Meisel (Ed.), *Bilingual First Language Acquisition: French and German Grammatical Development* (pp. 89–129). Amsterdam: John Benjamins.

Meisel, J. M. (2001). The simultaneous acquisition of two first languages: Early differentiation and subsequent development of grammars. In J. Cenoz, & F. Genesee (Eds.), *Trends in Bilingual Acquisition* (pp. 11–41). Amsterdam: John Benjamins.

Meisel, J. M. (2004). The bilingual child. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 91–113). Oxford: Blackwell.

Meisel, J. M. (2007). On autonomous syntactic development in multiple first language acquisition. In H. Caunt-Nulton, S. Kulatilake, & I.–H. Woo (Eds.), *Proceedings of the 31st Boston University conference on language development* (pp. 26–45). Somerville, MA: Cascadilla Press.

Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft*, 28, 5–34. doi: 10.1515/ZFSW.2009.002

Meisel, J. M. (2011). *First and Second Language Acquisition: Parallels and Differences*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511862694

Meisel, J. M., Elsig, M., & Bonnesen, M. (2011). Delayed acquisition of grammar in first language development: Subject-verb inversion and subject clitics in French interrogatives. *Linguistic Approaches to Bilingualism*, 1, 347–390. doi: 10.1075/lab.1.4.01mei

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism*. Amsterdam: John Benjamins. doi: 10.1075/sibil.39

Müller, N. & Hulk, A. (2001). Cross-linguistic influence in bilingual acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4, 1–21. doi: 10.1017/S1366728901310118

Nicoladis, E. (2006). Cross-linguistic transfer in adjective-noun strings by preschool bilingual children. *Bilingualism: Language and Cognition*, 9, 15–32. doi: 10.1017/S136672890500235X

Paradis, J. & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or independent? *Studies in Second Language Acquisition*, 18, 1–15. doi: 10.1017/S0272263100014662

Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics*, 14, 191–262.

Putnam, M. & Sánchez, L. (2013). What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3, 478–508. doi: http://dx.doi.org/10.1075/lab.3.4.04put

Quay, S. (1995). The bilingual lexicon: Implications for studies of language choice. *Journal of Child Language*, 22, 369–387. doi: 10.1017/S0305000900009831

Ronjat, J. (1913). *Le Développement Du Langage Observé Chez un Enfant Bilingue*. Paris : Librairie Ancienne H. Champion.

Sebastián-Gallés, N. & Bosch, L. (2005). Phonology and bilingualism. In J. F. Kroll, & A. M. B. de Groot (Eds.), *Handbook of Bilingualism* (pp. 68–87). Oxford: Oxford University Press.

Serratrice, L. (2013). Cross-linguistic influence in bilingual development: Determinants and mechanisms. *Linguistic Approaches to Bilingualism*, 3, 3–25. doi: 10.1075/lab.3.1.01ser

Sopata, A. (2011). Placement of infinitives in successive child language acquisition. In E. Rinke, & T. Kupisch (Eds.), *The Development of Grammar: Language Acquisition and Diachronic Change* (pp. 105–121). Amsterdam: John Benjamins.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, 1, 1–33. doi: 10.1075/lab.1.1.01sor

Sorace, A., & Filiaci, F. (2006) Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22, 339–368. doi: 10.1191/0267658306sr271oa

Taeschner, T. (1983). *The Sun is Feminine: A Study on Language Acquisition in Bilingual Children*. Berlin: Springer. doi: 10.1007/978-3-642-48329-5

Thomas, E. M. & Mennen, I. (Eds.) (2014). *Advances in the Study of Bilingualism*. Bristol: Multilingual Matters.

Volterra, V. & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language*, 5, 311–326. doi: 10.1017/S0305000900007492

Yip, V. & Matthews, S. (2000). Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism: Language and Cognition*, 3, 193–208. doi: 10.1017/S136672890000033X

# 29 The Development of Morphosyntax in Child and Adult Second Language Acquisition

GITA MARTOHARDJONO AND
ELAINE C. KLEIN

The Graduate Center, City University of New York

## Introduction

This chapter will focus on some of the broad issues in second language acquisition, along with the specific nature, development and processing of an important aspect of acquisition in both child and adult second language (L2) learners. From a wealth of L2 theories, we offer a generative and psycholinguistic perspective, examining how these contribute to our knowledge of the L2 development process. The chapter will also provide examples of current studies that have investigated some of the over-arching questions of how second language acquisition (L2A) is both similar to and different from first language acquisition.

## Child second language acquisition

In this section of the chapter, we first take the broad view by focusing on child L2A and asking the extent to which it not only differs from first language acquisition but how it is to be distinguished from *bilingual acquisition*, which is expanded in Chapter 28 of this volume. In this first section, we also provide an overview of some of the theoretical and methodological issues in child L2A and, finally, we take a narrower, deeper view: We do this by focusing on studies of the development of child L2 morphosyntax, the research on functional categories and the interface

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

between morphology and syntax, to illustrate how some of the theoretical and methodological issues in child L2A play out in experimental work.

## First language acquisition, bilingual acquisition, and child second language acquisition

Although there are many interesting issues to explore, the onset of child acquisition of a native language requires no elaboration when only one language is involved. We consider that to be *first language acquisition* (FLA or L1A). However, when the child is exposed to another language, generally around the age of three, scholars of child language acquisition need to make distinctions: If the child acquires both languages before three years old (others including Meisel, 2009 suggest slightly different age ranges), acquisition is usually considered to be *simultaneous* and, if successful, the child becomes a (simultaneous) bilingual, that is, having two "native" languages. Of course, complications arise because input and therefore development is not so neatly delineated: For example, there can be variation in exposure between the two languages, both qualitatively and quantitatively, along with age of onset of one versus the other, a matter examined by a number of researchers (e.g. Genesee, Paradis, & Crago, 2004; Meisel, 2009, Unsworth, 2013; Unsworth *et al.*, 2014).

McLaughlin (1978) and Bhatia and Ritchie (1999), among others, have distinguished bilingual acquisition from *second language acquisition,* which involves those cases when one of a child's two languages is learned later, that is, approximately after age three. Second language acquisition at its end point is also considered to be *sequential* or *successive* (rather than simultaneous) *bilingualism*. That is, the process involved in developing sequential or successive bilingualism comes under the study of second language development or acquisition, with most scholars agreeing that balanced bilingualism (native-like proficiency in two languages) may not, in reality, be attained except in exceptional cases. The importance of the distinction between simultaneous and successive bilingualism even in young children is not generally trivial; for example, it has been targeted in recent research by Unsworth *et al.* (e.g., 2014), who found that such differences mattered in the acquisition of some L2 Greek features but not those of L2 Dutch. Further, Tsimpli (2014) adds another dimension by examining the role of timing of L2 acquisition phenomena (i.e., early, late, very late in development), finding that such timing also has strong effects on development in simultaneous versus successive bilingual children.

The upper limit of child L2A, that is, when child L2A becomes adult L2A, is generally assumed to be around the early "teenage" (pre-pubescent) years, anywhere from 9 to 13 years of age, roughly speaking. That is, if the learner begins to develop another language after that time, we assume that adult L2 acquisition is taking place (to include adolescents), thereby tying physical development to acquisition. However, some researchers have suggested a lower age: Schwartz (2003) has argued that age seven be considered the end of child L2A rather than the onset of puberty. The issue of when "adult L2A" begins has been conflated with the arguments for and against a so-called Critical (or sensitive) Period Hypothesis for L2A, which we will not attempt to summarize here (but see Bialystock & Hakuta, 1999 and Meisel, 2009, for contrasting viewpoints); suffice it to say that there is not a sharp distinction between "child" and "adult" L2A but rather sensitive periods between the two. To reconcile conflicting views on the matter, Foster-Cohen (2001) offers a "sliding window" that considers a *continuum* between and within L1 and L2.

## Some theoretical and methodological issues in the study of child SLA

Within the generative and psycholinguistic paradigms, the following theoretical issues have been among the most important being investigated in second language studies of child development:

### *Innateness*

The assumption that the child learner is innately endowed for language underlies the generative approach by which such endowment is instantiated through Universal Grammar (UG). This explains why the child can acquire language, quickly and efficiently, in spite of the so-called *poverty of the stimulus*—an environment that supplies only impoverished input. The question for generative theorists is not whether nativism exists but to what extent UG is operative in L2A. Many studies have supported the Strong Continuity Hypothesis, which posits that UG is fully available to L2 learners from the start and continues throughout development (e.g., Epstein *et al.*, 1996; Grondin & White, 1996; Schwartz, 2003). On the other hand, opponents of UG have argued that children, whose production of L2 morphemes, for example, is non-target-like, are assumed to have impaired grammatical representations because such morphemes are not represented in their L1s and are, therefore, forever inaccessible, that is, an indication that these child L2 learners are not equipped with UG (Hawkins & Chan, 1997; Haznedar & Schwartz, 1997; Vainikka & Young-Scholten, 1996). This differs somewhat from researchers like Tsimpli (2014), Hawkins (2005), and Tsimpli and Dimitrakopoulou (2007) who have argued, with some differences, that "uninterpretable features" are inaccessible or resistent in L2A, particularly in adult L2A (see the second section of this chapter). Meisel (2009), in a slightly different mode, argues for fundamental differences between child and adult L2, following the Fundamental Difference Hypothesis originally proposed by Bley-Vroman in 1990, whereby children, though not adults, are endowed with an innate human "Language Making Capacity." He further claims that "child second language acquisition is indeed a type of acquisition in its own right" (2009, p. 30), sharing some features with child L1A and some with adult L2A.

## *Language transfer*

Another important theoretical question concerns the role of prior linguistic knowledge, that is, the extent to which the child's L1 impacts the development of the L2. The Full Access/Full Transfer Hypothesis (e.g., Haznedar, 2001; Schwartz & Sprouse, 1996; Schwartz, 2003), for example, posits that the child begins L2A with the entire L1 at her disposal at the earliest stages of L2 development and gradually refines her developing L2 grammar to match the L2 input. The extent to which this strong notion of *transfer* occurs has been the subject of much L2 research and much debate.

## *The role of input, age, and maturation in second language development*

This theoretical issue concerns the extent to which child L2A is similar to or different from adult L2A because of degree and quality of the input learners receive across time and biological maturation, with researchers generally acknowledging the confluence of the two (e.g., Tsimpli, 2014). That is, it is often difficult to disentangle the effects of the role of input and maturity or the age at which the learner begins and continues through the L2 acquisition process. As noted above, there are also hypotheses that argue that adults, alone, are burdened with a deficit or incompleteness of the full range of grammatical representations offered by UG due to post-puberty exposure to the L2 (e.g., the Failed Functional Features Hypotheses of Hawkins & Chan, 1997; Hawkins & Hattori, 2006). On a slightly different note, some recent studies in child acquisition have investigated outcomes when L2 input is delayed and L1 input is no longer available for these children; these are the cases of Internationally Adopted (IA) children, whose L1 development essentially ends when they are taken from their home environments, but L2 development may not begin until sometime later in their new country. In these studies, these (IA) children acquiring their L2 are essentially found to follow the same path as monolingual L1 learners, suggesting that the timing of exposure to the (L2) input is not as important a factor as the general availability of UG, which provides these children with access to the grammatical representations they need to build on.

## *The processes of L1A versus L2A*

For this theoretical issue, psycholinguistic and processing studies investigate the way in which children process their input along the continuum of language development as they develop a second language in contrast to a first language. Lakshmanan (2009), in an insightful critique, points out that—up until 2009 at least—there were virtually no studies using online processing in child L2A, leaving the child L2 developmental *procedures and processes* for further research. We report below on a recent L2 study involving online processing and examine how important such a methodological step has been to research in child SLA.

### *Variability between a learner's internal grammar and the learner's output*

The theoretical distinction between competence and performance has always been difficult to tease apart in actual research for the simple reason that all studies must necessarily measure a learner's knowledge through performance outcomes. However, it has also been strongly argued that a study's methodology can seriously undermine findings among L2 learners (Klein & Martohardjono, 1999; Klein, 2004; Lakshmanan, 2009).

This points to one of the major methodological issues in child L2 studies, where there is a focus on production (performance), which is often taken (mistakenly, in our view) to reflect the child's knowledge (competence). The focus on oral production, often spontaneous (e.g., Ionin & Wexler 2002), in child L2 research has skewed the data toward what is quite possibly (and indeed is, as shown below) an underrepresentation of what the child may actually know (see arguments along these lines in, e.g., Lakshmanan, 2009 and Chondrogianni & Marinis, 2012).

The case study approach, commonly used in child L2 studies, has the advantage of offering both quantitative and deep qualitative data; such studies are often longitudinal (a good thing, e.g., Pierce *et al.*, 2012) but generally rely on only a few individual children (not a good thing, e.g., Pierce *et al.*, 2012), making it difficult to generalize from such research. Also, few studies make use of elicited and comparison data from a reliable sample size of children (cf. Paradis *et al.*, 2008). Further, and very importantly, there have been very few studies in comprehension or online processing of child L2 learners, as noted above, areas that are more likely to reveal a child's internal grammatical system.

## Child L2 acquisition of morphosyntax

There are reported to be more studies of child L2 development of morphosyntax than any other area of language development (Miller, 2014, p. 31). Perhaps this is not surprising when we note that such studies, especially those involving the development of tensed morphemes, a subset of functional projections, have enabled scholars to examine child development from very targeted perspectives and theories, including the generative and psycholinguistic paradigms, both of which look at underlying rather than surface (only) manifestations of language.

Studies of functional categories have also enabled researchers to investigate some of the most important theoretical issues and questions in the field, including those described briefly above. Many studies of morphosyntactic development have compared child L2A and typical developing child L1A populations (e.g., Ionin & Wexler, 2002). Studies of child L2 morphosyntax have also invited comparisons with adult L2A, to be described more fully in the second half of this chapter, as well as comparisons with special populations such as internationally adopted children and atypical populations. We examine a sample of these studies next.

Investigation of child L2 acquisition of morphosyntax began with the early studies of Dulay and Burt (1973; Dulay, Burt, & Krashen, 1982) who found similar developmental orders of English morphemes by child L2 learners, across L1s. They also found great variability in the production of tensed and non-tensed morphemes, with production of the latter group far more accurate. In a later review of the early (1970s, 1980s) "morpheme studies," Zobl and Liceras (1994) conclude, in particular, that L2 learners from different language backgrounds acquire the non-inflected suppletive BE form before inflectional tense endings. Their overall conclusion is that functional categories are available in L2 grammars, though not necessarily affixal forms, from early on.

More recent studies in the child L2 development of morphosyntax (e.g., Ionin & Wexler 2002; Paradis, 2005; Paradis *et al.*, 2008) have also found that both L1 and child L2 learners have difficulty acquiring tense-marking inflections (i.e., English past tense *-ed*, or third-person singular *–s)* compared to affixes unrelated to tense-marking (i.e., progressive –ing; plural –s*)*. However, there are at least two ways in which L1 and L2 children differ in their morphosyntactic development (Pierce *et al.*, 2012): First, as earlier noted for L2 learners in general by Zobl and Liceras (1994), child L2 learners, in contrast to child L1 learners, acquire the BE tense-marking morpheme early (e.g., the copular: e.g., *I am sad,* and the auxiliary: e.g., *she is singing),* as later evidenced in studies by Haznedar (2001), Ionin and Wexler (2002), Paradis (2005) and Paradis *et al.* (2011, 2008). Secondly, while both L1 and child L2 learners produce more errors of omitted inflections (e.g., *she walk)* than errors of commission (e.g., *she are eating),* child L1 learners produce relatively few of the latter type; that is, among child L2 learners, commission errors are much more frequent than in child L1A (Paradis, 2005).

Interestingly, L2 child learners but not L1 learners, tend to make errors involving BE overgeneration (e.g., he is go) in which they erroneously use BE "as a type of all-purpose marker of tense or agreement" (Pierce *et al.*, 2012, p. 1077); that is, the claim appears to be that the child's production might generally be "he go" but that the child marks tense (or agreement) by adding a form of 'be' to the structure, in this case "is." Such overgeneration of the auxiliary BE was found by Ionin and Wexler (2002) to constitute approximately one quarter of all BE utterances of the twenty Russian children learning English that the researchers were studying. Considering the clear evidence of verbal morphology in child L2A, though not necessarily inflectional affixes early on, Ionin and Wexler conclude, along with others, that child L2 learners' grammars are fully equipped with functional categories from the very beginning, providing support for the theoretical position that UG is operative in early L2 child grammars and, more specifically, the Strong Continuity Hypothesis. They also argue that the omission of verbal affixes is potentially due to difficulty in the outward manifestations of surface morphology, rather than impairment of functional features, in accord with the Missing Surface Inflection Hypothesis (Prévost & White, 2000).

The Missing Surface Inflection Hypothesis (MSIH) has also been invoked to explain the findings of other child L2 studies, like that of Haznedar and Schwartz (1997) and Haznedar (2001) who studied a Turkish child learning L2 English. Their findings of non-target-like finite forms in the developing L2 grammar of their learner were explained as being substitutions for target finite forms, which were too difficult for surface mapping to take place, that is, the MSIH. Underlyingly, however, the researchers claim that knowledge of tense by the learner was apparent.

Importantly, evidence of features transferred from the L1 to the L2 have been found in some morphosyntactic studies of child L2 learners and not in others. In a study of a Turkish child learning L2 English, Haznedar and Schwartz (1997) argued that distinct properties of head-final Turkish were apparent in the child's early development of English, a head-initial language (an interpretation that was later disputed by Lakshmanan & Selinker, 2001).

In another psycholinguistic study using the case study approach, Kwon and Han (2008) investigated both transfer from the L1 to the L2 and reverse transfer from the L2 to the L1 in a single L1 Korean speaker (aged 3;6) learning L2 English. Using naturalistic data on the learner's L2 and L1 production over three phases, during which time the child either lived in Korea or in the United States, three morphosyntactic features were targeted because of their different realizations in the L1 Korean and L2 English: negation, the regular plural, and the possessive. Results showed different patterns for each of the features, with L1 transfer of the possessive a constant throughout the data collection. For the other two features, transfer in different directions occurred depending upon the dominance of the L1 or the L2 at any particular time: Thus direct correlations were found between L1 transfer and L1 dominance, on the one hand, and reverse transfer and L2 dominance, on the other. It is clear, then, that many factors influence the occurrence, the nature, the direction, and the effects of transfer in child L2A and research continues to investigate these complex issues.

Relationships between a child's L1 and L2 often invoke other important theoretical questions: For example, in a study of a Chinese child acquiring L2 German, Lee (2008) reports that the learner easily acquired the tense features of the L2 input, even though Chinese does not have morphological realization for tense. Thus, this study indicates that the L1 does not hinder the learner from access to the full range of functional features offered by UG and also does not support the notion that the child automatically assumes L1 functional features throughout early L2 development, confirming the Strong Continuity/Full Access Hypothesis but not supporting Full Transfer.

Chondrogianni and Marinis (2012) argue that time of exposure to the L2 input, not surprisingly, has an impact on the acquisition of functional categories: those children with at least three years in the ambient environment show more target-like structures than those with less exposure (Chondrogianni & Marinis, 2012). As important, however, is how the data are collected: That is, children in production tasks, even with more lengthy exposure to the L2, continue to exhibit missed inflections, which some researchers (e.g., Vainikka & Young-Scholten, 1996) have argued are due to deficits in their early grammars; others have argued that consistently or variably missed inflections are due to surface mapping issues, as posited by the MSIH. However, Chondrogianni and Marinis (2012) provide examples of research from offline and online comprehension and processing studies that show that these same learners understand much more than they produce (e.g., Marinis, 2008; Marinis & Chondrogianni, 2011).

The study reported in Chondrogianni and Marinis (2012) is an important case in point: In this study, the researchers directly compared (28) typically developing L1 English-speaking children with (39) Turkish children exposed to English L2 for more than three years, with the two groups matched for age and demographics. The children were tested on a controlled elicited production task along with an online processing and comprehension task called the online word-monitoring task (Montgomery & Leonard, 1998, 2006), which the researchers had formerly used in other studies (e.g., Marinis, 2008; Marinis & Chondrogianni, 2011). In this task, children see a picture at the beginning of each trial (e.g., *a cake*). Then they hear a sentence containing the pictured word along with another sentence that contains a tensed verb (either in correct form, e.g., *he bakes*, or incorrect form, *He bake*, as in *Tom's father is a great baker. Most nights he bake(s)* ***cakes*** *for Tom*.) The task for the child participants is to press a button as quickly as possible, when they hear the pictured word in the sentence. The word-monitoring processing and comprehension task is based on the premise that the child's sensitivity to grammaticality in the sentence will affect her reaction time (RT) in picture-word identification. That is, if sensitive, then her RT should be slowed down when she encounters an ungrammatical verb (he bake*)*, where the required verbal inflection is missing, while a grammatical verb (*he bakes*) should not slow down her RT. No difference between RTs in responses to sentences containing ungrammatical and grammatical verbs would show that the child is unaffected by the grammaticality of tensed morphemes.

The findings of Chondrogianni and Marinis (2012) show that the L2 children's performance on the production task was significantly lower than that of the L1 children. In the processing task, both groups were similar in their sensitivity to tensed morphemes (-s and –ed) in English, although the L2 children had longer RTs, which was not surprising. Both groups of children had longer RTs in the ungrammatical conditions than the grammatical conditions for both tensed and non-tensed morphemes. Most importantly, there was a strong asymmetry between the production and the processing of tense morphemes among the L2 learners. That is, these children showed sensitivity to the ungrammaticality of morpheme omission even though their performance (in the elicited production task) showed optional or other non-target-like production of the same tense morphemes. These results were the same for children no matter how long their exposure to L2 English; in cases of less than three years or more than three years, the L2 learners all comprehended and processed grammatical vs. ungrammatical forms at a much higher rate than they were able to produce those target-like forms. The authors conclude that a limited and superficial focus on production only would have severely under-represented the knowledge these learners have with respect to morphosyntactic tense in their L2, providing a cautionary note to further research among child L2 learners. Claims of functional deficits and under-representation of UG categories, according to the authors, must be reconsidered in light of new studies using comprehension and processing methodologies.

## Adult L2 acquisition of morphosyntax

Like in child SLA, the development of morphosyntax has been of considerable interest to adult SLA research, given the persistent variability in the use of L2 grammatical morphemes well into higher levels of proficiency. What continues to vex researchers is why this variability persists even with prolonged use and exposure to the target language (Franceschina, 2005, Lardiere, 1998, 2007). Early explanations appealed to differences in the L1 and L2 grammars, for example in cases where learners whose L1 does not contain articles (e.g., Russian, Chinese) acquire a second language that does (e.g., English, German). These grammatical differences were thought to result in irreparable deficits in the learner's representation of the L2 grammar, making native-like attainment of morphosyntax impossible. Today, so-called "deficit" accounts continue to provide the theoretical framework for many SLA studies, but recent years have also seen a burgeoning of research into *performance-based* explanations of variability, especially for the acquisition of inflection. This work involves closer investigations of *processing* differences between native and non-native speakers. In this section we will briefly describe the most important grammar-based accounts of variability in the L2 acquisition of morphosyntax and then proceed to a summary of some of the more recent psycholinguistic research investigating processing issues and parser strategies in adult L2.

## Grammar-based explanations of morphosyntactic variability

Three categories of hypotheses have been advanced to explain morphosyntactic variability in adult SLA: Syntactic accounts, Phonological accounts, and Mapping accounts. Syntactic accounts constituted the first grammar-based explanations of non-target morphosyntax, particularly the absence or optionality of agreement and tense markers, and determiners. Since in generative grammar these markers reside under the functional nodes of syntactic projections, it was reasonable to assume that errors in morphosyntax were the result of a compromised syntax. Specifically, the deficit was articulated in terms of failure to project all levels of syntax, (e.g., the Minimal Tree Hypothesis of Vainikka & Young-Scholten, 1998) or failure to instantiate or assign syntactic features to the Functional nodes (e.g., the Failed Functional Feature Hypothesis of Beck, 1997; Hawkins & Chan, 1997; or the Interpretability Hypothesis of Tsimpli & Dimitrakopoulou, 2007; Hawkins & Liszka, 2003). Researchers subscribing to those models were divided as to whether adult learners would be able to recover from syntactic deficit and build target-like L2 grammars. On the one hand Critical Period proponents argued against recovery, using the persistence of optionality at advanced levels of proficiency as evidence. On the other hand proponents of the Full Transfer/Full Access hypothesis (FTFA, Schwartz & Sprouse, 1996) proposed that while error patterns were caused by full transfer of the L1 grammar at the L2 initial state, full access to UG allowed for the possibility of recovery and target-like attainment, once the requisite input was available.

A methodological weakness in these early studies is their reliance on oral production data as a main source of evidence. Spontaneous oral production can certainly be argued to provide evidence of automaticity; and automaticity, in turn, can be taken as an indicator of underlying syntactic representation. But earlier in the chapter we saw convincing arguments that overreliance on production masks and vastly underestimates what learners know about the L2 (Lakshmanan, 2009, Chondrogianni & Marinis, 2012 for L2 children; Martohardjono, Valian, & Klein, 2011 for L2 adults).

Phonological accounts offered an alternative to purely syntactic explanations of morphosyntactic development, tracing errors and variability to transfer of L1 phonological constraints to the L2 (e.g., Goad, White, & Steele, 2003; Solt *et al.*, 2004). While offering a different domain of language as the source of difficulty, these approaches are equally representational in nature. For example, Goad, White & Steele (2003) and Goad and White (2005) propose that inconsistent use of inflectional endings in the L2 English of Chinese speakers is caused by differences in the prosodic constraints on the two languages. Specifically, they claim that English, but not Chinese, allows adjunction to the Prosodic Word, which would explain the observed difficulty Chinese learners of English have in producing inflections, especially when they consist of word final consonant clusters. Phonological accounts from an Optionality Theoretic perspective can also be found in the literature, primarily in the work of Broselow (e.g., Broselow, 2004 and Broselow, Chen, & Wang, 1998).

For Mapping accounts variability is the result of difficulties in the computational space between the lexicon and syntax (for early accounts, see Lust, 1994 for L1 and Flynn & Martohardjono, 1994 for L2). Two hypotheses in particular have been developed in enough detail to be empirically testable: The Missing Surface Inflection Hypothesis, or MSIH (Prévost & White, 2000a, 2000b), and the Feature Reassignment Hypothesis (Lardiere, 2007). Both assume differential feature representation in the L1 and L2 lexicons but allow for eventual recovery given the necessary input.

The MSIH explains variability through underspecification of L2 lexical items. Under Minimalism, lexical insertion in the syntactic tree takes place when features of a lexical item (e.g., *watched*: V [+fin + past]) are matched to the terminal node in the syntax, in this case the Tense node, bearing the same features. The MSIH proposes that the learner's grammar contains fully specified syntactic nodes, but that individual items in the L2 lexicon may be underspecified. To explain patterns of errors in L2 French and German, Prévost and White (2000b) posit underspecification of finiteness in the learner lexicon: non-finite forms (e.g., French "manger," *to eat*) are not specified for finiteness (i.e., neither +/- finite) and may be inserted into nodes bearing the feature [+finite]. Finite forms, on the other hand, (French "manges" second-person sg.) are always fully specified, that is, [+finite], and can therefore not be inserted into [-finite] environments. The MSIH predicts an asymmetry of error patterns: non-finite forms will be substituted for finite forms in early L2 acquisition, but finite forms will never replace non-finite forms. This prediction is borne out in data from L2 learners of French and German.

Today the most widely accepted mapping model is Lardiere's Feature Reassignment Model (Lardiere, 2007, 2008, 2009). Like the MSIH, Feature Reassignment builds on Minimalist conceptions of the lexicon as the locus where lexical items are stored with feature matrices, such as [number] [gender] [tense]. When such matrices differ for equivalent items across the L1 and the L2, the L2 learner lexicon will initially specify only the feature matrices that already exist for the equivalent L1 items. For example, going from L1 English to L2 French, the learner's lexicon will initially give the same set of features for the L2 definite article (modeled on English "the", thus lacking gender and number). Upon noticing multiple forms in the French input ("le" masc. sg. "la" fem. sg. "les" plural), the L1 feature set has to be dismantled and new feature matrices for these entries must be reassigned by adding [gender] and [number].

As we have seen, most grammar-based models critically assume full L1 transfer at the initial state of L2 acquisition, thereby explaining the common occurrence of L2 errors that resemble L1 outputs at lower levels of proficiency. The models diverge with regard to the resolution of L2 errors, depending on whether or not they adopt Full Transfer/Full Access (FTFA, Schwartz & Sprouse, 1996). In FT/FA L2 input that contradicts or conflicts with L1 constraints eventually triggers reanalysis. Since L2 grammars have full access to Universal Grammar, restructuring of the developing L2 grammar and eventual resolution of errors is possible. Thus, the FTFA offers a grammar-based account of two aspects of L2 acquisition: both initial divergence from and development toward target-like language use are explained via the representation of the interlanguage.

Not all grammar-based models that resort to transfer allow for resolution and some incorporate a critical period component (e.g., Johnson & Newport, 1989; but see Birdsong & Molis, 2001). The Interpretability Hypothesis, adopted in Hawkins and Liszka (2003), for example, crucially assumes that L2 learners, even at advanced stages of proficiency, cannot recover from negative L1 transfer in order to account for persistent errors in the production of inflections.

## Performance-based models of L2 development

In contrast to grammar-based models, performance accounts (e.g., Carroll 2001; Epstein, Flynn, & Martohardjono, 1996; Klein, 2004; Klein & Martohardjono, 1999) look to domains outside grammar and representation, such as processing or input factors, as significant sources of non-target patterns and L2 development. From a psycholinguistic perspective, such accounts are interesting because they lend themselves more directly to methodologies suited to parsing and online processing. These methodologies include self-paced reading, speeded grammaticality judgment, eye-tracking, and event-related potential (ERP) tasks. In this section we describe studies that use one or more of these methodologies.

Generative SLA research into L2 processing began to emerge in the late 1990s. Many studies at the time focused on L2 acquisition of *wh*-questions and movement constraints and noted lowered accuracy rates in L2 learners' detection of *ungrammatical* sentences, such as "*What did John hear the news that Mary did?"[1] Largely in response to claims that failure of detection was due to the L2 grammar, White and Juffs (1998) suggest that a processing-based account might be better suited to explain this phenomenon. In particular, they noted an asymmetry in detection between *grammatical* sentences, such as 1a. and b. below. In a judgment task, Chinese L2 learners of English performed better on object extractions (1a) than on subject extractions (1b).

(1) a. What does Mary believe John teaches ——? (Object extraction)
    b. Who does Mary believe —— teaches linguistics? (Subject extraction)

This asymmetry cannot be ascribed to the lack of movement constraints in the grammar, since grammatical sentences are by definition not affected by such constraints. White and Juffs were therefore led to conclude that a processing-based explanation might be more appropriate.

Using a moving windows task, Juffs and Harrington (1995) further demonstrated that L2 learners showed a dramatic increase in reading time at the verb "teaches" in 1b, just after the location of the subject gap. Thus, they had evidence for a parsing deficit (performance), rather than a grammar or representation deficit (competence) for lower accuracy on subject extractions compared to object extractions. The implication in general was that parsing could also provide a viable explanation for lower rates of accuracy on ungrammatical sentences.

More recent psycholinguistic studies, Hopp (2010, 2013) investigated knowledge of German inflectional markers (case inflection, subject-verb agreement) by advanced and near-native speakers who had learned German post-Critical Period.[2] Hopp found that the L2 speakers were indistinguishable from native speakers in an off-line grammaticality judgment task, but slightly different in on-line tasks that were more sensitive to processing. In particular, in a self-paced reading task of grammatical sentences differing in word order, such as 2a and 2b below, native and near-native groups showed robust slowdowns at critical points during incremental processing of case and agreement markers.

(2) a. Er denkt, dass **der** Hotelier im August **den** Gastwirt angezeigt hat. (Subject-Object)
    He thinks that the-NOM hotel owner in August the-ACC landlord sued has
    "He thinks that the hotel owner sued the landlord in August"
    b. Er denkt, dass **den** Hotelier im August **der** Gastwirt angezeigt hat. (Object-Subject)
    He thinks that the-ACC hotel owner in August the-NOM landlord sued has.
    "He thinks that the landlord sued the hotel owner in August"

Advanced L2 speakers also displayed a slowdown, though not as great as the other two groups. A speeded grammaticality judgment task yielded similar results, with near-natives and native speakers patterning together and advanced speakers trailing in accuracy rates. When processing load was increased for native speakers by speeding up the presentation of sentences, accuracy was reduced to below chance for some constructions. Interestingly, the error patterns made by native speakers at the highest speed of presentation paralleled those shown by the non-native groups at lower speeds. Hopp suggests that these results point to limitations in processing, rather than grammatical deficits in non-native speakers.

Another area of research of relevance are ERP studies (event-related potentials) that measure electrical flow at the surface of the scalp during language processing. Because ERP methodology is non-invasive and relatively cost-efficient (compared to MEG and fMRI), it has been used for some time to investigate language processing and in the past ten years increasingly so for L2 learners and bilinguals. The first studies into later-learned languages showed significant correlations with age of acquisition (e.g., Weber-Fox & Neville, 1996). Recently, however, there has been a shift away from the position that L1 and L2 neural processes are fundamentally different, with newer evidence pointing to the crucial role played by proficiency and use (Perani & Abutalebi, 2005, Kotz, 2009; Kotz, Holcombe, & Osterhout, 2008.)

In the domain of syntax, ERP studies on early and late learners (e.g., Kessler, Martohardjono, & Shafer, 2004) showed that the same ERP component (Late Positivity or P600) is evoked in learners and native speakers when presented with word order violations *(*John not is eating pizza*). Differences in the two groups were primarily found in the amplitude and latencies of the responses, indicating quantitative, rather than qualitative differences (see also Hahne, 2001). Inflectional errors, such as **John is eat pizza,* however, evoked a P600 only in a small subgroup of early learners at the highest proficiency levels. In a study of L2 German, Hahne, Müller, and Clahsen (2006), on the other hand, show similar responses in native and non-native speakers to inflectional errors (overapplication of the plural –s rule). Importantly, proficiency seems to play a major role in rendering ERP components evoked in L2 learners more similar to those evoked in native speakers. In a longitudinal learning study investigating the acquisition of morphosyntactic markers in L2 French, Osterhout *et al.* (2008) look at agreement violations in French, as exemplified in (3), where agreement mismatches evoke a P600 in native speakers.

(3) Tu adores*adorez le français.
'you-2-sg adore-2-sg\adore-2-pl the French
"You love French"

In classroom-instructed L2 learners, they found an interesting progression: while the same agreement violations evoked no responses at beginning stages of learning, an N400 response (indicating lexical and semantic difficulty, rather than the syntactic reintegration signaled by the P600) did emerge after 2 months of instruction. Finally, after one year of instruction, the same response as in L1 speakers, namely a P600, was evoked in the L2 learners. With this longitudinal design, Osterhout *et al.*'s study was able to test the effect of experience and proficiency on ERP components directly (unlike the studies by Kessler *et al.* and Hahne *et al.*), suggesting that the critical factor in evoking L1-like responses may be prolonged use.

Even though the focus of performance-based studies is primarily on processing mechanisms, there is nonetheless debate around the degree to which grammar or representation contributes to L1/L2 differences. On the one hand, there are claims that L2 processing is hampered by limited access to the developing grammar, in a way that L1 processing is not. This position was first articulated in Clahsen and Felser (2006) as the Shallow Structure Hypothesis (SSH), based on a summary of findings in the domain of morphology and morphosyntax. Built on the hypothesized distinction between L1 and L2 processing proposed by Ullman (2001), the SSH posits a fundamental difference between L1 child learners and L2 adult learners. Specifically, it claims that children learning their L1 use the same parsing mechanisms as adult native speakers, and are hindered only by extra-grammatical factors, such as limited working memory capacity and lexical access difficulties. Adult L2 learners, in contrast, are argued to use fundamentally different parsing strategies that rely more on meaning (semantic) than on grammatical (syntactic) information. The SSH is not without its problems and has engendered much debate (Crago & Goswami, 2006). Nonetheless, work in this paradigm continues. For example, focusing on two morphological processes, past tense –ed in a masked-priming task and plural inflection inside compounds in an eye-tracking study, Clahsen, Balkhair, Shutter, and Cunnings (2012) find that proficient learners of L2 English (L1 Arabic) differ from native speakers in sensitivity to inflectional decomposition and conclude that grammatical analysis is used less by the L2 parsing mechanism than by the L1 parsing mechanism. Furthermore, the broad claims of the Shallow Structure Hypothesis have been extended to other areas of grammar (e.g., Felser & Roberts, 2007 (wh-dependencies); Felser *et al.*, 2003 (sentence-ambiguity), Felser, Sato, and Bertenshaw, 2009 (Binding Principle A); Marinis *et al.*, 2005 (wh-dependencies); Papadopoulou and Clahsen, 2003 (relative clause attachment).

In contrast to the SSH, another set of studies maintains that L2 processing relies on the same structurally-guided procedures as L1 processing (e.g., Dekydstpotter & Miller, 2013; Juffs, 2005; Hopp, 2007; Lopez-Prego & Gabriele, 2014; Miller, 2011; Williams, 2006) and that L1/L2 differences are the result of non-syntactic factors such as processing capacity, proficiency, and lexical access. This position has been named the *computational* approach to L2 processing, relegating differences between native and non-native speakers to resources pertaining to the computation, rather than the representation, of language.

In a study investigating morphosyntactic markers of gender and number in L2 Spanish, Lopez-Prego and Gabriele (2014) use untimed and speeded grammaticality judgment tasks to test whether L2 learners are able to detect agreement errors involving both gender (4) and number (5) mismatches:

(4) * *Juan dijo que vio un colegio que era antigua en Londres.*
Juan said that he saw a school-M-Sg that was old-Fem-Sg in London

(5) * *Juan dijo que vio un colegio que era antiguos en Londres.*
Juan said that he saw a school-M-Sg that was old-M-Pl in London

As expected, accuracy rates varied across proficiency levels (low, intermediate, advanced) but all three groups patterned together and performed better in detecting number errors than in detecting gender errors. This was attributed to a difference between the L1 (English) and the L2 (Spanish), since English lacks gender marking, thus going back to the issue of L1 transfer.[3] Of particular interest are the results obtained from the native speaker group who received the same sentences under three speeded conditions. In the fastest speed condition, that is, when processing burden increased significantly, native speakers showed higher acceptance rates of number errors (i.e., failure of detection), similar to those seen in the learners. That is, when native speakers' processing burden is increased to the limit, they begin to behave like non-native speakers. As this can clearly not be attributed to a deficit in grammatical representation, it is likely that the learners' higher acceptance rates of mismatches is also the result of "stretched working-memory resources" (Dekydstpotter & Renaud, 2014, p. 150), rather than representational deficits.

Similar conclusions are drawn by Hopp (2012), who looks at L1 English late learners of L2 German at the steady state, namely when they have reached near-nativeness.[4] In this study, again, the critical variable is feature realization in the L1 English, which unlike the L2 German, does not have gender morphology. Thus, the primary research question is whether late learners can reach target-like performance rates in the L2, even in domains of language that are not instantiated in the L1. Using a visual-world eye-tracking paradigm, Hopp tests whether adult L2 learners, like native speakers, use gender features in the determiner (masculine *der*; feminine *die*; neuter *das*) predictively to identify nouns (e.g., *Topf*, "pot", masculine; *Tasse*, "cup", feminine; *Glass*, "glass", neuter). He finds that a subgroup of L2 learners who consistently assign target gender in production also process predictive syntactic gender agreement like native speakers, thus finding a link between perception and production.

The contribution that these studies make to our understanding of performance differences in L2 speakers is that they investigate adult language learning in the limit. It is in some sense a trivial observation that L2 learners become more target-like as they advance in proficiency. What is interesting is to discover those areas and abilities where adult L2 learners continue to diverge from native speakers, and thus look more closely at very advanced learners and near-native speakers. This is also the area of research that bridges second language learning and bilingualism.

A case in point are the studies by Sorace and her colleagues (Sorace, 2011) showing that discrepancies between native and near-native grammars arise primarily in the areas of language where the grammar interfaces with pragmatics, such as in the interpretation of pronouns involving topic shift. This occurs, for example, in the use of overt pronouns in null pronoun languages, such as Italian and Greek. In these language, the overt pronoun often signals a topic shift. Sorace terms this the Interface Hypothesis and argues that differential treatment of the same linguistic phenomena by monolinguals and bilinguals derive from higher processing costs for the latter.

## Summary

In attempting to explain one the most pervasive and persistent non-native patterns, namely those of morphosyntax in L2 acquisition, generative SLA theory has offered up a variety of hypotheses and models. As we have seen, these range from the purely syntactic to the predominantly processing-oriented. The question whether adult L2 learners can master L2 grammars in a target-like manner has largely been answered in the positive by studies focusing on post-critical period near-natives and bilinguals (Hopp, 2010, 2012; Sorace, 2011). What remains to be resolved is to what degree extra-grammatical factors related to processing the input determine the outcome of L2 learning. The studies described in this section have taken us forward in our understanding of these matters, by closely inspecting performance mechanisms involved in parsing second languages. Whether such mechanisms are fundamentally different (Ullman, 2001; Clahsen & Felser, 2006), or fundamentally the same (Dekydstpotter & Renaud, 2014) in L1 and L2 remains to be seen, and continues to be the focus of psycholinguistic research in SLA.

## NOTES

1 This sentence questions a word inside a noun-complement "the news that…" and constitutes a subjacency violation, prohibited by Universal Grammar.
2 Language proficiency of native and non-native participants was measured by two tasks, i.e. a cloze test and a speech elicitation task. See Hopp, 2010, p. 909.
3 We note that this might also be the result of perceptibility, considering the fact that the number mismatches necessarily involved an additional phoneme/grapheme (*colegio*/*antigu**os***) while the gender mismatches involved vowel substitutions (*colegi**o***/*antigu**a***).
4 Language proficiency was measured with a standardized 30-item placement test used by the Goethe Institut.

## REFERENCES


Beck, M.-L. (1997). Regular verbs, past tense and frequency: Tracking down a potential source of NS/NNS competence differences. *Second Language Research*, 13(2), 93–115. doi: 10.1191/026765897670780840

Bhatia, T. & Ritchie, W. (1999). The bilingual child: Some issues and perspectives. In T. Bhatia & W. Ritchie (Eds.), *Handbook of Child Language Acquisition* (pp. 569–643). San Diego, CA: Academic Press.

Bialystok, E. & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Erlbaum, Mahwah, NJ.

Birdsong, D. & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44(2), 235–249. doi: 10.1006/jmla.2000.2750

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis* 20, 3–49.

Broselow, E. (2004). Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism*, 8(1), 51–65. doi: 10.1177/13670069040080010401

Broselow, E., Chen, S. -I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, 20(2), 261–280. doi: 10.1017/S0272263198002071

Carroll, S. (2001). *Input and Evidence: The Raw Material of Second Language Acquisition*. Amsterdam: John Benjamins. doi: 10.1075/lald.25

Chondrogianni, V. & Marinis, T. (2012). Production and processing asymmetries in the acquisition of tense morphology by sequential bilingual children. *Bilingualism: Language and Cognition*, 15(1), 5–21. doi: 10.1017/S1366728911000368

Clahsen, H., Balkhair, L., Schutter, J-S., & Cunnings, I. (2012). The time course of morphological processing in a second language. *Second Language Research*, 29(1), 7–31. doi: 10.1177/0267658312464970

Clahsen, H. & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42. doi: 10.1017/S0142716406060024

Crago, M. & U. Goswami, eds. (2006). Special Issue of *Applied Psycholinguistics*, *27*(1).

Dekydstpotter, L. & Miller, A.K. (2013). Inhibitive and facilitative priming induced by traces in the processing of wh-dependencies in a second language, *Second Language Research*, 29(3), 345–372. doi: 10.1177/0267658312467030

Dekydstpotter, L. & Renaud, C. (2014). On second language processing and grammatical development: The parser in second language acquisition. *Linguistic Approaches to Bilingualism*, 4(2), 131–165. doi: 10.1075/lab.4.2.01dek

Dulay, H. & Burt, M. (1973). Should we teach children syntax? *Language Learning*, 23, 245–258. doi: 10.1111/j.1467-1770.1973.tb00659.x

Dulay, H., Burt, M., & Krashen, S. (1982). *Language Two*. Oxford: Oxford University Press.

Epstein, S., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Brain and Behavioral Sciences*, 19(4), 677–714. doi: 10.1017/S0140525X00043521

Felser, C. & Roberts, L (2007). Processing *wh*-dependencies in a second language: A cross-modal priming study. *Second Language Research*, 31, 9–36. doi: 10.1177/0267658307071600

Felser, C., Roberts, L., Gross, R., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, 24, 453–489. doi: 10.1017/S0142716403000237

Felser, C., Sato, M., & Bertenshaw, N. (2009). The on-line application of Binding Principle A in English as a second language. *Bilingualism: Language and Cognition*, 12, 485–502. doi: 10.1017/S1366728909990228

Flynn, S. & Martohardjono, G. (1994). Mapping from the initial state to the final state: The separation of universal and language specific properties. In B. Lust, M. Suñer, & J. Whitman (Eds.), *Syntactic Theory and First Language Acquisition: Cross-Linguistic Perspectives, Vol. 1: Heads, Projections, and Learnability* (pp. 319–336). Hillsdale: Lawrence Erlbaum Associates.

Foster-Cohen, S. (2001). First language acquisition…second language acquisition: What's Hecuba to him and he to Hecuba? *Second Language Research*, 17, 329–344. doi: 10.1191/026765801681495859

Franceschina, F. (2005). *Fossilized Second Language Grammars: The Acquisition of Grammatical Gender*. Amsterdam: John Benjamins. doi: 10.1075/lald.38

Genesee, F., Paradis, J., & Crago, M. B. (2004). *Dual Language Development and Disorders: A Handbook on Bilingualism and Second Language Learning*. Baltimore, MD: Paul Brookes.

Goad, H. & White, L. (2005). Representational 'deficits' in L2: Syntactic or phonological? In A. Brugos, M.R. Clark-Cotton, & H. Seungwan (Eds.), *Proceedings of the 29th annual Boston University Conference on Language Development, Vol. 1* (pp. 216–227). Somerville: Cascadilla Press.

Goad, H., White, L., & Steele, J. (2003). Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics*, 48(3/4), 243–263. doi: 10.1353/cjl.2004.0027

Grondin, N. & White, L. (1996). Functional categories in child L2 acquisition of French. *Language Acquisition*, 5, 1–34. doi: 10.1207/s15327817la0501_1

Hahne, A. (2001). What's different in second-language processing?: Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, 30(3), 251–266. doi: 10.1023/A:1010490917575

Hahne, A. Muller, J. L. & Clahsen, H. (2006). Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience*, 18(1), 121–134. doi: 10.1162/089892906775250067

Hawkins, R. (2005). Revisiting Wh-movement: The availability of an uninterpretable [wh] feature in interlanguage grammars. In L. Dekydtspotter *et al.* (Eds.), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)* (pp. 124–137). Somerville, MA: Cascadilla Proceedings Project.

Hawkins, R. & Chan, C. Y. (1997). The partial availability of Universal Grammar in second language acquisition: The "failed functional features hypothesis." *Second languge Research*, 13, 187–226.

Hawkins, R. & Hattori, H. (2006). Interpretation of English multiple *wh*-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research*, 22, 269–301. doi: 10.1191/0267658306sr269oa

Hawkins, R. & Liszka, S. (2003). Locating the source of defective past tense marking in advanced L2 English speakers. In R. van Hout, A. Hulk, F. Kuiken, & R.J. Towell (Eds.), *The lexicon-syntax interface in second language acquisition* (pp. 21–44). Amsterdam: John Benjamins. doi: 10.1075/lald.30.03haw

Haznedar, B. (2001). The acquisition of the IP system in child L2 English. *Studies of Second Language Acquisition*, 23, 1–39. doi: 10.1017/S0272263101001012

Haznedar, B. & Schwartz, B. (1997). Are there optional infinitives in child L2 acquisition? In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development* (pp. 257–268). Somerville, MA: Cascadilla Press.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, 120(4), 901–931. doi: 10.1016/j.lingua.2009.06.004

Hopp, H. (2012). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1) 33–56. doi: 10.1177/0267658312461803

Hopp, H. (2013). The development of L2 morphology. *Second Language Research*, 29(1), 3–6. doi: 10.1177/0267658312465304

Ionin, T. & Wexler, K. (2002). Why is 'is' easier than '-s'?: acquisition of tense/agreement morphology of child second language learners of English. *Second Language Research*, 18, 95–136. doi: 10.1191/0267658302sr195oa

Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99. doi: 10.1016/0010-0285(89)90003-0

Juffs, A. (2005). The influence of first language on the processing of *wh*-movement in English as a second language. *Second Language Research*, 21(2), 121–151. doi: 10.1191/0267658305sr255oa

Juffs, A. & Harrington, M. (1995). Parsing effects in second language sentence processing. *Studies in Second Language Acquisition*, 17(4), 483–516. doi: 10.1017/S027226310001442X

Kessler, K., Martohardjono, G., & Shafer, V. (2004). ERP correlates of age and proficiency in L2 processing of syntactic and inflectional information. In A. Brugos, L. Micciulla, & C.E. Smith (Eds.), *Proceedings of the 28th annual Boston University Conference on Language Development*, *Vol. 1* (pp. 294–305). Somerville: Cascadilla Press.

Klein, E. C. (2004). Beyond syntax: Performance factors in L2 behavior. In B. VanPatten, J. Williams, S. Rott, & M. Overstreet (Eds.), *Form-Meaning Connections in Second Language Acquisition* (pp. 155–177). Mahwah: Lawrence Erlbaum Associates.

Klein, E. C. & Martohardjono, G. (1999). Investigating second language grammars: Some conceptual and methodological issues in generative SLA research. In E. C. Klein & G. Martohardjono (Eds.), *The Development of Second Language Grammars: A Generative Perspective* (pp. 3–36). Philadelphia: John Benjamins. doi: 10.1075/lald.18

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain & Language*, 109(2–3), 68–74. doi: 10.1016/j.bandl.2008.06.002

Kotz, S. A., Holcomb, P. J., & Osterhout, L. (2008). ERPs reveal comparable syntactic sentence processing in native and non-native readers of English. *Acta Psychologica*, 128(3), 514–527. doi: 10.1016/j.actpsy.2007.10.003

Kwon, E. Y. & Han, Z-H. (2008). Language transfer in child SLA: A longitudinal case study of a sequential bilingual. In J. Philp, R. Oliver & A. Mackey (Eds.), *Second Language Acquisition and the Younger Learner: Child's play?* (pp. 303–332). Amsterdam: John Benjamins. doi: 10.1075/lllt.23.18kwo

Lakshmanan, U. (2009). Child second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The New Handbook of Second Language Acquisition*, (pp. 377–401). UK: Emerald Press. doi: 10.1017/S0272263109990076

Lakshmanan, U. & Selinker, L. (2001). Analyzing interlanguage. How do we know what learners know? *Second Language Research*, 17, 393–420. doi: 10.1177/026765830101700406

Lardiere, D. (1998). Case and tense in the 'fossilized' steady state. *Second Language Research*, 14(1), 1–26. doi: 10.1191/026765898674105303

Lardiere, D. (2007). *Ultimate Attainment in Second Language Acquisition: A Case Study*. New York: Routledge.

Lardiere, D. (2008). Feature-assembly in second language acquisition. In J. Liceras, H. Zobl, & H. Goodluck (Eds.), *The Role of Formal Features in Second Language Acquisition* (pp. 106–140). New York: Lawrence Erlbaum Associates.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, 25(2), 173–227. doi: 10.1177/0267658308100283

Lee, A. Y. (2008). Child second language acquisition: Acquisition of functional categories – a case study. Doctoral dissertation, Universität Wien.

Lopez Prego, B. & A. Gabriele. (2014). Examining the impact of task demands on morphological variability in native and non-native Spanish. *Linguistic Approaches to Bilingualism*, 4, 192–221. doi: 10.1075/lab.4.2.03lop

Lust, B. (1994). Functional projection of CP and phrase structure parameterization: An argument for the Strong Continuity Hypothesis. In B. Lust, M. Suñer, & J. Whitman (Eds.), *Syntactic Theory and First Language Acquisition: Cross-Linguistic Perspectives, Vol. 1: Heads, Projections, and Learnability* (pp. 85–118). Hillsdale: Lawrence Erlbaum Associates.

Marinis T., Roberts, L., Felser, C., & Clahsen, H. (2005) Gaps in second language sentence processing. *Studies in*

*Second Language Acquisition*, 27, 53–78. doi: 10.1017/S0272263105050035
Marinis, T. (2008). On-line processing of sentences involving reflexive and non-reflexive pronouns in L1 and L2 children. In A. Gavarro Alguero & M. J. Freitas (Eds.), *Language Acquisition and Development*. Proceedings of GALA 2007 (pp. 348–358). Cambridge Scholars Publishing.
Marinis, T. & Chondrogianni, V. (2011). Comprehension of reflexives and pronouns by child L2 learners. *Journal of Neurolingistics*, 24, 202–212. doi: 10.1016/j.jneuroling.2010.02.009
Martohardjono, G., Valian, V., & Klein, E. (2011). Missing tense in second language acquisition: Competence or performance? Poster presented at LING50*: 50 Years of Linguistics at* MIT. Massachusetts Institute of Technology, December 9–11, 2011.
McLaughlin, B. (1978). *Second Language Acquisition in Childhood*. Hillsdale, NJ: Erlbaum.
Meisel, J. (2009). Second language acquisition in early childhood. *Zeitschrift fur Sprachwissenschaft*, 28, 5–34. Walter de Gruyter. doi: 10.1515/ZFSW.2009.002
Miller, A. (2014). Child second language acquisition: What do we know? *Teacher's College Columbia University Working papers in TESOL & Applied Linguistics*, 12(1), 18–34.
Miller, A. K. (2011). Rethinking non-native processing constraints: Evidence from L2 French. In L. Plonsky & M. Schierloh (Eds.), *Selected proceedings of the 2009 Second Language Research Forum: Diverse contributions to SLA: Integrating the parts of a greater whole* (pp. 109–120). Somerville, MA: Cascadilla Press.
Montgomery, J. W. & Leonard, L. B. (1998). Real-Time inflectional processing by children with specific language impairment: Effects of phonetic substance. *Journal of Speech Language and Hearing Research*, 41, 1432–1443. doi: 10.1044/jslhr.4106.1432
Montgomery, J. W. & Leonard, L. B. (2006). Effects of acoustic manipulation on the eaal-time inflectional processing of children with specific language impairment. *Journal of Speech Language and Hearing Research*, 49, 1238–1256. doi: 10.1044/1092-4388(2006/089)
Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., French-Mestre, C., & Hirschensohn, J. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics*, 21(6), 509–521. doi: 10.1016/j.jneuroling.2008.01.001
Papadopoulou, D. & Clahsen H. (2003). Parsing strategies in L2 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, 25, 501–528. doi: 10.1017/S0272263103000214
Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with Specific Language Impairment. *Language, Speech and Hearing Services in the Schools*, 36, 172–187. doi: 10.1044/0161-1461(2005/019)
Paradis, J., Genesee, F., & Crago, M. B. (2011). *Dual Language Development & Disorders: A Handbook on Bilingualism & Second Language Learning* (2nd Ed.). Baltimore, MD: Paul H. Brookes Publishing Co, Inc.
Paradis, J., Rice, M. L, Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics*, 29, 689–722. doi: 10.1017/S0142716408080296
Perani, D. & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15, 202–206. doi: 10.1016/j.conb.2005.03.007
Pierce, L. J., Genesee, F., & Paradis, J. (2012). Acquisition of English grammatical morphology by internationally adopted children from China. *Journal of Child Language*, 40, 1076–1090. doi: 10.1017/S0305000912000402

Prévost, P. & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16, 103–133. doi: 10.1191/026765800677556046

Prévost, P. & White, L. (2000a). Accounting for morphological variation in second language acquisition: Truncation or missing inflection? In M.-A. Friedemann, & L. Rizzi (Eds.), *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics* (pp. 202–235). London: Longman.

Prévost, P. & White, L. (2000b). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16(2), 110–133.

Schwartz, B. (2003). Why child L2 acquisition? In J. van Kampen & S. Baauw (Eds.), *Proceedings of* GALA *2003* (Vol. 1, pp. 47–66). Utrecht: Netherlands Graduate School of Linguistics.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and Full Transfer/Full Access. *Second Language Research*, 12(1), 40–72. doi: 10.1177/026765839601200103

Solt, S., Pugach, Y., Klein, E. C., Adams, K., Stoyneshka, I., & Rose, T. (2004). L2 perception and production of the English regular past: Evidence of phonological effects. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *Proceedings of the 28th annual Boston University Conference on Language Acquisition, Vol. 2* (pp. 553–564). Somerville: Cascadilla Press.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1–33. doi: 10.1075/lab.1.1.01sor

Tsimpli, I. M. (2014). Early, late or very late? *Linguistic Approaches to Bilingualism*, 4(3), 283–313. doi: 10.1075/lab.4.3.01tsi

Tsimpli, I. & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, 23(2), 215–242. doi: 10.1177/0267658307076546

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, 2, 717–726. doi: 10.1038/35094573

Unsworth, S. (2013). Assessing age of onset effects in (early) child L2 acquisition. *Annual Review of Applied Linguistics*, 33, 21–50. doi: 10.1017/S0267190513000044

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., & Tsimpli, I. (2014). The age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics*, 1–41. doi: 10.1017/s0142716412000574

Vainikka, A. & Young-Scholten, M. (1996). The early stages in adult L2 syntax:additional evidence from Romance speakers. *Second Language Research*, 12, 140–176. doi: 10.1177/026765839601200202

Vainikka, A. & Young-Scholten, M. (1998). Morphosyntactic triggers in adult SLA. In M.-L. Beck (Ed.), *Morphology and Its Interfaces in Second Language Knowledge* (pp. 89–114). Amsterdam: John Benjamins. doi: 10.1075/lald.19.06vai

Weber-Fox, C. M. & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–256. doi: 10.1162/jocn.1996.8.3.231

White, L. & Juffs, A. (1998). Constraints on wh-movement in two different contexts of non-native language acquisition: Competence and processing. In S. Flynn, G. Martohardjono, & W. O'Neil (Eds.), *The Generative Study of Second Language Acquisition* (pp. 111–129). Mahwah: Lawrence Erlbaum Associates.

Williams, J. N. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition*, 9(1), 71–88. doi: 10.1017/S1366728905002385

Zobl, H. & Liceras, J. (1994). Functional categories and acquisition orders. *Language Learning*, 44, 159–180. doi: 10.1111/j.1467-1770.1994.tb01452.x

# 30 Signed Language Acquisition: Input

JUDY KEGL

University of Southern Maine

## Introduction

People often ask in what ways signed languages differ from spoken languages. For the most part, the answer is that they do not. Signed languages are dialects of the same human language instantiated by spoken languages. In the most superficial sense they are distinguished by the modalities in which they are expressed and received: visual and gestural (face, torso, arms, hands and eyes as opposed to vocal articulators and ears). There are implicational universals[1] following from the nature of visual processing, the size, movements and degrees of freedom of the articulators, and the recruitment of facial and postural language features that characterize the typological choices made in signed language grammars. For example, if a language is signed in the visual modality, it tends to avoid isolating morphology. Nonetheless, signed languages do not stray from the set of typological options available in the set of human languages most studied to date—spoken languages.

While signed languages are themselves no exception to the class of human languages, there is one area in which every signed language in the world distinguishes itself from spoken languages. The input to acquisition is mostly from non-native signers. This chapter focuses on input to the acquisition of *primary* signed languages like American Sign Language, French Sign Language or Nicaraguan Sign Language. These signed languages are naturally occurring and are used by large populations of signers who are deaf and unable to fully access co-existing languages in the auditory modality. I will not address *secondary signed languages*. These communication systems are secondary to spoken languages and do not fall into the class of naturally occurring human languages. They do not exhibit the full range of typological and universal grammatical options shared by primary signed languages. Such languages include Warlpiri Sign Language (related to spoken

*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.

Warlpiri, although with a more restricted word order (Kendon, 1988)) and Plains Indian Sign Language (a trade language that served as an interlanguage between mutually unintelligible Native American languages (Kegl & White Eagle, 1986)), They may also include manually-coded forms of spoken languages that have been developed for educational purposes.

This chapter provides references to articles that review some findings to date regarding signed language acquisition and then focuses upon the unique challenges deaf language learners face in gleaning linguistic information from an extremely heterogeneous and widely varying amount of input that differs greatly child to child.

## Unique input

Imagine the possibility of being born into a family that speaks a different language from the one you are pre-destined to acquire. We can imagine an immigrant situation in which the heritage language is withheld from a child with the mistaken goal of better guiding that child in the direction of acquiring the dominant language of the society in which they now live. Some immigrant parents today struggle to speak only the local language, in which they are not very proficient, when communicating directly with their children. Such linguistic decisions were commonplace until the 1950s and early 1960s, prior to the groundbreaking work of Lambert and his colleagues (Peal & Lambert, 1962, inter alia) with French-English children in Canada, indicating cognitive advantages associated with growing up bilingual. These hearing children, deprived of their heritage language, do have the ability to overhear conversations in the heritage language between adults and also have access via the community and media sources to native input in the dominant language of the society. These children typically become proficient in the local language, but may only become passive bilinguals, partial users, or non-speakers of the heritage language. For a more in depth discussion of heritage languages and incomplete bilingualism, see Meisel's chapter in this volume.

Now imagine instead that these child learners are deaf. They physically cannot access their hearing family's spoken heritage language as input to the acquisition process. They are also cut off from auditory input relevant to the dominant language in their society. The only language input they are able to access fully is input from a visual language—a signed language. However, such input is not readily available in the home. Like the immigrant family, the parents might attempt to learn a signed language and struggle to use it at home, perhaps sporadically when trying to directly communicate with their child. They might use it haphazardly in conjunction with simultaneously speaking, or even use a makeshift home gesturing system or fingerspelling. In the typical case, the parents do not learn a signed language at all and the child gets its first exposure to any language in school. This late exposure to a first language outside the home is not atypical.

Children who can hear typically learn their heritage language(s) at home from parents and family members who speak the language(s) natively. However, only

6–10% of deaf children are born into families with deaf parents. (Mitchell & Karchmer, 2004). The stable 6–10% of deaf children with deaf parents is characteristic of North American and European populations.[2] It is even lower in some other areas, like Nicaragua for example.

Deaf children with signing Deaf parents are exposed to their heritage language directly from their parents in the home from birth. But, even here, variation abounds. Some of these deaf parents are themselves from hearing families and are not natively fluent in ASL, others use a variety of signed communication ranging from manually coded versions of spoken languages to a full blown signed language. Once in school the variations in exposure continue depending upon the school. These variations include schools following an oral philosophy of lip reading and speech, mainstream schools where access to academics is through interpreters (who may or may not sign natively), residential schools with populations of deaf peers with varying signing backgrounds as well as a variety of signing models ranging from culturally Deaf and native signing Deaf teachers to non-signing teachers who speak and rely upon interpreters.

In summary, all deaf children encounter much more variability in input over their developmental life span than is typically the case with spoken language learners who can hear. A small percentage have acquired a native language base in the home and, even when in school, have reinforcement of fluent signing at home. Others, lacking auditory access to a language spoken at home, first experience input from a signed language in school and do not have reinforcement of their signing at home. Sometimes these children have access to a primary signed language as their input and other times children in schools encounter only forms of manually coded spoken languages or interpretation from non-native signing models.

## The issue of iconicity

The visual character and spatial richness of signed languages has often led people to attribute more of an influence of iconicity to language learning than is warranted. An early study by Roger Brown (1980) found that young, hearing non-signers who were explicitly taught individual iconic versus non-iconic signs were quicker to learn the iconic ones. However, subsequent studies have failed to find evidence for iconicity in either the acquisition of ASL signs by young deaf children acquiring ASL (Orlansky and Bonvillian, 1984) or in the development of ASL grammar. For an alternative view of the iconicity issue, however, see Perniss *et al.* (2010).

Newport and Meier (1987) cite data from Kantor (1982) and Supalla (1982) indicating that the highly morphologically complex verbs of motion that are characteristic of ASL grammar are fully mastered as late as age 9. Late acquisition suggests that these multi-morphemic classifier verbs of motion are not acquired in an analogue or holistic fashion as might be suspected if they were iconic. Rather, they are acquired in a morpheme-by-morpheme fashion. The lack of error free acquisition of morphology, even in forms such as these that appear to be highly iconic, also suggests a lack of attention to iconicity in the acquisition process.

Bernstein (1980) in a replication of Clark (1973) looked at the acquisition of *in*, *on*, and *under* in ASL. If iconicity affected acquisition, we would expect these signs to be acquired earlier in ASL. He found that despite an apparent iconic relation between these ASL verbs and the spatial relations they represent, children still made the same errors found in the Clark study, where children relied upon typical pre-linguistic conceptual relations between objects rather than the linguistically determined relations. In other words, told to put something *under* a table, they would put it *on* the table or told to put something *on* a box, they would put it *in* the box. In Berstein's study some children went so far as to punch a hole in a sealed box to put something *in* when asked to put it *on*.

Similar findings regarding the irrelevance of iconicity appear in Pettito (1987). She identifies shifters in the early acquisition of the pronoun system (e.g., *pick you up* for *pick me up*) despite the obvious iconicity available in the deictic pronominal signs.

## Prior reviews of the literature

While the first studies of signed language acquisition date back to the work of Dutch researcher Bernard Tervoort (1953, 1959), the bulk of the earliest work on signed language acquisition focused almost exclusively on deaf children's acquisition of American Sign Language (ASL). The first review discussed focuses on ASL, but the second brings in a more recent international perspective.

A comprehensive review of the literature and issues relevant to the acquisition of ASL is presented by Newport and Meier (1987). This chapter addresses linguistically salient features of ASL, the overall course of development, characteristics of parent-ese as a form of input, and typical errors.

Newport and Meier (1987) focus primarily on the acquisition of ASL by deaf children of Deaf parents, but also recognize considerable diversity in the deaf population with respect to the time at which they are first exposed to ASL and the consequences of age of exposure on linguistic and cognitive development. They also recognize the abundance of work on the acquisition of ASL morphology and a dearth of studies on the acquisition of ASL syntax. Both of these topics have received increased attention subsequently. Lillo-Martin, for example, has since focused upon the acquisition of pro-drop (Lillo-Martin, 1991) and *wh*-questions in ASL (Lillo-Martin, 2000).

Morgan and Woll (1992) edited a volume on signed language acquisition that reviews more recent acquisition studies and widens the coverage to reflect rapidly expanding work on signed languages other than ASL, including Italian Sign Language, Sign Language of the Netherlands, Brazilian Sign Language, British Sign Language, Nicaraguan Sign Language, and references to many more.

Their introductory chapter succinctly summarizes what we know of the similarities between spoken and signed acquisition. Infants babble in both spoken and signed languages and at close to the same time (Pettito & Marentette, 1991). When early signs are combined, the patterning resembles the morphosyntactic and syntactic patterning in the parents' signing (Chen, 1999). In contrast, with a more holistic adult second-language learning style, children analyze signs at the morphological level (Newport, 1990), picking up grammatical nuances and making generalizations that adult learners miss.

The Morgan and Woll book also compares linguistic frameworks and methodologies brought to bear on the analysis of signed language acquisition. Various programs for linking transcription of signed languages to videos of the raw data have recently enhanced our ability to transcribe and code the complex gestural and facial data that relays ASL grammar (*SignStream*, Neidle & MacLaughlin, 1998; *The Berkeley Transcription System*, Hoiting & Slobin, 1992); and most recently (*ELAN*, Lausberg & Sloetjes, 2009). The existence of viable transcription systems that capture the nuances of ASL grammar while preserving links to the raw data gives signed language researchers powerful tools for analysis and cross-linguistic research comparable to the way the *CHILDES* (MacWhinney, 2000) system revolutionized spoken language acquisition research. New transcription and database systems linked to video have offered signed language acquisition researchers a means of sharing data not only amongst themselves, but with spoken language researchers as well.

## *A spoken language research perspective*

In an afterword to Morgan and Woll's book, Elena Lieven (1992) addresses several critical issues in signed language acquisition research from the perspective of a spoken language researcher. For cross-linguistic comparison, signed language researchers are limited to a subset of the less than 6–10% of the deaf population—those who have at least one natively fluent deaf parent. This is a relatively small pool to draw from and is even smaller when we move internationally to deaf communities with a less stable genetic source of deaf children. Lieven notes, "… many of the questions I raise will have to await a time when a far larger proportion of deaf children, or the hearing children of signing parents, grow up with a sign language as their naturally acquired language." In "dense database studies" that Lieven is conducting of spoken language acquisition with her colleagues in Leipzig and Manchester, researchers are collecting 7–10% of what children say and hear. Even denser studies are underway. In contrast with longitudinal studies in signed languages that collect about one hour of data every three weeks, Lieven and her colleagues sample 1–2% of everything a child produces. Most longitudinal signed language acquisition studies are sampling on average an hour a month and then are coding an even smaller sample within those data. Behrens (2001) reported that in the development of plural marking in German a much denser level of sampling picked up overgeneralization errors that remain unattested in less rich data sets. There is little potential at present for signed acquisition studies to partake in such benefits of dense sampling studies, as the type reported in Behrens.

Lieven goes on to note numerous other discrepancies between research on the acquisition of signed and spoken languages. Children (hearing and deaf) are typically filmed in dyads interacting with an adult. Lieven questions how common this one-on-one interaction is for children across cultures. She questions whether some children's learning may be based more on observation or polyadic interaction, yet this is rarely studied. For deaf children who rely on eye contact for communication, the study of acquisition in such situations could be critical to understanding how they further develop, or even possibly pick up, a first language in school. To date, there are only a few signed language acquisition studies of this phenomenon (Maestas & Moores, 1980; Lieberman *et. al.*, 2014).

Ambient language plays a role in language learning for children who can hear. The cues for attending to ambient communication and interpreting its meaning from what can be accessed in the environment differs between hearing and deaf learners. Again, while hearing children can receive language input passively, deaf children must intentionally visually attend to any signed language input they receive. In addition, visual attention will not serve as a reliable means of access to the ambient spoken language in their environment.

Deaf children are also exposed to the dominant spoken language in their environment to varying degrees. Thus, in adulthood, they exhibit varying degrees of bilingualism in their spoken and signed languages, ranging from balanced bilingualism to only limited or fossilized mastery of the co-occurring spoken language. While Grosjean (1982) would identify all these individuals as bilingual, the nature and degree of their bilingualism varies greatly.

## *The issue of bilingualism*

Beppie van den Bogaerd and Ann Baker (2002) address the question of whether deaf children acquiring the Signed Language of the Netherlands are truly bilingual and come to a mixed conclusion—one that can be extended to many other cases of signed language acquisition as well. Criteria for bilingualism include amount, frequency of use, and communicative necessity (DeHouwer, 2007, 2009). Van den Bogaerd and Baker conclude that while interaction with Deaf primary care givers (in their case Deaf mothers) is biased toward their production of the signed language, there is evidence that deaf children are learning aspects of both the signed and spoken languages in their environment. Nonetheless, this acquisition is by no means synchronized across the two languages. The spoken language lags far behind. By age 3 these Deaf children have acquired some words in the spoken language but are typically using them in combination with signs. There is little evidence of the acquisition of the syntax of the spoken language at this time, in contrast with the syntactic development seen in the signed language. Remember that, in contrast with hearing bilingual learners, deaf children have little to no auditory access to the spoken languages in their environment. In addition, the vocabulary spurt, typical when word combinations appear, is not seen in the spoken language or manually coded spoken language of deaf acquirers at this time.

Most of the children in the van den Bogaerd and Baker study entered pre-school at age two and a half and at that time were exposed by teachers to a manually coded form of the spoken language, in this case Signed Dutch. This exposure and requirement to use Signed Dutch brings communicative necessity into the process as well as input of the dominant language in the environment from an independent source other than the deaf parent. If syntax were the criterion for bilingual development, these deaf acquirers would not yet be bilingual, since they have nothing but a small repertoire of lexical items. Under the broader criteria in Grosjean (1982), they would be bilingual. In some cases of intensive exposure to both the naturally occurring signed language and manually coded versions of the co-occurring spoken language, yet a third communication form can arise that incorporates aspects of both—a contact language between the signed and spoken language that varies in its use from speaker to speaker and context to context. Van den Bogaerd and Baker (2002) point out that substantive bilingualism is critical to the education of Deaf children and should begin in the very earliest stages in Deaf families.

It should be noted that signed language acquisition is not typically studied within the framework of bilingual acquisition, yet monolingual acquisition of a signed language is rare to non-existent. This poses an important issue for future research. See Grosjean (2010) for an overview of the issues.

Many studies of both signed and spoken languages have shown that the first language acquisition process occurs relatively independently from the tangible linguistic input to that process. In the canonical case of a child acquiring the spoken language of its native speaking parents, the input is never pristine. Input is inherently incomplete and noisy. Sentences are fragmented, interrupted or only partially perceivable as the result of competing noise. Comprehension precedes production, but it is not fully present. Early input is compromised by a lack of full comprehension even when the source language utterances are complete and grammatical. The human brain is designed to make viable language generalizations even in the face of noisy or incomplete data.

Language is an extremely resilient human capacity that expresses itself even under adverse conditions (Kegl, 2002; Goldin-Meadow, 2003). Children with limited cognitive capacities or limited intelligence can still acquire language, and in cases where language deficits do exist, they can be characterized in the context of the human linguistic system we know to be shared across all humans (Clahsen, 2008). Children of parents who are not themselves native users of the input language surpass their models and acquire language natively (Singleton & Newport, 2004). Children without auditory access to languages in their environment are still capable of acquiring visual languages when available (Pettito, 2000). Children who are deaf and blind, given adequate input, acquire languages via a tactile modality. While there are no acquisition studies to date, in my own work on the efficacy of communication via the tactile modality with Nathaniel Durlach and Charlotte Reed (Reed *et al.*, 1995), I worked with individuals born both deaf and blind who were educated in schools for the Deaf and their tactile ASL was a fluent as their sighted Deaf peers.

There is no question that input to the language acquisition process is not linked to a single modality. There is also no question that often the input to the language acquisition process can be both incomplete and noisy. The place where signed language acquisition offers unique insight pertains to the question of how much if any *language* input is necessary and what constitutes language-relevant input.

## The canonical (but rarer) case of signed language input

As in the canonical case of spoken language acquisition in the home from native speaking parents, the amount and nature of input in cases of culturally Deaf children learning ASL from their natively fluent parents is over-determined. Language input is rich and ever present in the environment. The trajectory of acquisition, while variable, is what would be expected for hearing children learning similarly structured agglutinative languages with rich agreement and sentential verbs. A rough trajectory of ASL acquisition that was originally compiled from the existing literature by this author in collaboration with Ruth Loew appears below. Much of these milestones are addressed in the Newport and Meier (1987) review. Many of the additional findings on the acquisition of pointing are drawn from dissertations by Loew (1984) and Hoffmeister (1978), as well as their subsequent works.

### *Babbling and gestures*

Within the first nine months, sign language babbling and the first copying of sign-related gross motor gestures of parents occur. Independent gestures (including those which are sometimes described as the first signs) occur at the end of this period.

### *9 months - 1;0*

Non-linguistic pointing to self, other people and objects appears.

### *1;0 - 1;5*

Pointing to people drops out in this period, although pointing to objects is maintained. The first true signs appear at this stage. There is often overgeneralization (e.g., CAR refers to cars and busses).

### *1;6 - 1;11*

Pointing to other people appears. Verbs appear in the lexicon, but there is no productive verb morphology, with only citation forms of verbs used (i.e., no subject or object agreement verbs, no use of classifiers in spatial verbs). There is no use of derivational morphology and consequently no morphological distinction between nouns and verbs. The first two sign utterances appear. In contrast to adult signing, where agreement with spatial locations, for example, is used to mark subject and object on agreement verbs, sign order is used to mark semantic relations.

## *2;0 - 2;5*

Phonology differs greatly from that of adult signers, with regular patterns of reductions of contrast and omissions of phonological features. There appears to be a universal pattern of handshape development, with maximally visually contrasting handshapes (e.g., fist, pointing hand, flat hand) appearing first. There has been less research on location and movement, but it appears that children substitute simple for more complex movements, and often exhibit perseveration in movement. Some research from ASL suggest that sign location within the center of the child's visual field (e.g., signs made on the face and body) is mastered earlier than signs in the periphery (e.g., signs located on the top of the head). Pointing to addressee (YOU) appears at about two years. Some children show evidence of self/addressee reversal errors, or shifters, (e.g., YOU PICK (meaning I PICK). Pointing to third person begins slightly later, and by 2;5, first, second and third person are correctly distinguished. Verbs requiring person agreement begin to be used, but are most often produced in citation form with agreement omitted or as unanalyzed rote forms. There is often over-generalization of the verb inflection rule, with plain verbs inflected, where this is not grammatical in adult ASL. The first morphological distinctions between nouns and verbs occur, but the contrast is made incorrectly.

## *2;6 - 2;11*

Classifiers first appear in the theme position of spatial verbs. However, these verbs appear to be unanalyzed wholes, with no evidence of productive use. These early classifiers often use unmarked or incorrect handshapes. Verbs do not yet show morphological marking of manner (whether through facial expression or altered movement). The first productive use of verb agreement occurs at the beginning of this period. Noun/verb pairs are distinguished but this is frequently in idiosyncratic, non-adult ways. For example, children may mark one member of the pair with a distinctive facial expression, body posture, or speed of movement.

## *3;0 - 3;5*

Inflection of spatial verbs for movement or manner occurs, but children do not yet combine these. Thus, if movement exhibits inflection, manner is signaled separately from the verb. The first correct use of classifiers occurs at this stage. Verb agreement is mastered in sentences where reference is made to objects present in the environment. However, omission of verb agreement with abstract spatial loci continues until well after 3;0. The first correct use of some number and aspect morphemes is found with spatial and person agreeing verbs.

### *3;6 - 3;11*

Lexical compounds are used, but these are articulated without the characteristic phonological pattern (i.e., both parts of the compound are stressed: BLÚE SPÓT as opposed to BLUE^SPÓT for bruise). Spatial and agreement verbs now have both movement and manner, but these are produced sequentially rather than simultaneously, as would be the case in the adult production. Verb agreement begins to be found with abstract loci, but this occurs without coordinated establishment of referents at those loci. Facial marking and syntactic use of *Wh*-questions appears.

### *4;0 - 4;11*

Innovative compounds appear, although they are not adult-like in either phonology or meaning. Overt establishment of loci associated with referents is still absent in the first part of this stage. A moderate degree of control of the use of abstract loci, including their establishment, use and maintenance, is achieved by 4;11. Children still make occasional over-generalizations of verb inflection rules, although agreement with single subjects are correctly marked. The noun-verb distinction is clear, but innovative forms are still seen in addition to correct forms.

### *5;0 - 5;11*

Most morphological processes are used with reasonable skill. However, the most complex polymorphemic forms still cause difficulty.

### *6;0 - adulthood*

Between the ages of 6 and 20 years, there is ongoing development of the narrative genre. While acquisition of most structures has been completed at the sentence level, the application of cohesion markers, use of narrative role, etc. is still developing during this period.

### *8;0 - 8;11*

The use of classifiers and spatial verbs is largely mastered, although some errors on complex forms are still noted.

### *9;0 - 9;11*

Mastery of the productive use of classifiers and spatial verbs is completed. The mastery of complex overlays of multiple grammatical facial expressions (*wh*- and yes/no questions, relative clauses, conditionals, etc.) in sentences is still in progress.

## Spoken language input

The canonical case of signed language input diverges from its spoken language counterparts in two key respects. First, as noted earlier, most Deaf children with Deaf parents, even those in relatively monolingual countries like the United States, are bilingual to some extent in their signed language and the dominant spoken language(s) in their environment. However, while access to a signed language can be complete, access to the spoken language is always partial.

The average reading level for deaf individuals in the U.S. still hovers at or below a fourth-grade level (Paul, 1998). Because of a strong first-language foundation, more natively fluent Deaf signers tend to be among the better readers in the dominant spoken language in their communities (Charrow & Fletcher, 1974). These Deaf language learners may or may not also be able to speak and lipread a spoken language. Lipreading skills vary widely among deaf individuals. Proficiency depends upon a variety of factors including residual hearing, amount of oral training in speech and speechreading, and use of coded systems for speech such as Cued Speech where speech is phonetically coded via a set of handshapes and placements on the face and neck that indicate biphones (semi-syllabic units indicating the pronunciation of consonants and co-occurring vowels, Cornett, 1967).

## Signed language input: Heterogeneity

When Deaf native language learners move into the larger Deaf community and into schools, they encounter a more highly varied range of signed language input, including contact forms of communication between the signed language and more spoken language-based forms of signing (specialized and typically initialized lexical signs developed for educational purposes, invented sign systems based on co-occurring spoken languages (see S. Supalla, 1991)), and widely varying levels of fluency, dialects and idiolects among their peers, many of whom are not native users of the signed language. In addition, depending upon educational choices, Deaf native signers may attend mainstream schools where they access their educational content via interpreters, many of whom exhibit varying levels of signing competency.

While Deaf children of signing Deaf parents do experience a far wider range of signing input and can adapt their signing to a wide variety of interlocutors, they also maintain a stable and native mastery over their heritage signed language. Varied signed language input becomes a bigger factor when the signed language learner does not have native signing input in the home.

## Cases of learners surpassing their input

Many deaf children are born into deaf families where the input to the acquisition process is highly inconsistent. Singleton and Newport (2004) addressed the case of Simon, where parental input was far less than native. Simon's Deaf parents had been schooled using a philosophy of oral education, but did not have the residual hearing to allow them to learn English with anything near native fluency. Both learned ASL after the age of 15, embraced it and were using it as their primary and preferred means of communication with each other and in the home. Simon was exposed to their communication as input. At the time of testing, his parents had been signing for 20 years. Simon himself was being educated in a deaf enclosed classroom in a public school. He was mainstreamed with hearing children for gym, art, and recess. His teachers and deaf peers were not ASL users. His teachers communicated with the other deaf children using a manually coded form of English and Simon's deaf peers were never observed to communicate using ASL. In addition, Simon's parents had a fairly limited social life outside the home. For these reasons, the authors concluded that Simon's primary input was the non-native ASL of his deaf, late-learner parents.

Singleton and Newport (2004) gave Simon (age 7) and his parents a morphological test battery that they had developed to look at the production of verbs of motion and location. As noted in the acquisition trajectory above, this aspect of productive ASL morphology is complex (much like sentential verbs) and requires careful choice of classifiers as well as spatial agreement. Despite clear non-native performance by Simon's parents when compared with native ASL signers (65–83% accuracy on the morphology), Simon's performance surpassed theirs and fell in the range of native signing children of native signing parents. Simon's parents scored even lower on handshape choices that would correlate with classifier choice (37–46%). Simon's performance surpassed them, but his score was still lower than his native signing peers (46–59%). Classifier choice is semantically and lexically determined (also allowing for a wider range of options) rather than morphosyntactic, suggesting that Simon and his parents may have developed an alternate analysis of classifiers than is typical. Nonetheless, evidence from cases such as Simon's demonstrates that native input is not necessary for a learner to induce the grammar of a signed language and natively master it. The complex yet recoverable word-internal morphology[3] of languages like ASL allows young learners to mine the internal structure of signs that are unanalyzable wholes to their parents to induce the grammar of the language in which these lexical items originally resided (Kegl & Schley, 1986).

Even for a child like Simon, exposed at home to only non-native, late-learners of ASL and in school only to manually coded forms of English, the input was sufficient to yield a full-fledged language that in most cases aligned significantly with the grammar of ASL. Deaf children of non-natively signing Deaf parents with the benefit of even more exposure to ASL surpass their parents and fall into the class of native signers. If there is a cut off to sufficient language input to allow for native language acquisition, it is below the input Simon received. However, we do know that there must be a cut off (at least with respect to critical period) since Simon's later-learning parents themselves did not demonstrate native mastery of ASL.

## *The heterogeneous cases of Deaf children with hearing parents whose primary language is spoken*

In search of input that is insufficient to support native language acquisition, we turn to the cases of deaf children with hearing parents. We can quickly discount the cases of Deaf children with hearing parents who are either native signers (children of Deaf parents themselves) or who fluently and consistently use a natural signed language with their children, even if their own signing skills are non-native. Such cases parallel the cases with Deaf children of Deaf parents discussed earlier. Similarly, deaf children with older signing siblings would also typically pattern as native signers.

The remaining group of deaf children of hearing parents present a highly varied population of language users that defies attempts at standardized testing or categorization in terms of their signed as well as spoken/written language competencies. Their language profiles range from no signing at all to native or near native fluency in their signed language. In terms of bilingualism and competency in the spoken language in their environments, these deaf children of hearing parents also vary between no mastery of the spoken language of their community to full mastery. Native or near-native signing fluency within this group is typically the result of an education from an early age in a residential school for Deaf children. The signed and spoken language competence within this group shows no systematic correlations. The least fluent signers can be equally disfluent in their spoken language or may be highly proficient.

## *Factors affecting the profiles of deaf signers who grew up in non-signing families*

A variety of factors affect the language profiles of deaf signers from non-signing families: the nature of the input to which they are exposed (natural signed languages versus artificially designed signing systems based on the spoken language), the amount and regularity of signed input, the nature of the source of the input (fluent Deaf signers, interpreters of varying degrees of fluency, teachers of varying degrees of fluency).

The most fluent Deaf signers from non-signing homes have typically attended residential schools for the Deaf from an early age where they are exposed 24/7 to signing—both in school and in dormitories with other Deaf children. In such cases, even when the schools use artificially designed manual systems based upon spoken language grammar, it is possible for these children to receive substantive natural signed language input from each other. Under such input conditions, and with enrollment at a very young age, these children of non-signing parents have the potential to develop native or near native fluency in a natural signed language.

At a time when residential schools for the Deaf are shrinking and placements in public schools with interpreters are increasing, the presence of natively signing Deaf children of Deaf parents in schools for the Deaf is no longer guaranteed.

Therefore, the exposure to native language input in contexts that are outside of the classrooms using primarily manual codes for spoken language can no longer be assumed, even in residential schools for the Deaf.

If we look back to the acquisition experience of Simon, whose language mastery surpassed that of his Deaf parents (Singleton & Newport, 2004), an alternate pathway to natural signed language acquisition presents itself. Children can analyze the frozen signs in the input of their late learner parents and recover enough data to allow them to infer the grammatical rules that their parents are neither aware of nor employ in their own signing.

S. Supalla (1986, 1991) examined modality constraints specific to signed languages and presented evidence that artificially created systems like Manually Coded English (MCE) violate the perceptual constraints on signed languages, making them inherently unlearnable and therefore unable to function as natural languages. However, MCE borrows lexical items from American Sign Language, provides a source of frozen lexical items that young children can analyze to infer the natural grammar of the language they were borrowed from. Supalla documented ways in which children exposed to MCE unpacked these signs and came up with some novel forms that did fall within the class of possible grammatical options in a natural signed language. Since MCE was learned in an educational context and not in the home or a residential setting, Supalla found evidence of children's attempts to impose more natural grammar on this artificial input, but not full restructuring into a full-fledged new signed language.

MCE, like the English it is based on, is typologically isolating and relies on linear affixation. However, isolating morphology and linear affixation are not compatible with processing constraints in the visual modality. Visual memory and visual discrimination do not parallel auditory memory and auditory discrimination. A natural language evolved to be processed in the auditory modality (English) but converted to a visual form and forced to be processed in the visual modality (Manually Coded English) becomes unprocessable. Without adaptation of the language to visual processing constraints, all users would exhibit processing deficits.

Unlike English, MCE is produced in the visual modality by larger articulators that move more slowly in space. Signs are produced more slowly than spoken words. Nonetheless, signed and spoken languages share the same processing window of 2.5 seconds. Baddeley (1986) identified a 2.5 second phonological loop for spoken languages. Wilson and Emmorey (1986) demonstrated a processing loop of the same length exists for signed languages. Auditory and visual processing also differ in terms of how close each item in a sequence can be before they perceptually fuse (flicker fusion; Poizner & Tallal, 1987). The required interstimulus interval is almost double. Therefore, in a visual language fewer items in sequence can get into short-term memory for linguistic processing.

These factors conspire to lead languages produced in the visual modality to select from a smaller typological set of options available in universal grammar (implicational universals stemming from modality differences). Specifically, in primary signed languages, isolating typology is avoided. Information is packed vertically into signs rather than strung together in a linear fashion.

In many countries, the majority of deaf children attend public schools, either in self-contained classrooms with some classes mainstreamed or completely in mainstream settings. In mainstream classrooms, children get their signed language input primarily from interpreters with second-language fluency in ASL or more often using signing forms based more on spoken language grammar. Native signing educational interpreters who can hear are rare, and those who are deaf are even rarer. Therefore, students in such educational settings are also posed with the processing challenges just mentioned and are less likely to develop native-like proficiency in a signed language.

## *Factors affecting the profiles of deaf signers who grew up in non-signing families*

Factors affecting spoken language competency include the age at which a deaf child lost the ability to hear, the amount of residual hearing retained, the kind of education experienced, and whether or not the deaf child comes to spoken language reading with a native-level foundation in their signed language. While many deaf individuals develop native fluency in the spoken language and others develop very limited fluency, it is still the case that the average deaf student typically graduates with roughly a fourth-grade reading level (Paul, 1998). A fourth-grade reading level or below is best understood as not using syntax to decode reading.

Given the heterogeneity in terms of spoken language competency, it is impossible to norm reading tests or any English language assessments on the wider population of deaf individuals. Language assessment is best done on an individual basis—single case studies.

## *Varying degrees of bilingualism*

The variation seen in signing and spoken language competence among deaf individuals speaks to equal variability in the degree and balance of bilingualism observed. Some individuals have only a fossilized command of the signed language, the spoken language, or both languages. Others are natively fluent in one, the other, or both. Every possible combination of language skills can be seen.

## *Late signed language acquisition*

Children who can hear have input from birth. Only in cases of feral children such as Itard's case of Victor, the wild boy of Averyon (Lane, 1976), or severe cases of deprivation and abuse such as Genie (Curtiss, 1977) can a hearing child's access to language input be delayed for years, even beyond the critical period. In contrast, children who are deaf can be lovingly cared for and nurtured within their hearing families, but can still be cut off from input to the spoken language of their parents and others simply by virtue of their inability to hear and cut off from signed language input simply because of their geographic location and the lack of signing within their families. These are the cases we will examine in the remainder of this chapter.

As has been discussed, the signed language profiles of deaf children in hearing families, are highly variable, ranging from extreme language deprivation in one or all of the languages to which they are exposed to fluent balanced bilingualism or multilingualism in all of their languages. In addition to the kind of input discussed thus far, the time of input is also an important factor in signed language acquisition. Studies of late signed language acquisition (or non-acquisition) most directly address the critical period for language acquisition proposed by Penfield and Roberts (1959) and popularized by Lenneberg (1967). Lenneberg first identified an abrupt endpoint of the critical period in adolescence. Subsequent work by Newport (1988, 1990), Mayberry (1993), and Kegl, Senghas, and Coppola (1999) have identified an earlier point at which critical period facilitation for language acquisition begins to gradually wane (age 5–7>) until the critical period window closes in adolescence. (See Morford & Kegl, 2000, for a review of these studies.) More recently, evidence for an early critical period for perceptual processing of signed languages parallel to Kuhl's work on speech (2005) has appeared in the literature (see, for example, Palmer *et al.*, 2012).

Mayberry and Eichen (1991) and Mayberry (1993) looked at the critical period effects and differences between first language acquisition after childhood, as well as the lasting advantages of learning a signed language in childhood. Mayberry (1993) studied first language acquisition at age 0-3 (native acquisition); age 5-8 (acquisition in childhood); and 9-13 (late acquisition). All her subjects were adults 20 years post acquisition. She found a gradual decline in sentence processing abilities correlated with increasingly later exposure to ASL as a first language. She also looked at late-deafened English users who had learned ASL as a second language after childhood. The late first-language ASL learners and the second language ASL learners had acquired the language at exactly the same age. The second language learners out performed the late-first language learners, demonstrating that exposure to a language within the critical period facilitates the later learning of a second language, even across modalities.

## Language acquisition in the absence of language input

In an effort to home in on how much language is sufficient to yield native language acquisition, we turn our attention to cases where the communication input to children falls short of what would be deemed a full-fledged language. These include on one end pure gestures and on the other end contact gesturing among groups of individuals in command of their own idiosyncratic home gesture systems.

### *Gesture*

In contrast with Armstrong, Stokoe, and Wilcox (1995), I recognize a dichotomous relationship between gesture and language. Gesture is *not* language; nor is it an initial step on an evolutionary trajectory from gesture to language. Gesture does not involve actors and actions coming together in propositions. Language, on the other hand, is not only propositional and compositional in nature; it is hierarchically structured in a characteristically human way. Language and gesture coexist, each with its own unique structure and function.

Gesture and language are distinct communication systems that function in very distinct ways. Gestures tend to be holistic and to refer to an entire event: "come over here," "get out of here," "stay back," "I have a fever," etc. They are closer to an animal call system than a language. Gestures transmit information in a complete and efficient manner. Warning someone of impending danger is best done through gesture where the message is immediate and complete, than through the slowly unfolding, infinitely variable message that we parse as we listen to language input produced through syntactic means.

Exposure to purely gesture or gesticulation (gestures concurrent with speech) as input will not yield language acquisition as an outcome. Gesturers who remain in the home with gesturing as their input will remain language-less. Some of those individuals who only have gesture as their input develop a communication system that varies little from its input—an inventory of holistic gestures that serve their basic needs. In this case, the caregivers in their environment circumvent the need for language using their shared knowledge and filling in the gaps. For example, a caregiver may produce the same single gesture (a wave at the mouth) for "are you hungry?," "have you eaten?," "that's a hamburger," "it's time to eat," and "that is edible." The same single gesture produced by a gesturing isolate may be interpreted by family members as having any of the above meanings and more, depending upon the context. The home gesture is interpreted with varied and complex meanings, but it did not use anything language-like to convey those distinct meanings. The interpretation come from shared knowledge among the interlocutors, not from syntax (Spitz & Kegl, in press).

Since 1985, we have been conducting a population study of deaf individuals in Nicaragua. In the context of that study Dr. Romy Spitz and I have identified over 450 language isolates raised in hearing homes with access only to gesture of the sort described above. A large proportion have grown into adulthood with communication systems of a very limited inventory of single gestures. These data indicate that there is a lower bound on input below which the language acquisition process is not triggered. To assure that the lack of language is the result of the input and not sequelae of independent cognitive factors, Dr. Romy Spitz has tested each isolate with a battery of cognitive tests that complement the communication samples we collect (Spitz & Kegl, 2005).

## *Home sign*

Susan Goldin-Meadow and her colleagues documented a different phenomenon occurring among deaf sign language isolates in some homes with hearing families who have decided not to educate their children in signing schools, but rather to raise them with an oral philosophy of education—the development of home sign systems (Goldin-Meadow & Mylander, 1984, 1990). Marie Coppola has documented cases of homesigning in Nicaragua as well (Coppola, 2002). Home sign seems to occur when there is more of a back and forth between deaf non-signing children and their parents. The children innovate certain elaborated gestures and their parents copy them. In this way, the communication systems evolve beyond simple gesture. One issue with Goldin-Meadow's studies to keep in mind is that while not exposed to a signed language, they are exposed to academic training in spoken languages.

Home signing is not language, but it is more elaborate than the gestures typically used among hearing people. One major difference involves the move from a single gesture per event to more than one gesture per utterance. With the expansion beyond a single gesture, the ordering of gestures begins to appear. The expansion to home sign includes the emergence of spatial descriptions, labeling of referents, pointing combined with action gestures, and labels for objects and actions that co-occur in a single valence referent-action pair.

Coppola and her colleagues as well as others have argued for attributing to some of these structures linguistic labels like subject, verb, pronoun, even relative clause (see Coppola & Newport, 2005; Coppola & So, 2005, 2006). There is agreement among researchers that homesign systems are not in the category of natural human languages, but there is disagreement as to whether certain structural and ordering regularities in a non-language system can be assigned the syntactic status of subject, verb, pronoun, etc. (See Kegl, 2000, 2002, 2008.) For the purposes of argumentation here, regularity of ordering and function (agent versus patient) are not considered sufficient to determine the syntactic status of components of a home sign system in the absence of a fully articulated grammar characteristic of the class of human natural languages.

I work from the premise that full-fledged human languages (even newly created languages such as Idioma de Señas de Nicaragua (ISN)), emerge fully formed from the acquisition process. Home sign systems are not fully-fledged languages. However, some of the phenomena such as sequencing, ordering, and periodicity that appear in home sign systems, may well be the same phenomena that are sufficient to entice *deaf children* exposed to such data to treat them as language input and, with the aid of their innate language expectations, to surpass their non-language input to create language. So, while a homesigner without exposure to other signers would not create language, deaf learners within the critical period for language acquisition exposed only to homesigners might have sufficient input to trigger the language acquisition process.

The elaborated features of home sign are not present in the gestures of the parents of homesigners. There are several reasons. The parents of the homesigners in the studies by Goldin-Meadow and her colleagues were raising their children with an oralist philosophy of education, which generally entails parents being told not to sign with their children. Holding to this mandate typically ends up in a withholding not only of signs, but also gestures to some extent. In addition, parents are speaking to their children while gesturing. McNeill (1992) studied a specific kind of gesture that always accompanies speech, which he terms *gesticulation*. When gesture and speech are produced simultaneously, they combine into one coordinated speech act. The gesticulation that cooccurs with speech is far less rich than gesture on its own. Of course, deaf children exposed to this input are accessing only the gesticulation portion of this total speech act, so their input is even more impoverished.

As was mentioned earlier, the homesigners studied by Goldin-Meadow were in an educational system geared toward teaching them English via lipreading, speech, reading, writing and in some cases Cued Speech. Even though auditory access to English was blocked, they were being taught English in school during the period in which their development of homesigning was being studied. The impact of education in a spoken language on the emergence of homesign has not yet been fully investigated.

Notably, many hearing parents communicate with their deaf children using gesture or gesticulation that does not exhibit any of the more enhanced structural characteristics found in homesign. However, these parents are capable of gesturing much more than they do. The hearing parents we studied on the Atlantic coast of Nicaragua were not told to withhold signing or gestures from their children. Nonetheless, they communicated with their children using speech with gesticulation or simple gestures, just like the parents in Goldin-Meadow's studies. As part of our battery of testing, we presented deaf gestural isolates with two 1.5-minute non-verbal cartoons and ask them to recount what they had seen in gesture. At first contact, these gestured narratives are minimally comprised of a few single gestures over the course of the entire narrative. We gave the same task to their parents, asking them to gesture the story without speaking. Their performance on this narrative recounting task was surprising. Their gestural narratives exhibited quite a bit of complexity and seemed almost driven by the spoken languages they used. They incorporated features of their spoken languages that don't occur in their gesture like nouns, verbs, etc. into their gestural stories.

Overall, the gestural communication that parents of the gestural isolates produced under the voice-off condition was as rich or richer than what we see in homesigners themselves (Morford & Kegl, 2000). These results were similar to what Singleton, Goldin-Meadow, and McNeill (1995) and Goldin-Meadow, McNeill and Singleton (1996) observed when hearing non-signers were given the task of describing short video vignettes without speaking. This makes sense when we consider that pre-existing knowledge of a language (their spoken language) could be driving their gesture.

Yet, ironically, in study after study of the input to deaf children from non-signing hearing parents, we only find limited gesture to be used. In longitudinal studies by Goldin-Meadow and her colleagues, even hearing parents who developed some homesigning skills lagged behind their children. Convincing parents to gesture without also speaking concurrently with their children could have beneficial consequences. The link between the elaborateness of a child's homesign system and the success of subsequent language acquisition has not yet been determined, but studies like Morford, Singleton, and Goldin-Meadow (1995) and Morford & Hänel-Faulhaber (2011) have shown that ASL signs for concepts expressed in a child's homesigning are more easily acquired than ASL signs for concepts not represented in their homesign.

## *Family signed languages*

One stated characteristic of homesigning is that it is idiosyncratic to the deaf child and the child's family and tends not to be transmitted to subsequent generations. Home sign itself may not be transmitted to subsequent generations of deaf signers while retaining the characteristics of home sign, but there is a transmission within families with multiple deaf siblings that may be informative. We know that home sign does not constitute a full-fledged language. But, is it sufficient input to a child to allow that child to develop a full-fledged language? For an answer to this question, we can look to the emergence of *family signed languages*. Family signed languages are distinct from *family sign systems*, which is another term for *home sign*.

In a family signed language, three factors typically co-occur. First, in a hearing family more than one sibling is deaf. Second, the ratio of hearing to deaf family members tends to be closer to half and half. And third, the family unit is isolated from a larger population of signers. Under these conditions, creation of a signed language can occur in the microcosm of the family unit. The emergent language can arise and disappear in the course of a generation, the family signed language can continue to be passed on, or the family members can eventually assimilate into a larger signing community.

For the past 30 years, I have been engaged in a population study in Nicaragua documenting the birth of a new signed language that arose in the late seventies and early eighties when deaf children from isolated settings were brought together into public schools for the first time. This is the first time, spoken or signed, that a language has been able to be documented while it was coming into being. To date we have tested over 3000 deaf individuals across the age span and across the country to capture a snapshot in time of what communication looks like for language isolates and fluent signers alike. In the process of this population study documenting the emergence of Idioma de Señas de Nicaragua (ISN; Nicaraguan Sign Language), we came across several instances of family signed languages in Nicaragua: a farming family with 12 children, six of whom are deaf; a coffee-harvesting family with three deaf members and an extended family with hearing members; and a family of lumber jacks with three deaf members in a family of 10. In all of these families, it is notable that both deaf and hearing family members sign.

Families with multiple deaf children fit the formula for language emergence. It may be that it is in these tiny enclaves of language that we want to look for the barebones expression of an emergent language. In family signed languages, we find the minimal case of individuals who serve as input to successive individuals in the family. For example, in each of these families, one finds an older deaf member with more of a limited home sign system and successively younger deaf members of the family who sign more fluently. In each of these families, the second deaf sibling is communicating in a way that is different in kind and complexity from the older sibling. The learner has surpassed its input. As we move down the birth order the younger deaf siblings are successively more fluent. Younger hearing siblings in these families have a greater signing fluency than is found among siblings of homesigners.

What happened in family signed languages looks like the kind of language emergence progressions we are documenting on a larger, nation-wide scale in Nicaragua (Kegl & McWhorter, 1997; Kegl, 2000, 2002, 2008), but it is happening within a single family. Successive learners in these families (both hearing and deaf) exhibit language features more characteristic of full signed language users as opposed to homesigners: expanded morphology involving reduplication for actions as well as modifications of signs for distance and time; verb agreement, utterances with multiple lexical signs for referents as well as multiple referent arguments in a single utterance, facial modification of signs, grammatically relevant use of eye gaze and body shift, etc.

Several additional behaviors suggest that these family signed languages are actual signed languages. Family signers have name signs for each other—some thing that doesn't happen with homesigners. Furthermore, when these family signers are brought into contact with signers of Nicaraguan Sign Language there is a mutual borrowing of signs and grammatical devices in both directions—something typical of languages in contact and not a signed language in contact with a homesign system.

In Bluefields, a signing community on the Atlantic coast of Nicaragua, the assimilation of isolated gesturers and home signers into the Bluefields community typically involved homesigners learning signs and grammar from fluent ISN signers. Their success depended on exposure within the critical period and upon other pre-cursors in their idiosyncratic homesign systems to grammatical features also used in ISN (see Morford, 1996; Morford & Kegl, 2000).

However, when the three deaf members of the lumberjack family mentioned earlier were brought into Bluefields for an initial contact with this Deaf community, the borrowing pattern was different. If anything, the deaf lumberjack family members were resistant to learning signs that would replace what they already had. Furthermore, Bluefields signers immediately appreciated many of the lexical and grammatical features of the signed language of the deaf lumberjack brothers. They had developed their own unique system for indicating varying extended periods of time and distance. Within a day, these morphological systems were spreading through the Bluefields signers and being generalized to a variety of forms. The lumberjack brothers had also developed their own unique signs for the different monetary bills for *cordobas* based upon the pictures on the bills: a man resting his head on his chin; an indian shooting a bow and arrow, etc. These were actually quickly borrowed into the signing on the Atlantic coast. Their behaviors are more in line with the kind of interchange and borrowing that happens with a foreign signer or signer from a different dialect of ISN joins the community.

These family signed systems have their own unique set of syntactic and morphological rules characteristic of a fully formed language. They are just very tiny languages. In the 1980s when Ken Hale observed the Deaf community in Nicaragua, and read accounts of the emergence of a signed language there, we were looking at whether there may be a critical mass for language emergence as well. At that time, he noted that in in his study of languages throughout Australia he had observed numerous cases of languages that had emerged in groups as small as six individuals. He was doubtful that a critical mass of language users was necessary, or more precisely that that critical mass needed to exceed more than a handful of individuals. As we have encountered isolated families with multiple deaf members, Hale's comments have become increasingly apropos.

The linguistic analysis of these family signed languages is our most recent research focus and awaits further analysis. However, preliminary findings suggest that Ken Hale was right about the non-issue of critical mass. The study of family signed languages brings into focus issues regarding critical proportion, the complexity to be found in the initial expression of the human bioprogram for language even with single gesturer/homesigner input, and the nature of the input that full-blown signed languages recognize as worthy of borrowing. These studies may aid us in putting a lower bound on the quality of linguistic input to the acquisition process.

While critical mass doesn't seem to be a prerequisite to the emergence of a family signed language, critical proportion does seem to be a major factor in the emergence of family signed languages. Critical proportion affects the power dynamics that arise when the deaf population in a family approaches or exceeds 50%. In many of these situations, hearing members of the families also acquired the signed language and as signers contributed to the linguistic make-up of the language. Fluently signing hearing family members, who in these cases also had in their repertoires native mastery of Spanish and possibly other indigenous languages, were a potential source of additional grammar input to the family signed language—influence from knowledge of a prior language. This is the same kind of input that hearing parents of deaf children have the potential to provide, but in the case of deaf gestural isolates in hearing families the proportion does not drive bilingualism in the hearing siblings and thereby language influence from their spoken language is less likely.

Critical proportion is also characteristic of those situations in which a creole emerges in a plantation situation. In a situation where a small minority of substrate speakers (in this case a few slaves speaking one or more African languages) exists in the context of colonization (referred to as a homestead society (*sociéte d'habitation* as opposed to a plantation-type society; Bourdieu, 1977; Singler, 2008), the outcome of language acquisition is not creolization, but rather bilingualism. The minority speaker learns a second-language learner version of the majority language (Mufwene, 1999). Or, in the case of a single deaf child in a family using a majority language that cannot be accessed via auditory means, it is possible that no language acquisition occurs. Creoles arise instead in contexts where the minority population approximates 50% of the overall group and where more than a single language is spoken among the members of the minority group.

So, let's apply the plantation analogy to the hearing family with multiple deaf children. The spoken language in use in the hearing family is like the superstrate (the dominant language of the colonizer). It is used on an as needed basis for basic necessities (an asymmetric pidgin), but not for the full range of topics discussed among family members (a language). The deaf family members are a substantive substrate who share with each other being deaf and not having access to the dominant spoken language. They also share with each other a visual orientation to the world and a desire to communicate with each other on more than an "as needed" basis. But, initially, they don't share a language. They also don't have command of the dominant spoken language or a language of their own. They use the gestures they have, borrow gestures from each other, repeat themselves and try alternate means to get their point across. They establish a visual communication among themselves (a symmetric pidgin) that they use to talk about a much wider range of topics than they would with hearing family members. Once this symmetric pidgin arises, it has certain periodic characteristics that resemble language and younger siblings treat this input as language data. They fill in the gaps with their innate language expectations and a family signed language arises. Once the family signed language exists hearing members of the family (particularly younger siblings) will learn it as well, just as they would any other languages in use within the family.

The existence of family signed languages raises a further question regarding the extent to which family signed languages played a part in the input to the deaf students in Nicaraguan schools in the 1980s. Because of the isolated situations of these families, such influence is unlikely. We do however have evidence of at least one case of assimilation of these family members into the Deaf community after adulthood, and subsequent to the emergence of ISN.

Emergent sign languages like Al-Sayyid Bedouin Sign Language (Sandler *et al.*, 2005) and Providence Island Sign Language (Washabaugh, 1986) are however likely candidates for such influence and worthy of focus in that regard perhaps with comparison to ISN, the signed language that emerged in Nicaragua. There are now enough cases of emergent signed languages available to begin typological studies and to look at the effect of the influence of signing hearing siblings (and signing hearing children) on language emergence.

## A transitional input between home sign and Nicaraguan Sign Language

The case of the emergence of a signed language in Nicaragua in the 1980s is well documented and can be accessed elsewhere (Kegl & McWhorter, 1997; Kegl, 2000, 2002, 2008; Kegl, Senghas, & Coppola, 1999). This case involves a different kind of input from the others discussed thus far, but it parallels the emergence of family signed languages. In Nicaragua, we argue for a transitional input between homesign and Nicaraguan Sign Language. The establishment of large schools with dedicated classrooms for deaf children along the Pacific coast of Nicaragua (particularly in Managua) resulted in the first significant large-scale contact between deaf individuals of varying ages who had heretofore been isolated in their homes. They came together in schools with an emphasis on an oral philosophy of education, but the important factor is that *they came together*.

Each student brought to the interaction a repertoire of single gestures gleaned from the gesticulations and limited gesturing of their parents—a system of communication that served them well at home when their gestures occurred in the shared context of family members who could infer the intended message without recourse to language. However, outside of the home and without recourse to these language handlers, students were forced to fend for themselves and establish a base of communication with other deaf students from similar backgrounds (See Kegl and Spitz (in press) for a discussion of the role that the demand for *communication accountability* plays in language emergence.). Quickly the inadequacy of single, idiosyncratic gestures that signal whole events became apparent. Students repeated themselves, copied and used the gestures of others, and strung together whatever gestures they had until the information was conveyed. Gradually a shared system of *contact gesturing*, still not a language, arose and became elaborated. It took on characteristics that in contrast with gesture were sufficient to draw the attention of young children in the schools who were still within the critical period for native language acquisition. The only problem is that there was no *language* to acquire. The only thing accessible was a highly variable system of gestural communication within the group and within each individual. The richness of this system lay in the various gestures and gestural strategies used among its members. Like a beehive, the complexity lay in the product of the group as a whole.

Contact gesturing did have many characteristics not encountered in simple gesturing. Most importantly, it was periodic in its form. Instead of single gestures that evoked an entire message at home. Contact gesturing with peers who also lacked a language resulted in much repetition and alternate attempts to communicate a message. Despite any consistency or grammar, this repetition and redundancy had a rhythm to it that is characteristic of all human languages and is the feature that invariably draws the language-ready and language-thirsty brain to its most likely source of linguistic input (Kegl, 2002).

Contact gesturing does exhibit some structural properties distinct from isolated gesture, including multi-gesture utterances, compounding of two gestures RECTANGLE+TURN-KNOB (radio), SCRATCH+SMALL-ANIMAL (cat), etc., single valence constructions involving an object and an action and a stringing together of such utterances to encode larger conceptual events (BOY LOOK-AT, GIRL POSE; MAN GIVE, WOMAN TAKE; etc. Contact gesturing also distinguishes itself from isolated gesture and home signing in having name signs for individuals.

In summary, we have the single *gestures* per event or the *gesticulations* that cooccur with speech that are typical of the dominant culture in an area that may be used between family members who can hear and single deaf family member. This is distinct from *home sign*, which also occurs in hearing families with a deaf child, but where the deaf child builds upon the gestures used at home and produces innovative signs and sign sequences in an elaborated gesture system. *Contact gesturing* moves communication out of the here and now setting where there is shared knowledge among all members of the family to a community level, as in schools. In this setting, gestures are repeated, borrowed from one signer to another, strung together in a variety of ways, and at times are conventionalized to allow for communication beyond the here and now, where shared background cannot be presumed.

Transitional states of communication like homesign and contact gesturing appear to constitute sufficient input into to a child's language learning process to yield language creation if they serve as input to young children. Such language creation occurred on a large scale along the Pacific Coast of Nicaragua. It occurs on a small scale in the creation of family signed languages, in the context of multiple deaf members in a single family. Further research comparing and contrasting large scale emergence as well as the emergence of family signed languages (as opposed to homesign systems) is critical to determining the nature of the lower bound on the input to language creation.

## Conclusion

I would like to conclude with a call for convergence in how we talk about input to the process of signed language acquisition. This chapter has reviewed numerous cases of language-relevant input to children as well as cases of communication that were insufficient to support language creation. By language creation, I refer to the output of any child's acquisition process that yields a full-fledged human language, whether that be a spoken language like English or Warlpiri or a signed language like American Sign Language or Nicaraguan Sign Language.

Fruitful discussion on the topic of input to signed language acquisition has been hampered by the lack of a shared nomenclature for talking about the varying types of communication systems that exist as language-relevant input. Whether researchers espouse a nativist view of language where there is a saltation between a language-relevant communicative input and what we would call *language* or whether researchers believe that that process is incremental and ever increasing in complexity, we need a line in the sand before which and beyond which we can all agree that what we are dealing with is an instantiation of the class of recognized, full-fledged, human languages. Only then, can the various types of communicative input to children discussed in this chapter begin to receive the rigorous attention they deserve. Only then can the apparent exceptionality of languages like Riau Indonesian (claimed to lack syntax, Gil, 2005); and Pirahã (claimed to lack recursion, Everett, 2005, 2009) be fully explored and brought to bear on our current shared assumptions about the range of variation in human languages.

Jackendoff and Wittenberg (2014) focus on grammars in communicative systems up to and including those languages we recognize as full members of the class of natural human languages. In the case of the "simplest grammars" they include one-word grammars, two-word-grammars, and concatenative grammars, all of which can exist without the presence of human language syntax as we know it. It should be noted, however, that "simplest" in terms of a lack of hierarchical organization does not necessarily mean simplest in terms of processing. Newport and Morgan (1981) found that concatenative artificial grammars were harder for humans to learn than those with hierarchical structure. Jackendoff and Wittenberg address varying options within the interface between form, meaning, and syntax. These early grammars they consider capture many of the structural phenomena we encounter when looking at single gestures for events, home signing, and the contact gestural communication that feeds language creation. The study of language emergence needs this kind of a meta-language for talking about grammars within the larger class of communication systems we have encountered, only some of which feed into the language creation process.

In seeking a common ground for the discussion of language emergence, nomenclature matters. When discussing less complex grammars, Jackendoff and Wittenberg have been careful not to use nomenclature from syntax such as parts of speech, syntactic categories like NP, VP or structurally defined concepts like subject and object. Whether they realize it or not, they have drawn a theoretical line in the sand regarding the use of nomenclature from syntax to talk about these grammars claimed to lack syntax. The literature on sign language emergence has demonstrated less care in this regard and, as a result, we lack a common ground for theoretical interchange. Such interchange is critical at this time because of the need for cross-comparison and typological analysis among the proliferation of candidates for communication systems potentially on or having completed their travels on the road toward language emergence. Much linguistic research has focused on determining grammars for the output of language production. In the field of emerging signed languages, we also need to bring these tools to bear on the input to the varying communication systems and languages we have encountered.

## NOTES

1 By implicational Universals (following Greenberg, 1963), I am referring to linguistic universals that are not general across all languages in all modalities, but rather follow from some specific subset of languages distinguished by modality-specific characteristics, in this case the subset of languages produced in the visual-gestural modality. According to Greenberg, most universals are implicational: If a language has property X, then it has property Y.

2 In the population studies we conducted in Nicaraguan the 1980s and 1990s, we found that while occurrence of hearing loss is higher (resulting from factors like illnesses with high fevers, use of ototoxic drugs, and so on (Polich, 2005)), the occurrence of deaf children born to deaf parents is actually much lower (a handful of cases in over 3,000 subjects studied). In Nicaragua in the 1980s, it was almost negligible and in the cases where it did happen, deaf children were often put up for adoption or raised by hearing, non-signing relatives. Similar observations are reported regarding other sites throughout South and Central America.

3 ASL, and most signed languages, have a highly productive system of word formation. The remnants of these productive word formation rules are still transparent even in frozen lexical items that second-language learners treat as unanalyzable wholes. Nouns are deverbal nominals derived from sentential verbs, often with other verbs embedded within them. The rich aspectual and verb agreement morphology packed into these signs is unpackable by the young learner and allows them to recover the grammar of ASL in a way that their parents cannot. For a more in depth discussion see Kegl (2002).

# REFERENCES


Armstrong, D. F., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the Nature of Language*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511620911

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Behrens, H. (2002). Learning multiple regularities: Evidence from overregularization errors in the German plural. In A. H. J. Do, L. Dominguez, & A. Johnson, (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 72–83). Somerville, MA: Cascadilla Press.

Bernstein, M. (1980). Acquisition of locative expressions by deaf children learning american sign language. Doctoral dissertation, Boston University, Boston, MA.

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.

Brown R. (1980). Why are signed languages easier to learn than spoken languages? In Stokoe, W. C. (Ed.), *Proceedings of the First National Symposium on Sign Language Research and Teaching* (pp. 9–24). Washington, DC: National Association of the Deaf.

Charrow, V. R. & Fletcher, J. D. (1974). English as the second language of deaf children. *Developmental Psychology*, 10, 463–470. doi: 10.1037/h0036711

Chen, D. (1999). Investigation of Word Order Acquisition in early ASL. In D. Braze, K. Hiramatsu, & Y. Kudo (Eds.), *Cranberry Linguistics*, [University of Connecticut Working Papers in Linguistics *10*].

Clahsen, H. (2008). Chomskyan syntactic theory and language disorder. In M. J. Ball, M. Perkins, N. Mueller, & S. Howard (Eds.), *The Handbook of Clinical Linguistics* (pp. 165–183). Oxford: Blackwell. doi: 10.1002/9781444301007.ch10

Clark, E. V. (1973). Non-linguistic strategies and the acquisition of word meaning. *Cognition*, 2, 161–182. doi: 10.1016/0010-0277(72)90010-8

Coppola, M. (2002). The emergence of grammatical structure: Evidence from family-based gesture (home sign) systems in Nicaragua. Doctoral dissertation, University of Rochester.

Coppola, M. & Newport, E. (2005). Grammatical subjects in homesign: Abstract linguistic structure in adult primary gesture systems without linguistic input. *Proceedings of the National Academy of Sciences*, 102(52), 19249–19253. doi: 10.1073/pnas.0509306102

Coppola, M. & So, W.C. (2005). Abstract and Object-Anchored Deixis: Pointing and spatial layout in adult homesign systems in Nicaragua. In A. Brugos, M. R. Clark-Cotton, & S. Ha, (Eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development* (pp. 144–155). Somerville, MA: Cascadilla Press.

Coppola, M. & So. W. C. (2006). The seeds of spatial grammar: Spatial modulation and coreference in homesigning and hearing adults. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 119–130). Somerville, MA: Cascadilla Press.

Cornett, R. O. (1967). Cued Speech. *American Annals of the Deaf*, 112(1), 3–13.

Curtiss, S. (1977). *Genie: A Psycholinguistic Study of a Modern-Day "Wild Child."* New York: Academic Press.

Everett, D. (2005). Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology*, 46, 621–646. doi: 10.1086/431525

Everett, D. (2009). Pirahã culture and grammar: A response to some criticisms.

*Language*, 85, 405–442. doi: 10.1353/lan.0.0104
Gil, D. (2005). Word order without syntactic categories: How Riau Indonesian does it. In A. Carnie, H. Harley, & S. A. Dooley (Eds.), *Verb First: On the Syntax of Verb-Initial Languages,* (pp. 243–263). Amsterdam: John Benjamins. doi: 10.1075/la.73.17gil
Goldin-Meadow, S. (2003). *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us About How All Children Learn Language*. New York: Psychology Press.
Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: removing the handcuffs on grammatical expression in the manual modality. *Psychological Review*, 103, 34–55. doi: 10.1037/0033-295X.103.1.34
Goldin-Meadow, S. & Mylander, C. (1984). Gestural communication in deaf children: the effects and non-effects of parental input on language development. *Monographs of the Society for Research in Child Development*, 49, 1–121. doi: 10.2307/1165838
Goldin-Meadow, S. & Mylander. C. (1990). The role of parental input in the development of a morphological system. *Journal of Child Language*, 17, 527–563. doi: 10.1017/S0305000900010874
Greenberg, J. (1963). *Universals of Language*. Cambridge: The MIT Press.
Grosjean, F. (1982). *Life with Two Languages*. Cambridge, MA: Harvard University Press.
Grosjean, F. (2010). Bilingualism, biculturalism, and deafness. *International Journal of Bilingual Education and Bilingualism*, 13(2), 133–145. doi: 10.1080/13670050903474051
Hoffmeister, R. (1978). *The development of demonstrative pronouns, locatives, and personal pronouns in the acquisition of american sign language*. Doctoral dissertation, University of Minnesota.
Hoiting, N. & Slobin, D. (1992). Transcription as a tool for understanding: The Berkeley Transcription System for Sign Language Research (BTS). In G. Morgan & B. Woll (Eds.), *Language Acquisition in Signed Languages* (pp. 55–75). Cambridge: Cambridge University Press.
Jackendoff, R. & Wittenberg, E. (2014). What you can say without syntax: A hierarchy of grammatical complexity. In F. J. Newmeyer & L. B. Preston (Eds.), *Measuring Grammatical Complexity* (pp. 65–82). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199685301.003.0004
Kantor, R. (1982). Communicative interaction: Mother modification and child acquisition of American Sign Language. *Sign Language Studies*, 36, 233–282. doi: 10.1353/sls.1982.0018
Kendon, A. (1988). *Sign Languages of Aboriginal Australia: Cultural, Semiotic and Communicative Perspectives*. Cambridge: Cambridge University Press.
Kegl, J. (2000). Is it soup yet? Or, When is it Language? In the *Proceedings of the Child Language Seminar 1999*. City University, London.
Kegl, J. (2002). Language emergence in a language-ready brain: Acquisition issues. In G. Morgan. & B. Woll (Eds.), *Language Acquisition in Signed Languages* (pp. 207–254). Cambridge: Cambridge University Press.
Kegl, J. (2008). The case of signed languages in the context of Pidgin and Creole studies. In J. Singler & S. Kouwenberg, (Eds.), *The Handbook of Pidgin and Creole Studies* (pp. 491–511). London: Blackwell. doi: 10.1002/9781444305982.ch20
Kegl, J. & McWhorter, J. (1997). Perspectives on an emerging language. *Proceedings of the Stanford Child Language Research Forum* (pp. 15–36). Palo Alto: Center for the Study of Language and Information.
Kegl, J. & S. Schley. (1986). When is a classifier no longer a classifier? In N. Nikiforidou, M. Van Clay, M. Niepokuj, & D. Feder (Eds.), *Proceedings of the Twelfth Annual Meeting of the Berkeley*

*Linguistic Society* (pp. 425–441), Berkeley, Calif.: Berkeley Linguistics Society. doi: 10.3765/bls.v12i0.1881
Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (Ed.), *Comparative Grammatical Change: The Intersection of Language Acquisition, Creole Genesis, and Diachronic Syntax* (pp. 179–237). Cambridge, MA: The MIT Press.
Kegl, J. A. & White Eagle, J. (1986). Plains Indian Sign Language. In S.P. Parker (Ed.), *Gallaudet Encyclopedia of Deaf People and Deafness* (pp. 97–100). New York: McGraw Hill Book Company.
Kuhl, P. K., Conboy, B. T., Padden, D., Nelson. T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period." *Language Learning and Development*, 1(3&4), 237–264. doi: 10.1207/s15473341lld0103&4_2
Lane, H. (1976). *The Wild Boy of Averyon*. Cambridge: Harvard University Press.
Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, & Computers*, 41(3), 841–849. doi: 10.3758/BRM.41.3.841
Lenneberg, E. (1967). *Biological Foundations of Language*. New York: John Wiley and Sons, Inc.
Lieberman, A. Hatrak, M., & Mayberry, R. I. (2014). Learning to look for language: Development of joint attention in young deaf children. Language Learning and Development, 10(1), 19–35. doi: 10.1080/15475441.2012.760381
Lieven, E. (1992). Afterword: A view from research on spoken language development. In G. Morgan & B. Woll (Eds.), *Language Acquisition in Signed Languages* (pp. 277–290). Cambridge: Cambridge University Press.
Lillo-Martin, D. C. (1991). *Universal Grammar and American Sign Language*. Dordrecht: Kluwer Publishing. doi: 10.1007/978-94-011-3468-2
Lillo-Martin, D. C. (2000). Aspects of the syntax and acquisition of *wh*-questions in American Sign Language. In K. Emmorey & H. Lane (Eds.), *The Signs of Language Revisited: An Anthology in Honor of Ursula Bellugi and Edward Klima*. Mahwah, NJ: Lawrence Erlbaum Associates.
Loew, R. (1984). *Role and Reference in American Sign Language*. Doctoral dissertation, University of Minnesota.
MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
Maestas y Moores, J. (1980). Early linguistic environment: Interactions of deaf parents with their infants. *Sign Language Studies*, 26(1), 1–13. doi: 10.1353/sls.1980.0003
Mayberry, R. I. (1993). First language acquisition after childhood differs from second-language acquisition: The case of American Sign Language. *Journal of Speech and Hearing Research*, 36, 1258–1270. doi: 10.1044/jshr.3606.1258
Mayberry, R. I. (2010). Early language acquisition and adult language ability: What sign language reveals about the critical period for language. In M. Marschark & P. E. Spencer (Eds.), *The Oxford Handbook of Deaf Studies, Language and Education*, Volume 2 (pp. 281–291). Oxford: Oxford University. doi: 10.1093/oxfordhb/9780195390032.013.0019
Mayberry, R. I. & Eichen, E. B. (1991). The long lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language*, 30, 486–512. doi: 10.1016/0749-596X(91)90018-F
McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
Mitchell, R. E. & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, 4(2), 138–163. doi: 10.1353/sls.2004.0005

Morford, J. P. (1996). Insights into language from the study of gesture: A review of research on the gestural communication of non-signing deaf people. *Language and Communication*, 16(2), 165–178. doi: 10.1016/0271-5309(96)00008-0

Morford, J. P. & Hänel-Faulhaber, B. (2011). Homesigners as late learners: Connecting the dots from delayed acquisition in childhood to sign language processing in adulthood. *Language and Linguistics Compass*, 5(8), 525–537. doi: 10.1111/j.1749-818X.2011.00296.x

Morford, J. P. & Kegl, J. (2000). Gestural Precursors to Linguistic Constructs: How input shapes the form of language. In D. McNeill (Ed.), *Language and Gesture* (pp. 358–387). Cambridge: Cambridge University Press.

Morgan, J. L. & Newport, E. L. (1981). The Role of Constituent Structure in the Induction of an Artificial Language. *Journal of Verbal Learning and Verbal Behavior*, 20(1), 67–85. doi: 10.1016/S0022-5371(81)90312-1

Morgan, G. & Woll, B. (Eds.) (1982). *Directions in Sign Language Acquisition*. Amsterdam/Philadelphia: John Benjamins Publishing.

Mufwene, S. (1999). On the language bioprogram hypothesis: Hints from Tazie. In M. DeGraff (Ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development* (pp. 95–127). Cambridge, MA: The MIT Press.

Neidle, C. & MacLaughlin, D. (1998). Signstream™: A tool for linguistic research on signed languages. *Sign Language and Linguistics*, 1(2), 111–114. doi: 10.1075/sll.1.1.14nei

Newport, E. (1988). Constraints on learning and their role in language acquisition: Studies of the Acquisition of American Sign Language. *Language Science*, 10, 147–172. doi: 10.1016/0388-0001(88)90010-1

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science*, 14, 11–28. doi: 10.1207/s15516709cog1401_2

Newport, E. & Meier, R. (1987). The acquisition of American Sign Language. In D. I. Slobin, (Ed.), *The Cross-linguistic Study of Language Acquisition, Volume 1: The Data* (pp. 881–938). Hillsdale, NJ: Lawrence Erlbaum Associates.

Palmer, S. B., Fais, L., Golinkoff, R. M., & Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Development*, 83, 543–553. doi: 10.1111/j.1467-8624.2011.01715.x

Paul, P. (1998). *Literacy and Deafness: The Development of Reading, Writing and Literate Thought*. Needham, MA: Allyn & Bacon.

Peal, E. & Lambert, W. (1962). The relationship of bilingualism to intelligence. *Psychological Monographs*, 76(27), 1–23. doi: 10.1037/h0093840

Penfield, W. & Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton, NJ: Princeton University Press.

Perniss, P., Thompson, R. L., & Vigilocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, 1(227). doi: 10.3389/fpsyg.2010.00227

Pettito, L. A. (1987). On the autonomy of language and gesture: Evidence from the acquisition of personal pronouns in American Sign Language. *Cognition*, 27, 1–52. doi: 10.1016/0010-0277(87)90034-5

Petitto, L. A. (2000). The acquisition of natural signed languages: Lessons in the nature of human language and its biological foundations. In C. Chamberlain, J. P. Morford, & R. I. Mayberry, (Eds.), *Language Acquisition by Eye* (pp. 41–50). Mahwah, NJ: Lawrence Erlbaum Associates.

Pettito, L. A. & Marentette, P. F. (1991). Babbling in the manual mode: Evidence for the ontogeny of language. *Science*, 251, 1493–1496. doi: 10.1126/science.200642

Poizner, H. & Tallal, P. (1987). Temporal processing in deaf signers. *Brain and Language*, 30, 52–62. doi: 10.1016/0093-934X(87)90027-7
Polich, L. (2005). *The Emergence of the Deaf Community in Nicaragua*. Washington, D.C.: Gallaudet University Press.
Reed, C. M., Delhorne, L.A., Fischer, S.D., & Durlach, N.I. (1995). A study of the tactual reception of sign language. *Journal of Speech and Hearing Research*, 8, 477–489. doi: 10.1044/jshr.3802.477
Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar in a new sign language. *Proceedings of the National Academy of Sciences*, 102(7), 2661–2665. doi: 10.1073/pnas.0405448102
Singler, J. V. (2008). The sociohistorical context of creole genesis. In S. Kouwenberg & J. V. Singler (Eds.), *The Handbook of Pidgin and Creole Studies* (pp. 332–358). Oxford: Wiley-Blackwell. doi: 10.1002/9781444305982.ch14
Singleton, J., Goldin-Meadow, S., & McNeill, D. (1995). The "cataclysmic break" between gesticulation and sign. In K. Emmorey & J. Reilly (Eds.), *Sign, Gesture, and Space* (pp. 287–311). Hillsdale, NJ: Lawrence Erlbaum.
Singleton, J. & Newport, E. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognition*, 49, 370–407. doi: 10.1016/j.cogpsych.2004.05.001
Spitz, R.V. & Kegl, J. A. (2005). Memory and visuospatial analysis with and without language: Lessons from homesigners in Nicaragua. Doctoral dissertation, University of Southern Maine, Portland, ME.
Spitz, R. V. & Kegl, J. A. (in press). Enhancing communication skills in persons with severe language deprivation: Lessons learned from the rise of a signing community in Nicaragua. In N. Glickman & W. Hall (Eds.), *Language Deprivation and Mental Health Care*. New York: Routledge.
Supalla, S. J. (1986). Manually coded English: The modality question in signed language development. Masters thesis, University of Illinois at Urbana-Champaign.
Supalla, S. J. (1991). Manually coded English: The modality question in signed language development. In P. Siple & S. D. Fischer (Eds.), *Theoretical Issues in Sign Language Research, Volume 2: Psychology* (pp. 85–109). Chicago: University of Chicago Press.
Supalla, T. (1982). Structure and acquisition of verbs of motion and location in American Sign Language. Doctoral dissertation, University of California, San Diego.
Tervoort, B. (1953). *Structurele Analyse van Visueel Taalgebruik binnen een Groep Dove Kinderen. [Structural Analysis of visual language use within a group of deaf children.] Deel 1*. Amsterdam: North-Holland Publishing Company.
Tervoort, B. (1959). Language development in young deaf children. *Proceedings of the 39th Convention of American Instructors of the Deaf* (pp. 120–124). Colorado Springs: National Association of the Deaf.
Van den Bogaerd, B., & Baker, A. (2002). Are deaf children bilingual? In G. Morgan & B. Woll (Eds.), *Language Acquisition in Signed Languages* (pp. 183–206). Cambridge: Cambridge University Press.
Washabaugh, W. (1986). *Five Fingers for Survival: Deaf Sign Language in the Caribbean*. Ann Arbor: Karoma Press.
Wilson, M. & Emmorey, K. (1997). A visuo-spatial "phonological loop" in working memory: Evidence from American Sign Language. *Memory & Cognition*, 25, 313–320. doi: 10.3758/BF03211287

# Index



*The Handbook of Psycholinguistics*, First Edition. Edited by Eva M. Fernández and Helen Smith Cairns.