
Theses and Dissertations

Spring 2010

A frequency and error analysis of the use of determiners, the relationships between noun phrases, and the structure of discourse in English essays by native English writers and native Chinese, Taiwanese, and Korean learners of English as a Second language

Jane E. Gressang
University of Iowa

Copyright 2010 Jane E Gressang

This dissertation is available at Iowa Research Online: <http://ir.uiowa.edu/etd/507>

Recommended Citation

Gressang, Jane E.. "A frequency and error analysis of the use of determiners, the relationships between noun phrases, and the structure of discourse in English essays by native English writers and native Chinese, Taiwanese, and Korean learners of English as a Second language." PhD (Doctor of Philosophy) thesis, University of Iowa, 2010.
<http://ir.uiowa.edu/etd/507>.

Follow this and additional works at: <http://ir.uiowa.edu/etd>

 Part of the [Linguistics Commons](#)

A FREQUENCY AND ERROR ANALYSIS OF THE USE OF DETERMINERS, THE
RELATIONSHIPS BETWEEN NOUN PHRASES, AND THE STRUCTURE OF
DISCOURSE IN ENGLISH ESSAYS BY NATIVE ENGLISH WRITERS AND
NATIVE CHINESE, TAIWANESE, AND KOREAN LEARNERS OF ENGLISH AS A
SECOND LANGUAGE

by
Jane E. Gressang

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Linguistics
in the Graduate College of
The University of Iowa

May 2010

Thesis Supervisor: Associate Professor Roumyana Slabakova

ABSTRACT

Second language (L2) learners notoriously have trouble using articles in their target languages (e.g., *a*, *an*, *the* in English). However, researchers disagree about the patterns and causes of these errors. Past studies have found that L2 English learners:

- Predominantly omit articles (White 2003, Robertson 2000),
- Overuse *the* (Huebner 1983, Master 1987, Parrish 1987, Tarone & Parrish 1988, Thomas 1989, Ionin 2003), or
- Overuse *a* (Leung 2001).

Previously proposed explanations of the causes of article errors include:

- Learners have incorrect or incomplete semantic representations (Tarone & Parrish 1988, Hawkins & Chan 1997, Goto Butler 2002, Ionin 2003), or
- Learners have complete, correct semantic representations for articles, but difficulty choosing the lexical form during production due to stress on mental processing or phonological limitations (Lardiere 1998, Bruhn de Garavito & White 2000, White 2003, Goad, White, & Steele 2003).

Prior studies have focused on articles, which identify discourse relationships, but have not considered other morphemes that do so as well, such as pronouns and demonstratives. Furthermore, they have focused on L2 errors in isolation and not in the context of a full discourse or contrasted with first language (L1) input. This study examined the use of articles and other discourse morphemes in 20 L1 and 20 L2 English essays. L2 essays were produced by L1 Chinese and Korean writers at two proficiency levels. The essays' noun phrases (NPs) were marked for part-of-speech, co-reference, syntactic position, and other discourse-relevant features. L2 errors were identified and categorized.

Frequency data showed that L2 proficiency level more often indicated significant differences in discourse construction than L1. No significant difference between L2 and

L1 writers was when considering all articles together. Breaking this down, students used *a/an* significantly less than L1 writers, but the use of *the* was not significantly different. In contrast, the error analysis showed most L2 mistakes being made in the use of *the*, with almost none in the use of *a/an*. Together the frequency and error data give a richer understanding of discourse and article use in L2 production.

Abstract Approved: _____
Thesis Supervisor

Title and Department

Date

A FREQUENCY AND ERROR ANALYSIS OF THE USE OF DETERMINERS, THE
RELATIONSHIPS BETWEEN NOUN PHRASES, AND THE STRUCTURE OF
DISCOURSE IN ENGLISH ESSAYS BY NATIVE ENGLISH WRITERS AND
NATIVE CHINESE, TAIWANESE, AND KOREAN LEARNERS OF ENGLISH AS A
SECOND LANGUAGE

by

Jane E. Gressang

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Linguistics
in the Graduate College of
The University of Iowa

May 2010

Thesis Supervisor: Associate Professor Roumyana Slabakova

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Jane E. Gressang

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Linguistics at the May 2010 graduation.

Thesis Committee: _____
Roumyana Slabakova, Thesis Supervisor

William Davies

Elena Gavrusseva

Jason Rothman

Jerzy Rubach

To Chance

An impossible result should lead to infinite surprise.

William L. Hays, *Statistics*, Fifth Edition

ACKNOWLEDGMENTS

I would first like to thank my advisor, Roumyana Slabakova, and all of my committee members—Bill Davies, Elena Gavrusseva, Jerzy Rubach, and Jason Rothman—for their kind and timely consideration of my work. My appreciation also goes out to those in the Department of Linguistics who have assisted me over the course of this process, including Maureen Burke, Barb Hermeier, Alice Davison, Jill Beckman, Cathie Ringen, and Marc Light. I would also like to thank Bob Arens for his kind assistance with some of the vagaries of Python and Perl.

Much gratitude and many thanks also go to Joni Lakin for the dissertation writing group. I am broker, but more full of coffee and chocolate. I would also like to thank Hua Ou for being a consistent support in the group.

Finally, there are no words to adequately thank all my family and friends who have traveled on this journey with me for so long, so I will express this simply. Thank you Mom, Dad, Chance, Debi, John, and Katie. Also, thank you to Melissa Meisterheim, Craig Dresser, Erum Nadeem, Christine Prince, Vicki Andrienas, Valerie Pruvost, Peter Rohrbough, Ivan Ivanov, Tomomasa Sasa, Jessica Goetz, LlynnAnn Lluellen, Sarah Muirhead, Jen Polinski, and Lauren Eby. What are we going to talk about now?

Thanks again to all for your helpful advice, unending support, and unfailing optimism.

ABSTRACT

Second language (L2) learners notoriously have trouble using articles in their target languages (e.g., *a*, *an*, *the* in English). However, researchers disagree about the patterns and causes of these errors. Past studies have found that L2 English learners:

- Predominantly omit articles (White 2003, Robertson 2000),
- Overuse *the* (Huebner 1983, Master 1987, Parrish 1987, Tarone & Parrish 1988, Thomas 1989, Ionin 2003), or
- Overuse *a* (Leung 2001).

Previously proposed explanations of the causes of article errors include:

- Learners have incorrect or incomplete semantic representations (Tarone & Parrish 1988, Hawkins & Chan 1997, Goto Butler 2002, Ionin 2003), or
- Learners have complete, correct semantic representations for articles, but difficulty choosing the lexical form during production due to stress on mental processing or phonological limitations (Lardiere 1998, Bruhn de Garavito & White 2000, White 2003, Goad, White, & Steele 2003).

Prior studies have focused on articles, which identify discourse relationships, but have not considered other morphemes that do so as well, such as pronouns and demonstratives. Furthermore, they have focused on L2 errors in isolation and not in the context of a full discourse or contrasted with first language (L1) input. This study examined the use of articles and other discourse morphemes in 20 L1 and 20 L2 English essays. L2 essays were produced by L1 Chinese and Korean writers at two proficiency levels. The essays' noun phrases (NPs) were marked for part-of-speech, co-reference, syntactic position, and other discourse-relevant features. L2 errors were identified and categorized.

Frequency data showed that L2 proficiency level more often indicated significant differences in discourse construction than L1. No significant difference between L2 and

L1 writers was when considering all articles together. Breaking this down, students used *a/an* significantly less than L1 writers, but the use of *the* was not significantly different. In contrast, the error analysis showed most L2 mistakes being made in the use of *the*, with almost none in the use of *a/an*. Together the frequency and error data give a richer understanding of discourse and article use in L2 production.

TABLE OF CONTENTS

LIST OF TABLES	xii
LIST OF FIGURES	xvii
LIST OF EXAMPLES.....	xix
CHAPTER	
I. INTRODUCTION	1
A Brief Discussion of Theoretical, Generative Linguistics, Corpus Linguistics, and First Language Acquisition	1
Review of the Literature and Prior Studies.....	5
Common Hypotheses Regarding Article and Determiner Acquisition	7
Feature-Based Discourse Models.....	7
Functional Morphology Error Theories and Studies	8
Functional Morphology Error Theory 1: The Failed Functional Features Hypothesis—Hawkins and Colleagues.....	8
Functional Morphology Error Theory 2: The Missing Surface Inflection Hypothesis	9
Functional Morphology Error Theory 3: The Prosodic Transfer Hypothesis	12
Functional Morphology Error Theories: Conclusions	16
Studies and Theoretical Explanations Focusing Only on Articles	18
Article-Focused Theory 1: Discourse Rule Transfer— Robertson (2000)	18
Article-Focused Theory 2: The Article Choice Parameter and the Fluctuation Hypothesis—Ionin, Wexler, and Colleagues	20
Questions about the Theoretical Foundation of Binary Feature- Based Discourse Models.....	21
Hawkins (1978)	21
Binary Features Checked in Syntax.....	26
Feature-Based Discourse Theories: Summary	29
Alternative Models of Discourse.....	30
Discourse Models	30
Anaphoric Hierarchies	32
Accessibility Theory—Ariel (1990)	33
The Givenness Hierarchy—Gundel, Hedberg, and Zacharski (1993, 1998, 2001, 2003)	35
Transition Models	40
Centering Theory	43
A Simple Transition Model—von Stutterheim, Mangold- Allwinn, Baratelli, Kohlmann, and Kolbing (1993)	47
Semantics-Focused Models	49
Discourse Representation Theory—Kamp and Reyle (1993)	49
Text-Type Models.....	52
Modes of Discourse—Smith (2003)	52

Von Stutterheim et al. (1993)	56
Summary and Evaluation of Discourse Models	58
L2 Studies Using Alternative Models of Discourse	59
L2 Use of Null Subjects—Saunders (1999)	60
Overall Discourse Cohesion—Strömquist and Day (1993)	61
L1 Transfer of Nominal Reference—Hendriks (2003)	63
Null Pronouns and Over-Explicit Reference—Nistov (2001)	65
Summary of Findings	65
Research Goals	67
The Three Main Research Aims	67
Research Questions	68
Questions Related to Goals One and Two	68
Questions Related to Goal Three	69
Questions for Future Research: Goal Four	70
Questions for Future Research: Goal Five	71
II. METHODS	72
Data Collection	72
Collection of the Corpora of Essays	72
Collection of the Learner Corpus	72
Description of the Learner Essays and Proficiency Exams	72
The EPE: For Admitted Students	73
The IIEP Exam: For Un-Admitted Students	73
Similarities between the EPE and the IIEP Exam	74
Essay Topics	75
Advantages and Complications Stemming from the Method of Collecting the Learner Corpus	76
Collection of the Native Corpus	78
The Native Writer Testing Situation	78
Essay Topics	79
Methods of Coding Data	80
Part-of-Speech Tagging	83
Defining Parts-of-Speech	84
Human Classification of Parts-of-Speech	84
How Machine Classification of Parts-of-Speech is Different from Human Coders'	90
The Computer Algorithm Used by MontyLingua-2.1 to Tag Parts-of-Speech	94
How Automatic Part-of-Speech Taggers Function	95
Three Common Approaches to Automatic Tagging	96
Rule-Based Taggers	96
Stochastic Taggers, Including Hidden Markov Models	97
Transformation-Based Taggers	101
The Brill (1994) Transformation-Based Tagging Method	102
Summary	107
Tag Correction and the Accuracy of Automatic Taggers	107
Adapting and Correcting MontyLingua-2.1 Output for This Study	109
Tag Removal	110
Tag Correction	112
Phrase Correction	114
Co-reference Tagging	120

	Characteristics Shared by Commonly Used Co-reference	
	Tagging Schemes.....	120
	The MUC-7 Co-reference Tagging Scheme	123
	Tag the Largest Phrasal Unit Possible	125
	Embedded NPs and Co-reference in PPs and Relative	
	Clauses	125
	Conjoined NPs	130
	Appositions.....	133
	Meaning Connections that Cannot be Tagged.....	134
	NPs after Copula Verbs.....	138
	Meaning Functions.....	140
	Summary	142
	Statistical Analysis Methods	143
III.	RESULTS, GOALS ONE AND TWO	145
	Total Number of Words	145
	Native Speakers vs. Student Writers, Number of Words	145
	IIEP vs. EPE Students and Korean vs. Chinese/ Taiwanese Students,	
	Number of Words.....	148
	Total Number of Noun Phrases	150
	Native Speakers vs. Students, Noun Phrases	150
	IIEP vs. EPE Students, Noun Phrases	155
	Korean vs. Taiwanese/Chinese Students, Noun Phrases	155
	Co-reference Chains.....	158
	Number and Length of Co-reference Chains for Individual Writers	160
	Other Patterns in Co-reference Chains	163
	Group Comparisons of Co-reference Chains.....	171
	Natives vs. Students, Number of Chains per Total Number of	
	Words.....	172
	Transitions	178
	New Referents or Transitions	178
	Native vs. Student Writers, Number of New Referents.....	178
	IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese	
	Students, Number of New Referents	181
	IIEP vs. EPE Students.....	181
	Korean vs. Chinese/Taiwanese Essays	187
	Maintained Referents or Transitions	188
	Native vs. Student Writers, Number of Maintained Referents.....	188
	IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese	
	Students, Number of Maintained Referents	194
	IIEP vs. EPE Students, Maintained Referents.....	194
	Korean vs. Chinese/Taiwanese Students.....	196
	Re-mentioned Referents or Transitions	199
	Native vs. Student Writers, Number of Re-mentioned Referents ...	199
	IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese	
	Students, Number of Re-mentioned Referents.....	200
	IIEP vs. EPE Students.....	202
	Korean vs. Chinese/Taiwanese Students.....	205
	Conclusions about Transitions	205
	Syntactic Position.....	208
	Classification System for Syntactic Position	209
	Subjects.....	209
	Verbal Objects	210

Objects of Prepositions	210
Genitive Specifiers.....	211
Copula Complements.....	212
Appositions.....	212
Comparisons	213
Titles	214
Native vs. Student Writers, NPs in Various Syntactic Positions	215
IIEP vs. EPE Students, Syntactic Positions of NPs.....	216
Korean vs. Chinese/Taiwanese Students, NPs in Various Syntactic Positions.....	217
Conclusion	222
Total Number of Distinct Words (Word Types).....	224
Native vs. Student Writers, Number of Types	225
IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese Students, Number of Word Types	237
IIEP vs. EPE Students, Number of Word Types	238
Korean vs. Chinese/Taiwanese Students, Number of Word Types.....	239
Articles, Demonstratives, and Pronouns	241
Student L2 English vs. Native English Writers	241
Frequency of Use of <i>the</i>	243
Number of Pronouns.....	244
Frequency of Use of Demonstrative Determiners	247
Use of <i>a/an</i>	247
Comparison of IIEP and EPE Student Sub-groups and Korean and Chinese/Taiwanese Students.....	252
Conclusions about Articles, Demonstratives, and Pronouns	253
Proper Nouns, Common Nouns, and Other Kinds of Noun Phrases	256
Summary of Significant Findings, Research Goals One and Two	263
Native English vs. L2 English/Student Writers	263
IIEP (Lower Proficiency) vs. EPE (Higher Proficiency) Students	264
Korean vs. Chinese/Taiwanese Students	265
 IV. RESULTS, RESEARCH GOAL THREE.....	 270
Overall Number of Errors in NPs	270
A Note on Optionality	271
An Examination of Possible Patterns in the Overall Error Percentages	276
Number of Errors in NPs in Co-reference Chains	281
Number of Errors in NPs that are Not in Co-reference Chains	283
Common Error Patterns.....	285
Over-use of <i>the</i>	285
Double Determiners.....	285
Names	286
Generic Statements	288
Specific Indefinites	290
Summary of the Over-use of <i>the</i>	295
Omission of <i>the</i> and Definite Determiners	298
Names	299
Unique or Defined by the Inherent Semantics of the Referent or Shared Background Information	300
Unique or Defined by the Context of the Essay	302
Another Kind of Omission Error.....	303

Summary of Omission of Definite Determiners.....	305
Noun Number, Singular and Plural Nouns, and Count or Non-Count Nouns.....	309
General Description or Noun Phrase Structure.....	318
Conclusions about Error Analysis	320
Final Conclusions.....	322
 APPENDIX A: TEST QUESTIONS AND INSTRUCTIONS GIVEN TO NATIVE ENGLISH-SPEAKING WRITERS FOR CREATION OF THE NATIVE CORPUS ...	323
 APPENDIX B: PENN TREEBANK PART-OF-SPEECH TAGS	325
 APPENDIX C: PYTHON PROGRAM USED TO INITIATE PART-OF-SPEECH TAGGING BY MONTYLINGUA-2.1	327
 APPENDIX D: PERL PROGRAM FOR AUTOMATICALLY REMOVING NON-NOUN OR DETERMINER PHRASE RELATED PART-OF-SPEECH TAGS.....	328
 APPENDIX E: CO-REFERENCE CODING RULES	329
 APPENDIX F: STUDENT AND NATIVE WRITER ESSAYS.....	349
 REFERENCES.....	376

LIST OF TABLES

Table

1.	Summary of Theoretical Perspectives of Functional Morphology Error Theories	17
2.	Strengths and Weaknesses of Morphological Error Theories	17
3.	Robertson’s Discourse Rules	19
4.	Predictions of L2 Article Use in English Using the Fluctuation Hypothesis and the Article Choice Parameter	21
5.	Robertson’s Taxonomy of Determiner Use	22
6.	Ariel’s Accessibility Hierarchy	33
7.	Gundel, Hedberg, & Zacharski’s Givenness Hierarchy for English.....	36
8.	Gundel, Hedberg, & Zacharski’s Cross-Linguistic Comparison of Discourse Forms.	41
9.	Centering Theory Transitions	44
10.	Strömquist & Day’s Anaphoric Hierarchy	62
11.	L2 Studies Using General Discourse Models and their Findings.....	66
12.	Essay Topics in the Learner Corpus.	75
13.	Essay Topics Given to Native English Writers	79
14.	Criteria for Determining Parts-of-Speech in English—Four Commonly Used Lexical Categories	87
15.	Criteria for Differentiating Parts-of-Speech in English—Other Frequently Assumed Categories	89
16.	The PennTreebank’s Eight Categories that Correspond to the Category “Verb”.....	91
17.	Comparison of the Main Features of Three Part-of-Speech Tagging Methods.....	103
18.	Total Number of Words per Essay	146
19.	Average and Median Total Number of Words per Essay	148
20.	Average and Median Total Number of Words per Essay, IIEP vs. EPE and Korea vs. China/Taiwan.	149
21.	Total Number of Noun Phrases per Essay.....	151

22.	Ranks of Total Number of Noun Phrases per Total Number of Words	153
23.	IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test.....	155
24.	Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Wilcoxon Rank Sums Test.....	156
25.	Ranks of the Total Number of Noun Phrases, Chinese/Taiwanese Students vs. Korean Students.....	157
26.	Number and Length of Co-reference Chains by Participant.....	159
27.	Data Table for the Previous Figure: Extra Lengths on Extremely Long Co-reference Chains	164
28.	Average Length of Co-reference Chains in Essays, Wilcoxon Rank Sums Test Results.....	172
29.	Mean Length of Co-reference Chains: A Comparison of IIEP and EPE Students.....	173
30.	Number of Co-reference Chains in Student vs. Native Speaker Essays, Wilcoxon Rank Sums Test Results	174
31.	Number of Co-reference Chains in IIEP vs. EPE Students' Essays, Wilcoxon Rank Sums Test Results	176
32.	Average and Median Total Number of Words, NPs, and Co-reference Chains per Essay, IIEP vs. EPE.....	176
33.	Number of Co-reference Chains in Korean vs. Chinese/Taiwanese Students' Essays, Wilcoxon Rank Sums Test Results	177
34.	Total Number of New Referents per Essay, Natives vs. Students.....	179
35.	Average and Median Total Number of New Transitions per Essay, Native vs. Student Writers.....	180
36.	Ranks of Total Number of New Referents per Total Number of Words, Native vs. Student Writers.....	182
37.	Average and Median Total Number of New Referents per Essay, IIEP vs. EPE and Korea vs. China/Taiwan.....	184
38.	IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test.....	184
39.	IIEP Essay Topics	185
40.	EPE Essay Topics	186
41.	Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Wilcoxon Rank Sums Test.....	187

42.	Total Number of Maintained Referents per Essay.....	189
43.	Average and Median Total Number of Maintained Referents per Essay, Student vs. Native Writers.....	190
44.	Ranks of Total Number of Maintained Referents per Total Number of Words, Native vs. Student Writers.....	192
45.	Average and Median Total Number of Maintained Referents per Essay, IIEP vs. EPE and Korea vs. China/Taiwan.....	195
46.	Selected Averages for Statistically Significant Differences between IIEP and EPE Students.....	195
47.	IIEP Essays as Compared to EPE Students' Essays, Maintained Referents, Wilcoxon Rank Sums Test.....	196
48.	Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Maintained Referents, Wilcoxon Rank Sums Test.	197
49.	Total Number of Re-mentioned Referents per Essay, Native vs. Student Writers.....	198
50.	Average and Median Total Number of Re-mentioned Referents per Essay, Native vs. Student Writers.....	200
51.	Ranks of Total Number of Re-mentioned Referents per Total Number of Words.....	201
52.	Average and Median Total Number of Re-mentioned Referents per Essay, IIEP vs. EPE and Korea vs. China/Taiwan.....	203
53.	IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test, Re-mentioned Referents.....	204
54.	Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Wilcoxon Rank Sums Test, Re-mentioned Referents.....	206
55.	Summary of Averages for Statistically Significant Differences between Groups.....	207
56.	Comparison of Statistical Differences between Native and Student Writers' Syntactic Positions of NPs.....	216
57.	Comparison of Statistical Differences between IIEP and EPE Student Writers' Syntactic Positions of NPs.....	217
58.	Syntactic Positions of NPs in Native and Student Essays.....	219
59.	Syntactic Positions of NPs in IIEP and EPE Student Essays.	220
60.	Syntactic Positions of NPs in Korean and Chinese/Taiwanese Student Essays.....	221

61.	Comparison of Statistical Differences between Korean and Chinese/Taiwanese Student Writers' Syntactic Positions of NPs	222
62.	Group Averages for Significant Findings Relating to Syntactic Positions of NPs.....	223
63.	Total Number of Distinct Word Types per Essay.....	226
64.	Average and Median Total Number of Word Types per Essay, Native vs. Student Writers.	227
65.	Ranks of Word Types Based on the Number of Tokens of Each Word Type, Native vs. Student Writers.....	228
66.	Three English Corpora and their High Frequency Function Words.....	231
67.	Ten Most Common Word Types in Each Essay.....	232
68.	Percentage of Word Types with Only One Token, Native vs. Student Writers.	235
69.	Percentage of Word Types with Exactly Two Tokens, Native vs. Student Writers.	236
70.	Average and Median Total Number of Word Types per Essay, IIEP vs. EPE and Korea vs. China/Taiwan.....	238
71.	IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test, Number of Word Types.	239
72.	IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test, Percentage of Word Types with Only One or Two Tokens.....	239
73.	Average and Median Total Number of Distinct Word Types per Essay, IIEP vs. EPE and Korea vs. China/Taiwan.....	240
74.	Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Wilcoxon Rank Sums Test, Number of Word Types, Percentage of Types with Only One or Two Tokens.	241
75.	Number of Articles, Demonstratives, and Pronouns per Essay.....	242
76.	Average Number of Articles, Demonstratives, and Pronouns for Native and Student Writers.	244
77.	Comparison of the Pronouns Used Most Often by Student and Native Writers.	245
78.	Use of <i>one</i> in Student Essays.....	249
79.	Summary of Statistical Findings for Differences between Groups in Articles, Demonstratives, and Pronouns, Wilcoxon Rank Sums Test.....	254
80.	Group Averages for Significant Differences Found in Articles, Demonstratives, and Pronouns.	255

81.	Number of Common Nouns, Proper Nouns, and Other NPs per Essay.....	257
82.	Average Number of Common Nouns, Proper Nouns, and Other NPs for Native and Student Writers.....	259
83.	Summary of Statistical Findings for Differences between Groups in Common Nouns, Proper Nouns, and Other NPs—Wilcoxon Rank Sums Test.....	260
84.	Group Averages for Significant Differences Found in Common Nouns, Proper Nouns, and Other NPs.....	261
85.	Summary of Statistically Significant Differences by Group.	266
86.	Summary of Group Averages for Significantly Different Findings.	268
87.	Summary of Number of Noun Phrases with Problems in Student Essays.....	272
88.	Noun Phrase Errors in the Essays of Native English Writers.....	274
89.	Error Percentages Compared with Other Judgments of Students’ Proficiencies.....	278
90.	Summary of Number of Noun Phrases in Student Essays in Co-reference Chains that have Problems.....	282
91.	Summary of Number of Noun Phrases in Student Essays NOT in Co-reference Chains that have Problems.....	284
92.	Comparison of Features of NPs in Generic Statements with the Over-use of <i>the</i>	291
93.	Predictions of L2 Article Use in English Using the Fluctuation Hypothesis and the Article Choice Parameter.	294
94.	Determiners in <i>a/an</i> Contexts.	296
95.	Mistakes in Contexts Appropriate for Zero Determiners.	297
96.	Examples of NPs Missing Definite Determiners.....	306
97.	Determiners used by Students in <i>the</i> Contexts.	308
98.	Number Mistakes in Noun Phrases in All Essays.....	312
99.	Summary of Common Error Patterns in Student Essays.	321
B1.	PennTreebank Tags for Part-of-Speech Tagging.....	325

LIST OF FIGURES

Figure

1.	The Prosodic Hierarchy of Goad and White.....	13
2.	Prosodic Structure in Turkish: The Article is an Affixal Clitic (a Prefix).....	13
3.	English Prosodic Structure: Articles are Free Clitics	14
4.	A DRS for the Sentences: Jones owns a car. He loves it.....	50
5.	A DRS for the Sentence: Mary owns the book.....	50
6.	Text Type, Viewpoint, and Type of Article.....	57
7.	Saunder’s (1999) Anaphoric Hierarchy.....	60
8.	Finite State Machine that Produces <i>prayed, played, pruned, and plumed</i> , but not <i>*pluned</i>	99
9.	Finite State Machine that Produces <i>prayed, played, pruned, plumed</i> , and <i>*pluned</i>	100
10.	Bayes’ Rule.....	101
11.	Iteration in Brill’s (1994) Transformation-Based Learning Tagging Algorithm ...	105
12.	Tree Structure Representation of the Initial Parsing of the Last Example.	116
13.	Tree Structure Representation of the Changed Parsing of the Last Example.....	117
14.	Tree Structure for the Meaning Highlighting the Timing of the “Elimination” in the Phrase “the elimination of general education requirements in today’s universities”	118
15.	Tree Structure for the Meaning Highlighting the “Today-ness” of the General Education Requirements in the Phrase “the elimination of general education requirements in today’s universities”	119
16.	Comparison of the Grammatical Reduced Relative Clause Structure and Learner Phrase Structure.....	129
17.	Test Statistic for the Wilcoxon Rank Sums Test.....	144
18.	Comparison of Extra Length of Extremely Long Chains in the Essays of Most Participants.....	163
19.	Student 7, A Mixture of Short and Long Co-reference Chains in the Same Writing.....	167

20.	Native 25, A Mixture of Short and Long Co-reference Chains in the Same Writing.....	168
21.	Student 9, Book-Ended Referents in the Same Writing.	170
22.	The ‘Teacup’ Pattern of Discourse in Student 9’s Essay—General-Specific-General Idea Presentation.	171
23.	Comparison of Student Essay Trends in Length	174
24.	Comparison of Native Essay Trends in Length.....	175
25.	A Zipf Curve Based on Word Distributions in Wall Street Journal Articles	230
26.	Comparison of the Percentage Use of Different Word Types, Native vs. Student Writers.	237
27.	Comparison of Native and Student Writer Use of Pronouns.....	246
28.	Comparison of Native and Student Writer Use of Demonstratives.....	248
29.	Percentages of Proper Nouns, Common Nouns, and Other Kinds of NPs in Student and Native Essays.....	262

LIST OF EXAMPLES

Example

1. Anaphoric Use of the Definite Article in L2 Learner Dialogue	23
2. Native Speaker Dialogue—No Over-Explicit Reference	25
3. Ionin, Ko, & Wexler’s Distinction between Specific and Non-Specific Definites.....	27
4. Lack of Cohesion in L2 English by a Native Japanese Learner.	32
5. Native Japanese L2 English Text under the Accessibility Hierarchy.....	34
6. <i>That</i> Making a Referent More Salient	39
7. A Story that is Not Fully Cohesive.....	42
8. Story Coded Using Centering Theory	45
9. Dialogue Showing Centering Theory’s Explanation of Pronoun Use is Too Limited.....	46
10. Effects of Time and Distance on Sentence Topic.....	46
11. L2 Learner Writing Coded with the Simple Transition Model	48
12. Samples of General Statives	53
13. Information Mode: Text Progression through Primary Referents	54
14. Criteria for Determining the Primary Referent.....	54
15. Samples of Abstract Entities.....	55
16. The Criteria for Identifying the Topic Phrase.....	56
17. Sample of a Description from Smith	58
18. Influences on L2 Learner Use of Null Subjects.....	61
19. Ways to Increase Overall Cohesion of a Narrative.....	62
20. The Category “Foreign Word” is Tagged by Intuition	92
21. PennTreebank List for the Part-of-Speech Category Modal Verb	93
22. Semantic Meaning in Tagging: Wh-Adverbs.	94
23. Two Sentences Illustrating Example Trigrams in Part-of-Speech Tagged Test or Model Corpus	98

24.	A Brill Transformation Rule Template.....	104
25.	Jurafsky & Martin (2000)'s Example of a Machine-learned Rule Used by a Transformation Based Tagging Program.....	106
26.	Sample of Part-of-Speech Tagging, Student 1.....	111
27.	The Same Sample with Tags Removed.....	111
28.	The Sample from Student 1 Again after Tag Correction.....	112
29.	A Fully Tagged Sample from Student 2.....	113
30.	The Same Sample with Corrected Tags.....	113
31.	The Sample from Student 1 with Tag Correction.....	114
32.	The Same Sample from Student 1 after Phrase Boundary Correction.....	114
33.	The Uncorrected Sample from Student 2.....	119
34.	The Same Sample from Student 2 with Corrected Phrase Boundaries.....	120
35.	NP + PP with Possessive Pronoun from Student 1.....	126
36.	An Extended Description by Native 28.....	127
37.	NP + Malformed Relative Clause from Student 12.....	128
38.	Malformed Structure from Student 19.....	129
39.	Malformed Structure from Student 19.....	130
40.	An Extended Description from Native 23.....	131
41.	Conjoined NP with Determiner by Native 26.....	131
42.	Three Conjoined NPs from Student 1.....	132
43.	Three Conjoined NPs from Student 1.....	133
44.	An Apposition from Student 2.....	133
45.	Noun Phrase Heads Cannot be Tagged Independently.....	134
46.	Heads of Compound Nouns and Collocations.....	135
47.	Names Cannot be Split When Tagging Co-reference.....	135
48.	A Pre-nominal Modifying NP that Refers Separately.....	136
49.	A Modifying Noun Used Elsewhere as a Head Noun by Student 15.....	137
50.	Predicate Nominal Following the Copula from Student 12.....	138

51. Post-Copula Nominal with Indefinite Determiner from Student 14.....	139
52. Predicate NP Following Have, from Student 5.....	139
53. Negated Predicates from Two Students.....	140
54. Co-reference Tagging of a Function.....	141
55. A Function in the Essay of Student 6	142
56. Sequence of Related Noun Phrases by Native 25 which are Not Coded as Co-referring According to MUC-7 Rules (given in order of appearance).	162
57. Syntactic Position Affects Anaphor Meaning Resolution.....	208
58. NP Subjects in Student and Native Essays	209
59. NP Verbal Objects in Student and Native Essays.....	210
60. NP Objects of Prepositions in Student and Native Essays.	211
61. NP Genitive Specifiers in Student and Native Essays.....	211
62. NP Complements of Copula Verbs in Student and Native Essays	212
63. NP Appositions in Student and Native Essays	213
64. NP Comparisons in Student and Native Essays.....	214
65. NPs in Titles in Student and Native Essays.....	214
66. Error in an NP Complement of a Copula Verb in a Korean Student’s Essay.....	218
67. Quantificational, Pronominal, and “Someone” Uses of <i>one</i> in Student 2’s Essay	250
68. Possible Indefinite Uses of <i>one</i> in Student 15 and Student 17’s Essays.....	251
69. Double Determiners in Student Essays.....	286
70. Names Incorrectly Used with <i>the</i>	287
71. Over-use of <i>the</i> in General Statements/ with Generic Reference	288
72. Optionality Shown in the Use of <i>the</i> with the Lexical Item <i>baby</i>	289
73. Over-use of <i>the</i> in Specific, Indefinite NPs	293
74. Over-use of <i>the</i> with a Non-Specific, Indefinite Referent.....	295
75. Superlatives are Definite.....	300
76. NPs Uniquely Identifiable Due to Shared Background Knowledge and Semantics.....	302

77.	Noun Phrase Description Uniquely Identifies a Referent/ A Native Writer's Mistake of Omitting <i>the</i>	303
78.	Omission of <i>a/an</i> by Student and Native Writers	304
79.	Over-use of <i>a/an</i> in <i>the</i> Contexts	307
80.	Pronoun Number and Referent Number do Not Match.....	316
81.	Word Order and Genitive Expressions in Noun Phrase Descriptions	319
82.	An Ill-Constructed Restrictive Relative Clause.....	320

CHAPTER I. INTRODUCTION

A Brief Discussion of Theoretical, Generative Linguistics, Corpus Linguistics, and First Language Acquisition

Linguistics is a field of study composed of many different pursuits, researching sounds, grammar, language learning, language history, conversational implicatures, turn-taking and more. Theoretical approaches that investigate the basic structure and generative nature shared by all human language rely on data, but it seems that the focus is on the breadth of data—what are the boundaries of what is possible or acceptable in a particular language or in all languages. In contrast, computational linguistics, and in particular corpus linguistics, seems to rely on frequency, or the depth of data. Knowing what is most common is highly valued, both for the existing practical applications of this knowledge (e.g., search engines, automatic parsing, speech recognition), but also for what can be understood from it about how language is actually used.

As a student of linguistics, I have studied language teaching, generative linguistic theory (phonology and syntax), language acquisition, historical linguistics, and a minute amount of computer programming and computational linguistics. In any discipline studied deeply, there comes the moment of transition from viewing the big picture to immersing oneself in the minutiae. In generative, theoretical explorations, questions can arise due to the reliance on grammatical judgments. For example, students question whose judgments should be considered correct. At times, the difference in some cases between saying a particular construction is ungrammatical (*) versus really strange (??##) can seem really small. In addition, with dialect and idiolect differences, what sounds bad or strange varies from person to person. It is also problematic that the context surrounding a particular statement makes a difference when deciding the grammaticality, especially because statements created to explain an analysis are often given without

context. However, when the grammaticality judgments of study seem questionable, the theoretical points made seem questionable.

Very abstract analyses lead to other concerns. Students often find it hard to accept multiple null elements in analyses, or proposals that seem counter-intuitive. For example, it is hard to support a hypothesis such as verbs derive from adjectives when adjectives feel much less important. These concerns lead to a desire for less abstract linguistic pursuits. When I reached this point, I stumbled into a class on corpus linguistics. Corpus linguistics focuses on frequency data and often practical uses of research. However, the need to simplify and often focus solely on what is most expedient for a particular purpose can move too far away from abstraction and a pursuit of understanding how language works. If trying to make an analysis with frequency data, a researcher can see what writers or speakers do, but not what meaning is attached to language or why. In addition, if a complicated method of analysis that is closer to what may be how humans mentally represent language does not get results significantly higher than a simplified process like just looking at what words are next to each other, then the complicated method is ignored in favor of the simple process. The need to get the most accurate results from the simplest method (but not the most accurate results overall) can lead to frustration. It may be easier for a machine to parse language in a certain way, but if this method is not something that can be approximated by humans without computers, it can seem unclear what is really being learned or discovered about human language.

First language acquisition researchers actually often use frequency data to inform their theoretical explanations. Children are often recorded speaking, and then the range of forms used and the frequency of certain structures are used to estimate what children of around that same age have as underlying structures or knowledge of their first language. Frequency data is in part used because young children cannot complete complicated tests requiring their grammatical judgments (it is at least very difficult to design such tests for certain language features). However, it is also used because

language learning involves optional use of forms and structures as new rules are internalized. Frequency data allows optionality to be examined. Frequencies are also important for studying language learning because how often a learner is exposed to a structure in the input can affect how they learn the use and limitations of the target language's structures.

Thinking of these kinds of considerations has led me to this study in which I use frequency data to look at the production of noun phrases in both native and learner writing. The frequency data cannot by itself indicate what students truly believe about the meaning assigned to the article and determiner forms in their interlanguage. Frequency data cannot indicate what the actual structure of the underlying form is. It is true that in the real world, language learners are often judged on what they can answer on tests about what they understand, but I would argue that they are judged most often by what they can produce in context under conditions of real language use. If a learner understands a structure, but cannot produce it in context, they will not succeed in the environs of the second language. Therefore, what their production looks like in comparison with that of native speakers is important. Typically, though, students are judged against the perceptions or preferences of only one or two native speakers (as with grammatical judgments—if one does not fully agree with the author's perspective, the findings of a study may be discounted much more easily). In this study, what students write is partly judged against what natives actually write. This is in addition error analysis based on one person's opinion. The frequency comparisons allow my individual grammatical judgments to be supplemented by the comparison of the *use* of several native speakers to the *use* of the students.

Getting back to the idea of production and errors, in contrast to students who understand the use but make production mistakes, if a student has meanings assigned to functional morphemes that are not native-like, but due to production factors, discourse influences, and semantic interpretations they perform fairly accurately in real-world

tasks, they will be successful. This second possibility may of course not be likely to result, but for the actual application of the knowledge of what students are doing, a consideration of production and not just underlying competence is necessary. Frequency data, while not perfect, is one way to examine the gross differences between the use of articles and determiners by native and student writers. When evaluating student writing, aside from looking at strange constructions, it is possible that readers' evaluations are influenced by frequencies of constructions compared to what is generally seen in native writing.

Aside from using the native writing as a standard against which to judge the frequency patterns in the student writing, the native essays allow some general guessing about the kind of input students might in part be exposed to. Textbook explanations of the grammar of article and determiner use in English are highly simplified. Rarely is a context of more than two or three sentences given. Complex, multi-clause noun phrases and sentence constructions are rarely used in examples or exercises. The role of demonstrative determiners and pronouns in the system of nominal reference in discourse is not explicitly explained in most textbooks. The question is whether this is a simplification of real native speaker use or not. If it is, this kind of simplified input might help students understand some of the semantic and discourse meanings of articles, but it would not really teach them native-like production. I think that the examination of native essays clarifies that the structures in the native writing are very complex, and the noun phrases are very long and more varied in style in general when compared with those of the students.

To summarize, using frequency data is not the perfect way to examine what students understand about the semantic and discourse meanings of determiners. Carefully constructed tests requiring grammaticality judgments are more accurate ways of doing this. However, frequency data comes from students' own production, which test data does not. Furthermore, error analysis can add another dimension to such frequency

data, and looking at the frequency data of native speakers provides some insight into what students' input might look like, or against what standard students are truly being judged by other native speakers. The next section begins to explain more specifically the theories and studies from which this project developed.

Review of the Literature and Prior Studies

Second language (L2) learners notoriously have trouble using articles¹ in their target languages, even if their native language also has articles. This is extremely frustrating to adult students of languages, and can have repercussions in the workplace, at school, and in other arenas in which complete grammatical accuracy is highly prized and non-native speakers are judged against their native speaker counterparts. Adult learners seem to have difficulty acquiring accurate usage of articles on their own, though, and current teaching methods have limited success in eradicating production errors. In order to improve these methods, the cause of the errors and the patterns of errors have to be better understood.

Past linguistic studies have found a variety of error patterns in L2 article use in English. Studies have found that learners:

- predominantly omit articles (see, for example, White 2003 or Robertson 2000),
- tend to overuse *the* (see, for example, Huebner 1983, Master 1987, Parrish 1987, Tarone & Parrish 1988, Thomas 1989, or Ionin 2003, Ionin et al. 2003), or
- tend to overuse *a* (see Leung 2001).

Different theoretical explanations have been put forward to try to account for these different patterns. Two common hypotheses are:

¹ Throughout this paper I will differentiate between the term *articles* and the term *determiners*. *Articles* will be used to refer to morphemes parallel to *the* and *a/an* in English. *Determiners* will be used to refer to morphemes parallel to *the, a/an, this, that these, those, my*, etc. in English.

- learners have incorrect or incomplete semantic representations linked with articles (see, for example, Tarone & Parrish 1988, Hawkins & Chan 1997, Hawkins & Liszka 2002, Goto Butler 2002, Ionin 2003, Ionin et al. 2003), or
- learners have complete and correct semantic representations for articles, but trouble choosing the right form during production due to stress on mental processing or phonological limitations (see, for example, Lardiere 1998, Bruhn de Garavito & White 2000, Prévost & White 2000, White 2003, Goad, White, & Steele 2003, Goad & White 2004).

The main concern regarding the conclusions of these previous studies is that they have focused solely on articles. Articles identify discourse relationships, but other morphemes do so as well, such as pronouns and demonstratives. Articles, pronouns, and demonstratives all communicate how nominal referents relate to the previous discourse—whether referents are old or new, whether they should be accessible in memory, and other relationships. Although not all languages have articles, all do have some morphological reflexes to express these discourse relations. It is therefore important to ask: Is looking only at article use in a study sufficient?

Researchers commonly search for the transference of patterns from the first language (L1) to the L2. Is it possible that only looking at patterns of article use is obscuring indications of transfer occurring or not?

Due to these considerations, this project will explore the use of articles and other discourse morphemes in L1 and L2 English essays. A variety of different theories of how to classify and explain discourse meanings will be used in order to determine where there may be significant differences between the L1 and L2 methods of reference and what might be the root of production errors.

Common Hypotheses Regarding Article and Determiner Acquisition

This section compares several hypotheses that have been proposed to account for article errors and why they linger in the interlanguage of advanced adult L2 learners. These theories have been divided into groups based on what they have in common. First, studies relying on Feature-based Discourse Theories will be discussed. Feature-based theories describe the meaning given to articles in the discourse with semantic features, such as \pm definite, that must be checked in syntax. These theories can be further divided into two types—theories that strive to explain lingering errors in many kinds of functional morphology, and theories that focus solely on articles. Three Functional Morphology Error Theories will be examined here:

- The Failed Functional Features Hypothesis (Hawkins and others),
- The Missing Surface Inflection Hypothesis (Haznedar & Schwartz, Prevost & White and others), and
- The Prosodic Transfer Hypothesis (Goad & White).

Two Theories Focusing Solely on Articles will be contrasted with these:

- Discourse Rule Transfer (Robertson), and
- The Article Choice Parameter and the Fluctuation Hypothesis (Ionin, Wexler, and others).

After examining these, a more in-depth discussion of Feature-Based Discourse Models will be provided in preparation for looking at theories and studies using Alternative Models of Discourse.

Feature-based Discourse Models

As mentioned, feature-based theories describe the discourse and semantic meanings of articles with features that must be checked in syntax such as \pm specific. Two types of these theories will be examined—some looking at all kinds of functional

morphology and some only looking at articles. In general, it is concluded that although these theories make insightful claims about how L2s are acquired, their foundation models of discourse are limited and even faulty in some respects.

Functional Morphology Error Theories and Studies

The three theories described below are intended to be applicable as explanations for all/any L2 functional morphology errors. They have in common the use of generative L2 acquisition theories such as L1 transfer, access to Universal Grammar (UG) after the critical period, optionality, phonological interference, fossilization, and more. However, these theories differ in how these are seen to play out and result in the surface forms produced by learners.

Functional Morphology Error Theory 1: The Failed

Functional Features Hypothesis—Hawkins and Colleagues

Hawkins & Chan (1997), Hawkins (2000), and Hawkins & Liszka (2002) seek to explain why article errors linger for adult L2 learners with the Failed Functional Features Hypothesis (FFFH), which is based on work by Smith and Tsimpli (1991, 1995). The FFFH states that adult L2 learners have impaired meaning. They cannot build the semantic representations for articles that need to be checked in syntax. This meaning impairment results because adults do not have full access to UG, rather their only access is through their L1 settings. Therefore, the L1 greatly influences learners' interlanguage through transfer of features and parameters, and if the L1 lacks certain functional features that need to be checked in syntactic representations, L2 morphological errors result. In cases in which the L2 has articles but the L1 has no articles, learners will lack a semantic element that they need to map to the L2 lexical item, such as the feature [+definite], and their use of the L2 articles will therefore never be fully target-like in the underlying structure (Hawkins & Chan 1997:199).

Adult L2 production may approximate the target surface structure, though, despite missing a relevant feature or employing an L1 parameter setting. This is because:

Some other operation which is not parameter resetting must be involved in producing the observed restructuring of the learner's grammar away from the L1 and towards the L2. (Hawkins & Chan 1997:200)

In other words, adult learners can rely on other cognitive learning skills to approximate target structures. White (2003) takes issue with this stipulation of the FFFH by stating that it results in there being no evidence that could disprove the FFFH. The problem is: When learners fail to produce appropriate L2 morphology due to differences between the L1 and L2, this is evidence for the FFFH. At the same time, when these same learners *do* have L2 target-like morphology, some other cognitive learning technique could be involved that is unrelated to underlying features. This is not evidence against the FFFH.

Being thus opposed to the Failed Functional Features Hypothesis, White has examined or proposed alternative theories to account for the fossilization of functional morphemes like articles and past tense markings in adult L2 production. Two of these will be described in the next two sections.

Functional Morphology Error Theory 2: The Missing Surface Inflection Hypothesis

The Missing Surface Inflection Hypothesis (MSIH) (Haznedar & Schwartz 1997, Prévost & White 2000) differs from the FFFH in that it does not claim that learners have incomplete or faulty semantic representations. Adults have full access to UG, and are not limited to only those features and parameter settings of the L1. Therefore, underlying syntactic structures are correct, and the necessary semantic features that need to be checked are present and accurate. However, the resulting surface functional morphology is not target-like due to problems mapping the correct morphology onto the feature representation in syntax. The MSIH thereby does not deny that there could be a great

deal of L1 transfer in L2 production. These mapping problems are caused by an overload of the learner's mental processing system.

In Prevost & White (2000), Distributed Morphology (DM) (Halle & Marantz 1993) explains the mapping of morphology to the syntactic/semantic structure. It is stated that learners have a different rule for what morphology can match the feature set than native speakers. Prévost & White (2000:127) explain that for L2 learners: "The features of the lexical item do not need to exactly match all the features of the hosting node: it is sufficient that they form a proper subset of the feature bundle of that node." So, learners have several functional morphemes to potentially match to any feature set it.

The MSIH is supported when L2 production shows:

- missing or incorrect functional morphology (past tense, determiners, grammatical gender, etc.), and
- correct verb or adjective placement, use of plurals, or selection of nominal pronouns.

These co-occurring characteristics are important because they should be affected by the same functional features. If features are missing or incorrect, then learners should not be able to accurately produce these other structures.

Bruhn de Garavito & White (2000)² conclude that their data supports the MSIH because the learners' L1s are irrelevant to predicting their errors. The FFFH predicts that French speakers learning Spanish should more accurately produce gendered morphemes than English speakers learning French because both Spanish and French have grammatical gender, while English does not. Overall, Bruhn de Garavito & White found:

² Looking at nominal inflection and articles, Bruhn de Garavito & White (2000) compare their data of native French speakers learning Spanish to data from Hawkins (1998) of English speakers learning French. Both studies look at the order of nouns and adjectives and gender marking on determiners, while Bruhn de Garavito & White also compare gender on adjectives and nouns.

L2 learners at higher levels performed better. Proficiency was more important in determining errors than L1.

In fact, learners produced accurate gender morphology almost 80% of the time on average, although it was found that for some reason they were more accurate in producing gender agreement for definite article phrases than indefinite article phrases (Bruhn de Garavito & White 2000:170-1). This last finding is problematic because it cannot be explained by either the MSIH or the FFFH. The MSIH has also been criticized as being a post-hoc solution. White (2003:139) explains this criticism in this way:

For example, such proposals do not predict inevitable variability in suppliance of overt L2 morphology but seek only to account for such variability as is found. This contrasts with the position arguing for grammatical impairment (e.g., Hawkins 2000, 2001), where problems of overt L2 morphology are predicted in those cases where the L1 and L2 differ as to which abstract features are represented in the grammar.

The pattern of usage by L2 learners is not fully explained due this lack of predictive power. Another potential issue with the MSIH is that the causes, characteristics, and co-effects of “a mapping or processing problem” are not clearly defined.

Lardiere (1998, 2000) proposes a theory similar to the MSIH in which surface morphology is inaccurate due to mapping and not underlying structure errors. The same criticisms as above apply to the mapping problem Lardiere describes. These criticisms have led both White and Lardiere to seek the source of errors at the interface between syntax and phonology, as will be explained in the next section.³

³ It will not be discussed explicitly, but in Lardiere (2003), she provides a phonological explanation for an adult L2 learner’s production of functional morphology. The main difference between Lardiere’s phonological theory and the one to be discussed is that it looks at the possibility that transfer of L1 segment clustering constraints, and not transfer of L1 prosody, limits what L2 functional morphology can be produced.

Functional Morphology Error Theory 3: The Prosodic Transfer Hypothesis

The Prosodic Transfer Hypothesis (PTH) (Goad, White, & Steele 2003, Goad & White 2004) hypothesizes that the transfer of L1 phonological constraints causes learners to inaccurately produce L2 functional morphology. Specifically, when L1 rules for prosodifying function words and inflection differ from the L2's prosodic rules, learners will either be unable to produce the L2 morphology or will variably omit it. However, they will produce the morphology when it is possible to use valid L1 prosodic structures to pronounce it (see Goad, White, & Steel, 2003:254). Like the MSIH, the PTH claims that adult learners have access to UG, and that learners' underlying syntactic structures are accurate and complete.⁴

Goad & White (2004) explain how the PTH works by looking at the production of a native Turkish speaker learning English. The PTH predicts Turkish speakers should have difficulty producing English articles because⁵:

- Turkish is a zero or one article-language and not a two article-language like English. The word *bir* 'one' is used as an indefinite article when unstressed.⁶
- The prosodic structure used to pronounce this Turkish article is not the same as that used in English to pronounce articles.

These differences mean that learners' only recourse to attain proficiency close to full attainment and avoid fossilization would be to 'minimally adapt' L1/Turkish prosodic

⁴ It is important to note that the PTH does not claim that the differences in prosody between the L1 and the L2 would have some kind of filtering effect resulting in impaired syntactic structure. See Goad & White 2004:178.

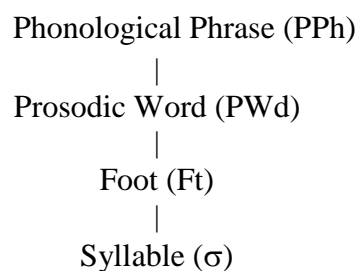
⁵ The FFFH differs in that it would not predict difficulties in Turkish-English interlanguage because Turkish does have articles.

⁶ There is some disagreement about this fact, but this is the stance that Goad & White assume.

structure to produce the L2/English morphology (Goad & White 2004:181). To ‘minimally adapt’ L1 prosodic structure is to use L1 structure when it is not licensed in the L1, or to combine L1 structures in ways not licensed in the L1 to create a target-like L2 prosodic structure.

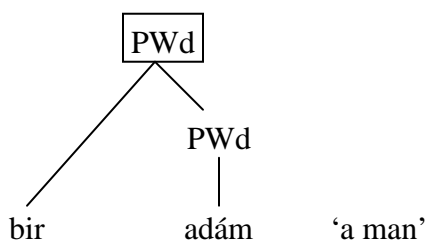
The prosodic hierarchy Goad and White assume is in the following figure.

Figure 1: The Prosodic Hierarchy of Goad and White (Goad & White 2004:180)



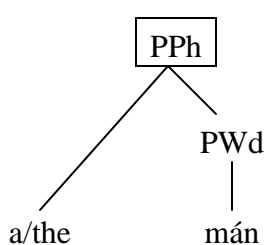
When prosodic structures are being built, the least marked are those in which elements are only dominated by the next type up in the hierarchy, with binary branching.

Figure 2: Prosodic Structure in Turkish: The Article is an Affixal Clitic (a Prefix)



Language specific rules, though, determine if marked structures are allowed, when they are possible, and how they are constructed. Goad & White propose the structures that follow for Turkish and English articles. Turkish articles are described as prefixes, and English articles as free clitics.

Figure 3: English Prosodic Structure: Articles are Free Clitics



Goad and White use vowel harmony and word order in Turkish to argue that unstressed *bir* is adjoined to the lower prosodic word. For English, word order and stress are used to argue that articles must be attached to the phonological phrase, and cannot be adjoined to the lower prosodic word.⁷ Goad and White claim that Turkish does not allow free clitics at either the right or left edge of words, so Turkish learners cannot produce a structure such as that in (3) due to prosodic transfer from their L1. Therefore, Turkish learners should have a tendency to omit articles in English.

However, the fact that the PTH includes the option of minimally adapting L1 prosodic structure for L2 production confuses the situation. The claim that articles will be omitted is not well supported by the data Goad and White collected. The speaker in the case study (SD) omits articles only in 32-34% of all the contexts in which Goad and

⁷ See the arguments in Goad & White 2004:183-184 for further explanation.

White deem they are obligatory (Goad & White 2004:180). So, in the majority of cases, SD produces articles. Goad & White state that SD could be ‘minimally adapting’ Turkish prosody to produce English articles in three possible ways: (i) SD could be adjoining articles directly to the prosodic word as with Turkish *bir*, (ii) SD could be placing functional morphology within the lowest prosodic word, as is done for past tense morphemes in both English and Turkish at the right end of the word, or (iii) SD could be treating articles like stressed determiners, placing them in their own prosodic word, not simply adjoining them.

Goad and White only claim to have found evidence for the first option.⁸ As an affixal clitic, *bir* cannot be separated from the noun, so it cannot precede the adjective in a phrase with both a noun and an adjective. The Turkish order would be: adjective-unstressed *bir*-noun. This means that the learners should have trouble producing articles when an adjective appears in the noun phrase. Goad and White in fact did find significantly fewer *indefinite* articles produced before an adjective, but two problems remain. First, the PTH cannot explain why SD showed a difference in her use of indefinite versus definite articles. Second, the PTH does not discuss how learners move from their L1 prosodic structures to L2 ones, or how this is possible. For example, how does SD learn to produce any articles in the adjective-noun context if the English structure is not initially allowed and she must rely on her L1 prosodic structures?

Some other issues also remain unresolved. For instance, the PTH only explains omissions of determiners and other inflectional morphology, while it has commonly been found that L2 learners of English overuse *the* (see, for example, Huebner 1983, Master 1987, Parrish 1987, Tarone & Parrish 1988, Thomas 1989, or Ionin 2003, Ionin et al.

⁸ The option in which articles are stressed is ruled out by the fact that SD did not produce any stressed articles. There was not enough data of a particular type to make any determination about whether SD used the second option.

2003). It is not clear why *the* would be easier to produce than *a* when both morphemes would have to be represented similarly in the prosodic structure.⁹ Furthermore, if phonology is the source of errors, and underlying competence is unaffected, what can be said about learners omitting functional morphology in writing? Hawkins & Liszka's (2002) study found errors in the *written* production of verb forms. Goad and White (2004) specifically state that phonology does not act as a filter on the acquisition of syntax or semantic features. For these reasons, although the PTH seems promising, it is unclear that it explains determiner errors effectively.

Functional Morphology Error Theories: Conclusions

The FFFH, MSIH, and PTH take various perspectives on some of the key concepts of generative L2 acquisition, such as access to Universal Grammar and first language transfer. The table that follows summarizes these differences. All three theories try to account for apparent fossilization and optionality in L2 production. Unlike the FFFH, the MSIH and PTH specify that full access to UG (not just what is instantiated in the L1) is possible for adults. Because study findings vary so widely, no definitive choice can be made amongst these theories. Each one is supported by some evidence and refuted by other evidence, and each one has its strengths and weaknesses. See the table that follows.

One aspect shared by the FFFH, MSIH, and PTH that is not shown in the tables is that they all rely on the same underlying model of discourse.

⁹ There are obviously differences in the markedness of the syllable types and individual phonemes of these morphemes, but these types of phonological issues are not discussed by the Prosodic Transfer Hypothesis.

Table 1: Summary of Theoretical Perspectives of Functional Morphology Error

Theories

<i>Theory</i>	<i>Adult access to UG?</i>	<i>Influence of the L1?</i>	<i>Fossilization? Ultimate Attainment?</i>	<i>Optionality present on surface?</i>
FFFH	no	great influence	may be able to mimic target language, but will never have accurate underlying structure	yes, learners will show L1-type form, or else mimic L2
MSIH	yes	influence only apparent in beginning stages	full attainment possible, but surface structures influenced by processing problems	yes, many forms can be selected to match one underlying representation
PTH	yes, for syntax (unclear for prosody)	great influence in the prosody, none in syntax or semantic features	may be able to mimic L2 prosodic structure, but if not present in L1, it's likely to fossilize	yes, but no clear explanation of why this happens

Table 2: Strengths and Weaknesses of Morphological Error Theories

<i>Theory</i>	<i>Strengths</i>	<i>Weaknesses</i>
FFFH	explains differences in article use by native speakers whose L1 has no articles vs. those whose do	no evidence can disprove the theory
MSIH	can explain misuse of forms, not just omissions	post-hoc solution
PTH	explains the learning process and acquisition of the use of morphology	cannot explain misuse of article forms, unclear what it predicts for writing

Discourse relations are expressed as binary semantic features (such as $[\pm\text{specific}]$) that need to be checked in syntax. This study will argue that such a model is undesirable. This will be discussed at the end of this Feature-based Discourse section. The next

section looks at theories focusing only on articles and not multiple types of functional morphology.

Studies and Theoretical Explanations Focusing Only on Articles

This section will look at two studies using generative second language learning concepts to hypothesize the cause of *only* L2 article errors and not all kinds of functional morphology. Robertson (2000) focuses on the transfer of discourse rules as a source of article errors. Ionin (2003) and others explain that the setting of the Article Choice Parameter and the Semantic Fluctuation Hypothesis cause article errors.

Article-Focused Theory 1: Discourse Rule Transfer—

Robertson (2000)

Robertson (2000) uses a rule-based approach to explore L2 errors. L1 transfer is hypothesized to be the cause of L2 article errors, but no clear statement about whether adult L2 learners have access to UG is made, and the rules given are language-specific and not universal. Furthermore, because Robertson's main goal is examining when articles are omitted, error rates for the misuse of those articles supplied are not discussed.

Robertson's rule-based analysis and classification has two or three parts. First, an English-specific classification system is used to label syntactic or background knowledge contexts in which the use of *the* or *a/an* is obligatory. What kind of mistakes L2 learners made in each context was then evaluated. As part of this, article use was further classified by what Robertson describes as pragmatic contexts—echo and non-echo situations. In echo situations, what was just said is repeated for clarification, but L2 learners may not produce an exact copy and sometimes omit an article in the repetition. Because the focus is on specific English syntactic environments, it is not clear what Robertson's results from this part of the analysis say about L2 acquisition in general. Not all languages have articles, nor do they use articles in the same syntactic positions.

In the second part of the analysis, Robertson uses a set of Chinese-specific discourse rules to analyze L1 transfer by Chinese learners of English. These rules are listed in the table that follows. These rules are said to explain most (but not all) of the situations in which the native Chinese speakers omitted or misused determiners in their L2 production. No theoretical explanation of why these rules would be present in Chinese while not present in English is given, although it is commonly acknowledged that Chinese and English differ in regards to pro-drop and anaphors. It is also not overtly discussed what other languages would be predicted to have or not have similar rules, although again it may be assumed to be related to pro-drop and anaphors. These discourse rules, again because of their language-specific nature, cannot be easily used to predict errors when other L1-L2 combinations are involved.

Table 3: Robertson's Discourse Rules (Robertson 2000:135)

#	<i>Discourse Rule</i>
(i)	a syntactic principle of 'determiner drop', whereby an NP with definite or indefinite reference need not be overtly marked for [\pm definiteness] if it is included in the scope of the determiner of a preceding NP
(ii)	a 'recoverability' principle, whereby an NP need not be marked for [\pm definiteness] if the information encoded in this feature is recoverable from the context; and
(iii)	a 'lexical transfer principle', whereby some of these learners are using demonstratives (particularly <i>this</i>) and the numeral <i>one</i> as markers of definiteness and indefiniteness respectively.

The next study also focuses solely on articles, but unlike Robertson's rule-based description, it focuses on cross-linguistic applicability and predictive power.

Article-Focused Theory 2: The Article Choice Parameter
and the Fluctuation Hypothesis—Ionin, Wexler, and
Colleagues

Ionin (2003), Ionin, Ko, & Wexler (2003), and Ko, Ionin, & Wexler (2004) propose that L2 article errors result from faulty underlying structure, similar to the FFFH discussed previously. However, unlike the FFFH, Ionin, Ko, and Wexler believe that adults have access to UG. The problem is that learners cannot properly set the parameters they can access in UG, resulting in surface errors. The parameters cannot be properly set due to the Fluctuation Hypothesis, which Ionin, Ko, & Wexler propose for L2 grammars. Individual learners fluctuate between parameter settings, causing them to improperly select the semantic features that must be checked in the syntax, and resulting in selection of the wrong functional morphology.

The Fluctuation Hypothesis is general to all parameters, but article errors result because learners cannot set the semantic parameter for articles proposed by Ionin, Ko, and Wexler that states that any language that has two articles (parallel to *the* and *a/an*) will have those articles either distinguish definiteness or specificity. Articles in English, for example, are seen to express definiteness (*the*) and indefiniteness (*a/an*). In contrast, Samoan articles distinguish specificity, with the article *le* used for specific referents and the article *se* used for nonspecific referents (Ionin, Ko, & Wexler 2003:4). L2 learners then alternate between using L2 articles to distinguish specificity or definiteness. For English learners, this leads to the predictions of possible errors shown in the table that follows.

One problem with this theory is that although it explains mistakes in usage, it does not make predictions for omissions of articles. Also, unlike the PTH, it does not seem to predict differences between proficiency levels of learners.

Table 4: Predictions of L2 Article Use in English Using the Fluctuation Hypothesis and the Article Choice Parameter (Ionin, Ko, & Wexler 2003)

<i>Context</i>	<i>Definite referent: target the</i>	<i>Indefinite referent: target a</i>
<i>Specific referent</i>	correct use of <i>the</i>	overuse of <i>the</i>
<i>Non-specific referent</i>	overuse of <i>a</i>	correct use of <i>a</i>

Questions about the Theoretical Foundation of Binary

Feature-Based Discourse Models

The theoretical foundations of the studies presented above use binary features such definite and indefinite to describe the meanings of articles, although there are two commonly used perspectives on this. One is that of Hawkins (1978), and the other focuses on Feature Checking in Syntax.

Hawkins (1978)

Robertson (2000) uses a discourse model that has three parts:

- language-specific syntactic or background knowledge contexts in which the use of different articles is obligatory,
- pragmatic contexts—either echo or non-echo situations¹⁰, and
- language-specific discourse rules.

The first part is typical of the discourse classification systems used in many studies, and is based on Hawkins (1978). The focus is on specific syntactic environments in which

¹⁰ Echo situations are those in which learners repeat what was just said for clarification, but may not be exactly the same as what was said. For instance, a learner may hear a phrase with an article, but repeat an approximation of the phrase without one. This is a limited way to look at discourse primarily because echo contexts are rare outside of certain kinds of situations, such as giving spoken instructions.

particular articles tend to occur. The meanings of articles are described by the binary feature \pm definite (definite for *the* and indefinite for *a/an*).

Table 5: Robertson's Taxonomy of Determiner Use (Robertson 2000:145-149)

<i>Definite NP environments (+ definite)</i>		
D1	Anaphoric use of referring NP	referent was used before with indefinite article
D2	Immediate situation use of referring NP	referent present in immediate situation and existence is known by speaker and hearer (ex., <i>the red pen</i>)
D3	Larger situation use of referring NP	referent uniquely identifiable due to shared background knowledge (ex., <i>the left hand side of the paper</i>)
D4	Head noun of an associative clause NP	referent in clause of two NPs joined by <i>of</i> (ex., <i>the bottom of the sea</i>)
D5	Unexplanatory use of definite NP	ex., <i>the same N, the first N, the best N</i>
D6	NP with nominal modifier	ex., <i>the letter A, the number 3</i>
D7	NP with establishing relative clause	referent is followed by relative clause (ex., <i>the first line that you drew</i>)
<i>Indefinite NP environments (-definite)</i>		
I1	Use of NP in existential predication	referent is stated to exist as object of <i>there is/are, have, got</i>
I2	Use of NP as object of transitive verb or complement of copulative construction	ex., <i>Then you draw a horizontal line., Is it a big one?</i>
I3	Generic use of singular NP	ex., <i>It's square, like a floor, you know?</i>

Because the focus is on specific syntactic patterns, these classifications cannot be easily compared with or extended to other languages. Not all languages have articles, nor do they use articles in the same syntactic positions. Robertson focuses on articles,

considering the use of *one* and demonstratives in passing. No overt discussion of features such as \pm specific is made, although some of the specifications discuss background knowledge and previous mention in discourse. The version of Hawkins that Robertson (2000) uses appears in the table on the previous page.

This list of contexts, aside from being specific to English, raises concerns because, in its limited scope, it does not model natural noun phrase use completely accurately, and it cannot be used to detect differences in acceptability of reference use. To explain this concern, the category “anaphoric use” for the definite article will be examined in more depth. Tokens in this category indicate that the English learner recognizes the intended referent because the referent was previously introduced in the discourse by an indefinite noun phrase (Robertson 2000:145). Robertson’s example of anaphoric use from the speech of actual L2 English learners is italicized in the next example. Previous mentions with the indefinite article are underlined.

Example 1: Anaphoric Use of the Definite Article in L2 Learner Dialogue

- A: And then after that you draw a square with the red...
 B: Square?
 A: Yeah, a square.
 B: What does *the square* draw like? (Robertson 2000:145)

The italicized “the square” is appropriate according to Robertson because the referent has been mentioned before in the discourse (in each of three prior turns). The question is that although a native speaker might find “the square” acceptable, would they perhaps prefer something else? Robertson himself brings this up and states, “Notice that a demonstrative (*this square*) would be perfectly acceptable in this context, although perhaps the definite article would be preferred (Robertson 2000:145).” In fact, alternative models of discourse predict that an unstressed overt pronoun would be the

best way to represent this referent in English, since it is the center of attention due to having been the only nominal referent repeated in the last three utterances. These other theories do not deny that either *the square* or *this square* could be appropriate, but this would only be under extenuating circumstances—such as there being two same gender, same number referents being centered on at the same time, which is guaranteed to cause confusion without being overly explicit. This shows that the Robertson/Hawkins model is not looking at some possibilities for production or considerations for usage. The model does not look at the use of pronouns.

In fact, the kind of over-explicitness seen in the example above (use of an article + noun instead of a pronoun) has been commonly found in other L2 studies (see, for example, Saunders 1999, Chini 1998, Ahrenholz 1998, Muñoz 2000, Hendriks 2003).¹¹ However, this is something that Robertson's taxonomy fails to recognize, thereby showing this model to be imprecise and obscuring some differences between L1 and L2 production. In the native English speaker example from Brennan (1998) below, there is no over-explicitness of repeated referents without a cause. Brennan had one speaker describe a taped basketball game to another who could not see the TV, and who had to note information about the players and the score. In this part of the game, there were two players being described—number 42, and Griffin (number 20).

Looking at the noun phrases in italics in speaker B's first turn, *Griffin* is repeated even though speaker A just said it. At this point, though, being overly explicit clarifies the nominal reference since the dialogue is referring to two male basketball players. Using *he* could be confusing. Unlike the example from Robertson's excerpt, there is a reason to be more careful in specifying the nominal reference. Furthermore, speaker B follows this reference with a pronoun—normal for elements that are center of attention.

¹¹ Over-explicitness here means using a full noun form instead of a pronoun, giving more information than a listener would need to identify the referent.

Example 2: Native Speaker Dialogue—No Over-Explicit Reference

- A: passing it to forty-two₁,
 who₁'s drive- making a drive,
 Ø₁ missed and was fouled by twenty, Griffin₂
- B: *Griffin*₂—When did *he*₂ come in
- A: umm
 uh huh huh huh, didn't notice
 two, *he*₂ has two personal fouls on *him*₂,
*Griffin*₂ does
- B: OK
- A: yah, *he*₂ jus-s-s slapped his₁ arm
 they've shown the replay twice now
 so they took Griffin₂ out and put in fifty-four
 (Brennan 1998:242)

Speaker A's second turn continues with the same topic, and a pronoun is used to refer to the player Griffin. However, since speaker A realizes there is potential for confusion, he ultimately clarifies with the more explicit name of the player. Again, there is no obvious reason for the speakers in Robertson's example to use a full noun phrase rather than a pronoun as there is no other item which could be confused with the square.

Ariel (2004) presents another argument for models of discourse that allow for the examination in differences of explicitness of reference over describing 'obligatory' contexts. She explains that psycholinguistic research has not advanced far enough to definitively state what level of abstraction speakers actually use when interpreting and communicating discourse relations. Speakers may rely on syntactic patterns, or on more abstract concepts such topic or focus, or both at once when determining how to linguistically express nominal reference. Therefore, Ariel advocates evaluating multiple model-types until a more definitive answer is reached. Many prior studies, though, simply rely on language-specific, context-specific classifications such as those of Hawkins (1978) and Robertson (2000).

Another concern with the Hawkins/Robertson methods of classification is the reliance on binary features (such as definite and indefinite) to describe the semantic and discourse meanings of articles. This will be discussed in detail in the next section.

Binary Features Checked in Syntax

The majority of current theories on articles—like the FFFH, MSIH, PTH, Hawkins' taxonomy and the Article Choice Parameter discussed above—rely on a model of discourse using binary semantic features checked in syntax, such as [\pm definite] and [\pm specific], to represent discourse relations. Accepting this idea requires accepting the idea that the meanings expressed by articles is very easily defined and perhaps that it is computed on the individual sentence level.

To examine this issue in more detail, the binary features model chosen by Ionin, Ko, and Wexler will be described. Instead of relying solely on the features \pm definite, Ionin, Ko, and Wexler (2003) adopt the features [\pm definite] and [\pm specific],¹² but specificity is defined differently than by most other researchers using these features. Ionin, Ko, & Wexler (2003) explain that they follow Fodor & Sag (1982) in believing that when an NP is specific, this entails that the speaker intends to refer to the real world referent—not announce its existence. Thus, unless speakers *know* the referent themselves—have seen, touched, or experienced the specific instance—the referent is *not* specific. Ionin, Ko, and Wexler highlight this distinction because many others using this framework would question whether it is possible to have a [-specific, +definite] referent, but the existence of this referent type is essential for the Article Choice Parameter to work. The examples below (from Ionin, Ko, & Wexler 2003 and Ko, Ionin, & Wexler 2004) are intended to clarify what a [-specific, +definite] referent would be.

¹² Some describe these features as primitive, in other words there is only, for example, [+definite] and the lack of this feature. Others use binary features. For consistency and ease of description, I will mention these always as binary features.

Example 3: Ionin, Ko, & Wexler's Distinction between Specific and Non-specific

Definites

- (a) [+specific, +definite]: I want to talk to the winner of this race—
she is my best friend.
- (b) [-specific, +definite]: I want to talk to the winner of this race—
whoever that is.

In (a), since the speaker actually has met the winner before, it is possible to intend to refer to the real world referent. On the other hand, in (b), since the speaker has not met the winner before, it would not be possible to intend to refer, and specific referents cannot merely announce existence.

It is not clear, though, that making this distinction for referents with the definite determiner is desirable. Do speakers of English feel there is a distinction in meaning? Is this distinction more explicitly expressed in other languages? Ionin argues that this distinction is codified in Samoan and certain other languages. However, other researchers have described the meaning of *the* as something like 'there can be only one thing you are referring to.' In both of the sentences above, the speaker knows there is only one person who fits the description *winner of the race*, regardless of whether he or she has met this person before. If this second view is a more accurate description of the meaning of *the*, there is no difference in specificity between these two sentences. Both refer to one specific person. It is generally agreed upon that the referents of noun phrases co-occurring with *a/an* can be either specific or non-specific, but this is proven by examining the distribution of the indefinite, but referential, demonstrative *this*. No such argument can be made in regards to *the*, though, and it seems that the only reason to state that both specific and non-specific nominal referents can follow *the* is to set the stage for Ionin, Ko, & Wexler's Article Choice Parameter. In fact, Ionin, Ko, & Wexler's (2003) results fail to support this distinction. They predicted that there should be an overuse of

a/an for [+definite, -specific] referents, since this would show that the specificity setting of the parameter was being used. This setting would be demonstrated by non-specifics (definite or indefinite) being represented by *a/an*—in a non-target-like fashion, but one that still would correspond with UG. However, this predicted overuse was unattested in their study.

This problem discussed above is specific to Ionin, Ko, and Wexler's analysis, and does not apply to all studies using binary features to represent discourse relations. However, one problem that does apply for any use of the binary features model of discourse is that determiners other than articles, and other linguistic means of describing nominal reference, are excluded or treated as unimportant. This is despite the fact that researchers using this model of discourse do not really show that they believe that these other forms are unimportant. For example, the Article Choice Parameter is defined as only being relevant for two articles and two article languages, but Ionin, Ko, & Wexler (2003) use the indefinite use of the demonstrative determiner *this* to explain how *a + N* could represent either a specific or non-specific indefinite referent.

Even if other determiners are considered relevant, it is not clear how they could be represented differently from articles without adding more features. If both *this* and *a* can be used to indicate a specific indefinite referent, how could they be distinguished with these features? Both would be labeled as [+specific, -definite]. Ionin, Ko, and Wexler cite research showing there is a distinction between *a/an* and indefinite *this*. Indefinite *this* is most often used to introduce a nominal referent that a speaker intends to refer to again shortly (in the next few clauses). This shows that there is a distinction between the discourse meaning relayed by these determiners, but it is one that cannot be described by only two features.

The FFFH (Hawkins & Chan 1997, Hawkins 2000, Hawkins & Liszka 2002) also causes concerns related to the small number of features in this model. Evidence for the FFFH is examined by comparing Chinese and English. These differ in that Chinese does

not have a lexical item that correspond directly to *the*, but Chinese does have demonstrative determiners. Under the FFFH, if the lack of *the* is explained by stating that Chinese lacks the feature [+definite], then what feature would be checked by a definite demonstrative, such as *zhèi* ‘this’ or *nèi* ‘that’? On the other hand, if Chinese learners do have access to a feature [+definite] in their L1, then what explains their article errors? The need to indicate how to resolve nominal reference is universal¹³, so perhaps this means that certain features have to be universal. Definiteness may be one of these, but if this is stipulated, then some other feature would have to be used to explain article errors and differences in the existence of articles in different languages.

There are still other issues with a binary features model of representing discourse. For example, feature models cannot be used to make statements about what determiners or noun phrases are expected where. Binding Theory describes the distribution of pronouns in relation to full NPs, but it is not related to the semantic features used to express discourse relations. It is commonly held that nominal referents initially appear as indefinite NPs and later on in the same piece of writing appear in the form of definite NPs or pronouns. This is a discourse issue, but it has not been well accounted for using the features model. The Hawkins/Robertson model stipulates prior mention as necessary for the licensing of *the*, but the other feature-based studies focus on the sentence and not on the discourse level.

Feature-Based Discourse Theories: Summary

In looking at the studies above, it was concluded that these theories rely on inadequate models of discourse. In these models, the universal nature of certain discourse meanings is often overlooked, and many determiners and forms used to signal

¹³ This is not to say that all languages express discourse relations with the same amount or type of morphology.

information about nominal reference in a way similar to articles are disregarded. The next section will examine more precise models of discourse that take into account issues like topic, focus, and center of attention.

Alternative Models of Discourse

Discourse meanings are built from a variety of linguistic features, and most current models of discourse only account for part of the story. These theories differ in:

- how much information is considered when predicting the flow of discourse relations and nominal reference,
- the facets of discourse focused on (for instance transitions, nominal reference and determiner semantics, or structure of text and its effect on nominal reference),
- what parts of sentences are seen as important,
- how topic and focus are defined, and
- what is seen as being universal or language-specific.

The first part of this section defines different models of discourse that are more detailed and can be used to examine more morphemes than those already discussed. The second part of this section looks at L2 studies that use these kinds of models.

Discourse Models

The models of discourse described below try to explain what linguistic elements lead to a sense of ‘cohesion’ over several sentences. Poor choices of lexical items, word orders, and structures can make a discourse seem awkward, unacceptable, and confusing. For example, Dressler et al. (2004) describe the poorly cohesive writing of right-brain damaged patients as “verbose, non-informative, irrelevant, and repetitive (Dressler, Stark, H., Vassilou, Rauchensteiner, Tomic, Weitzenauer, Wasner, Pons, Stark, J., & Brunner 2004:210).” In cohesive discourse, ease of processing should make discourse features and structure unremarkable or unnoticeable. In trying to determine what linguistic features and structures add to cohesion, various linguistic features are considered—

features such as grammatical function, distance between referents, or communicative purpose. Some models consider more factors. Some examine an entire piece of writing or whole conversation, while others focus only on paragraphs, sentences, or clauses. Models including the study of nominal reference reflexes consider the overall cohesion, the movement of reference and transitions from sentence to sentence, or the changes in forms realizing the same referent across many sentences.

There are four basic classes of models that will be exemplified below. *Anaphoric hierarchies* rank noun phrase forms based on how much information they give to specify a referent. Some lexical forms indicate more information about a referent so that a representation can be built in memory when it does not exist. Others do not give as much indication of how to define a referent, since it may already be the focal point or in memory. *Transition theories* model how the spotlight of the discourse and nominal reference change from sentence to sentence or across entire chains of reference. These theories concentrate on fewer distinctions between nominal forms than anaphoric hierarchies, highlighting the difference between given and new referents, or referents in and out of the center of attention. Both anaphoric hierarchy models and transition models try to simulate how discourse is perceived by an addressee.

Semantics-focused models also concentrate on the distinction between given and new referents. These models build formal representations of what exists in the discourse, and may also focus on evaluating truth values or semantic entailments. *Text-type models*, the last type to be considered, do not try to model how addressees perceive discourse relations. Instead, these theories focus on how the different communicative goals of various texts affect the choice of lexical forms for such discourse-related items as nominal reference. For example, the different purpose of narrative versus description leads to differences in what kinds of referents are selected and how a text progresses over time. The examples given below are meant to illustrate these different perspectives on simulating how discourse meanings are built, and are not meant to be an exhaustive list.

Anaphoric Hierarchies

Anaphoric hierarchies are scales in which NP and determiner forms are ranked according to what kind or how much information is passed on to the audience to help them recognize the referent. For example, pronouns provide less help in distinguishing a referent than full noun phrases. These models hope to answer: How do the specific lexical items (nouns, pronouns, and determiners) chosen to represent nominal referents add to or detract from a sense of cohesion? The L2 writing below has a cohesion problem due to the choice of articles or noun phrases.

Example 4: Lack of Cohesion in L2 English by a Native Japanese Learner¹⁴

Once upon a time, there was *an old man_i* and *an old woman_j* in village. One day *an old man_{i?k?}* went to the mountain to work. And *an old woman_{j?l?}* went to the river to wash the cloths. (Sophomore Female #003)

The linguistic forms do not indicate that the old man in the first sentence and the old man in the second sentence are the same person, although Grice's maxim about relevance leads to the assumption that they might co-refer. Anaphoric hierarchies would claim that if the two noun phrases co-refer, the problem for cohesion is that a lexical form from a different level of the hierarchy was not chosen for the second referent.

Two specific anaphoric hierarchies are:

- Ariel's (1990) Accessibility Theory, and

¹⁴ The Japanese student texts are all from the 'Momotaro' section of the Japanese Learners of English Corpus. These Japanese high school students were asked to retell the fairy tale of Momotaro, a boy born from a peach.

- Gundel, Hedberg, and Zacharski's (1993, 1998, 2001, 2003) Givenness Hierarchy.

Although these are both anaphoric hierarchies, they have certain key differences.

Accessibility Theory—Ariel (1990)

Ariel's (1990) Accessibility Theory is a universal hierarchy of noun phrase forms.

Table 6: Ariel's Accessibility Hierarchy (as cited in Sanders & Gernsbacher 2004:81)

<i>Order and relationship between referent types</i>	
↑ <i>lowest accessibility</i>	(1) full name >
	(2) long definite description >
	(3) short definite description >
	(4) last name >
	(5) first name >
	(6) distal demonstrative >
<i>highest accessibility</i> ↓	(7) proximate demonstrative >
	(8) NP >
	(9) stressed pronoun >
	(10) unstressed pronoun >
	(11) cliticized pronoun >
	(12) zero

Different lexical forms of noun phrases are hypothesized to indicate different levels of accessibility in long term or short term memory. Referents that can be quickly accessed because they are activated in short term memory have 'high accessibility.' For example, referents realized as pronouns are highly accessible. Nominal referents become highly accessible when they are directly related to the topic of the discourse, appear in a parallel

grammatical role in the immediately preceding utterance, or having an antecedent nearby.

The lexical forms in Ariel's hierarchy ascend from markers of the lowest accessibility (full names) to those that mark high accessibility (zero anaphora). Note that this hierarchy focuses only on definite NPs.

Ariel claims that the order of the hierarchy is universal, although not all languages may have all the forms, or at least not all can be freely used in all contexts.

The example below shows some L2 text with the forms referring to Momotaro classified according to Ariel's Accessibility Hierarchy.

Example 5: Native Japanese L2 English Text under the Accessibility Hierarchy

- a. *The baby* (3) was named *Momotaro* (5) after peach and *he* (10) grew up soon. (Sophomore Female #011)
- b. They named *the boy* (3) "*Momotaro*" (5). *Momotaro* (5) grow *the nice boy* (3). One day, *Momotaro* (5) knew about "Oni". Oni lives in Onigashima. They were bad monsters. So *Momotaro* (5) went to Onigashima in order to fight them. (Sophomore Female #016)

In (a) above, the lexical instantiations of the referent ascend the hierarchy from less accessible to more accessible. When *he* is used, the referent Momotaro is highly accessible because there is an antecedent nearby, and because Momotaro is the topic of the discourse. The example in (b), though, does not have repeated references to Momotaro realized by lexical items higher on the hierarchy, and unless the third and the last sentence are read as starting new paragraphs or sections of discourse, this excerpt sounds strange. The student who wrote (b) is being overly explicit—using a form to signal a less accessible referent when the referent is highly accessible due to its being

related to the topic, recently mentioned, and in a parallel grammatical role to previous sentences.

Although Ariel's hierarchy could be used to make predictions about L2 use of nominal references, these predictions are limited. Ariel's scale does not include indefinites, which leaves out a whole category of nominal reference that is a source of L2 errors. In addition, the predictive capability is limited to stating that nearby subsequent references (for example, in the same paragraph) should be higher on the scale (if they meet some of the three criteria for being more accessible). There is no indication of how much higher on the scale subsequent references should be, though. The scale could be hypothesized part of the knowledge of UG, though learners would have to acquire which forms can be regularly realized in their L1 and L2 and where, and the scale does not give any indication of what combinations of forms are possible in a language. If the full hierarchy is universal, then learners making errors in determiner use would have to be explained as having a mapping problem similar to that described by the Missing Surface Inflection Hypothesis. Alternatively, their production could be affected by phonology, as hypothesized by the Prosodic Transfer Hypothesis. The next version of an anaphoric hierarchy to be looked at is set up very differently from Ariel's, and does include indefinite forms (although generic NPs are excluded).

The Givenness Hierarchy—Gundel, Hedberg, & Zacharski

(1993, 1998, 2001, 2003)

Gundel, Hedberg, and Zacharski's (1993, 1998, 2001, 2003) anaphoric hierarchy for English is shown in Table 7 on the following page.

Table 7: Gundel, Hedberg, & Zacharski's Givenness Hierarchy for English

<i>Attention/ memory state</i>	<i>Archetypal lexical item(s) for referent</i>	<i>Definition/ example</i>
in focus (high accessibility) (center of attention)	it	The referent is not only in short-term memory, but is also the current center of attention. This status is necessary for the appropriate use of zero and unstressed pronominals. The entities in focus at a given point in the discourse will be that partially-ordered subset of activated entities which are likely to be continued as topics of subsequent utterances. <i>My neighbor's dog bit a girl on a bike. <u>It's</u> the same dog that bit Mary last summer.</i> (GHZ 1993:279-280)
activated (in working memory)	this, that, this N	The referent is represented in current short-term memory. Activated representations may have been retrieved from long-term memory or they may arise from the immediate linguistic or extralinguistic context. <i>I couldn't sleep last night. (barking) <u>That</u> kept me awake.</i> (GHZ 1993:278) <i>The authors developed a more detailed version of the original and a draft was prepared in 1986. <u>This draft</u> has been in wide circulation.</i> (Gundel 1998:188)
familiar (in memory)	that N	The addressee is able to uniquely identify the intended referent because s/he already has a representation of it in memory (in long term memory if it has not been recently mentioned or perceived, or in short-term memory if it has). <i>I couldn't sleep last night. <u>That dog</u> (next door) kept me awake.</i> (GHZ 1993:278)
uniquely identifiable	the N	The addressee can identify the speaker's intended referent on the basis of the nominal alone. Identifiability may be based on an already existing representation in memory, but it does not have to be based on previous familiarity if enough descriptive content is encoded in the nominal itself. <i>I couldn't sleep last night. <u>The dog</u> (next door) kept me awake.</i> (GHZ 1993:277) <i><u>The first paper</u> on centering was published in 1983 (Grosz, Joshi and Weinstein 1983). Subsequently, <u>the authors</u> developed a more detailed version of <u>the original</u>.</i> (Gundel 1998:187)

Table 7—Continued

<i>Attention/ memory state</i>	<i>Archetypal lexical item(s) for referent</i>	<i>Definition/ example</i>
referential	(indefinite) this N	The speaker intends to refer to a particular object or objects. The addressee not only needs to access an appropriate type-representation, s/he must either retrieve an existing representation of the referent or construct a new one by the time the sentence has been processed. <i>I couldn't sleep last night. <u>This dog</u> (next door) kept me awake. (GHZ 1993:276-7)</i> <i>And there was <u>this temporary</u> when George went over to Econ. And he was in my office when she came over to borrow something. (Gundel 1998:186)</i>
type identifiable (low accessibility)	a/an N	The addressee is able to access a representation of the type of object described by the expression. They understand the semantics of the kind of noun. <i>I couldn't sleep last night. <u>A dog</u> (next door) kept me awake. (GHZ 1993:276)</i> <i>Susan gave Betsy <u>a pet hamster</u>. (Gundel 1998:194)</i>

Note: The items in italics are meant to represent a type of lexical item. For example, *it* would represent all unstressed personal pronouns. In addition, there is an entailment relationship in that any noun phrase entails all those below it in the table. For example, any NP that is uniquely identifiable, must also be referential and type identifiable. An NP that is type identifiable if the audience understands the semantics of the common noun, able to pick out what kind of thing the referent is.

This hierarchy is different from Ariel's because lexical forms are matched with mental states of attention or indications of how the referent is stored in memory. Just as with Ariel's scale, though, different lexical forms of noun phrases are hypothesized to indicate different levels of accessibility in long term or short term memory. Going down the table, the lexical items decrease in accessibility. One other difference between Gundel, Hedberg, and Zacharski's (GHZ) scale and that of Ariel is that GHZ's scale implies an entailment relation, not just a simple ranking. This means that if a noun

phrase is in focus, it entails that it is also activated, familiar, uniquely identifiable, referential, and type-identifiable.¹⁵

Gundel, Hedberg, & Zacharski do not fully specify what algorithm addressees use for determining mental attention states. This is because:

While linguistic form plays an important role in determining what will be brought into focus, actual inclusion in the ‘in-focus’ set depends ultimately on pragmatic factors, and is not uniquely determinable from syntax. (Gundel, Hedberg, & Zacharski 1993:280)

GHZ’s rough coding guidelines look at syntax and how recently mentioned a form is, but some coding decisions must be based on perceived background or common knowledge. In their studies, most coding disagreements arose when deciding whether a referent was familiar vs. activated or activated vs. in focus (GHZ 1993:291, GHZ 2001:282). GHZ state, “We believe this is because the boundaries between statuses involving attention are not discrete, even though they map onto discrete forms (GHZ 1993:291).” Despite not having rigid guidelines, trained coders in these studies still achieved a fairly high level of accuracy, with agreement in coding between 77-90% overall.¹⁶

Gundel, Hedberg, and Zacharski also explain why often more than one type of determiner (or a pronoun instead of an NP) is appropriate in a given context by stating that their scale interacts with other pragmatic considerations, such as Grice’s Maxims (Grice 1975). Whenever a lexical form is used that is not typically related with the mental state of the referent, Grice’s Maxims will lead the hearer to resolve the difference by

¹⁵ Note that *focus* here refers to the center of attention, or what amounts to the topic of the discourse according to other linguistic researchers. Given information is *in focus*, and focus does not refer to new information for Gundel, Hedberg, and Zacharski.

¹⁶ In the 1993 study, the two raters agreed in 90% of the cases (GHZ 1993:291). In their 2001 study, four raters varied between 77-88% agreement. The agreement rate varied based on what kind of text was being coded. In addition, in the 2001 study GHZ were just classifying referents as familiar (focus, activated, familiar) or non-familiar (uniquely identifiable, referential, type-identifiable) (GHZ 2001:282).

attaching some special meaning to the situation or lexical item. For example, their scale does not have the meaning of “make a referent salient” associated with the demonstrative determiner *that*. However, this meaning is possible. See the next example.

Example 6: *That* Making a Referent More Salient (GHZ 2001:277)

[It is dusk and John and Mary are returning from a shopping trip. As John is parking the car, Mary exclaims:] Good God! Look at *that incredibly bright light*.

Under the Givenness Hierarchy, the light is at most uniquely identifiable, since there is no representation of it in memory—this is the first time it is being encountered. Normally, uniquely identifiable referents are paired with *the* in English. Hearers know that the referent is uniquely identifiable, and the fact that there is a non-typical lexical item paired with it leads them to add the salience meaning to the demonstrative determiner. This is because Grice’s Maxim of Quantity tells the hearers that the speaker is not being more or less informative than necessary. In this case, using the lexical item from the status ‘familiar’ (moving up the scale) achieves the special effect of making the hearer search for a familiar referent nearby, and also changes the focal point of the conversation by activating a new prominent referent (GHZ 2001:277).

The Givenness Hierarchy in the last table is specific to English, but the link to states of attention or memory makes it easy to compare the hierarchies of different languages. The table that follows shows a comparison of five different language’s lexical items used to express discourse relations. The highest variation between languages lies in what forms (if any overt forms) are used to differentiate between type identifiable, referential, and uniquely identifiable referents.

Using this model and a table such as the GHZ form-comparison table, it is easy to make predictions about what mistakes L2 learners would make if they were transferring discourse forms and meanings from their L1. A bare noun in Chinese, Japanese, and Russian is ambiguous—it could be either definite and uniquely identifiable, or indefinite and either referential or merely type identifiable. This predicts that if learners do have L1 transfer, they would be likely to omit articles, or since Chinese has an indefinite article, they may use the corresponding ‘one.’

Similar results have been found and explained in previous studies (such as Robertson 2000), but with this framework it would be more clear if learners are making mistakes with all kinds of noun phrases and referents, or only those that are unactivated or not clearly distinguished in memory. This framework also recognizes the importance of center of attention, subject position, and other matters important for determining the form of an NP. These other influences cannot be accounted for with, for example, the binary features approach. The GHZ framework could also be used to evaluate whether learners are overly explicit in their use of noun phrases—meaning they are not using the forms for activated referents as often as native speakers.

The next section surveys transition models. Transition models and anaphoric hierarchies both focus on nominal referents, but transition models focus on how center of attention changes, not the activation status of a particular referent.

Transition Models

Transition models of discourse look at the forms of noun phrases to see how the center of attention changes over a piece of discourse. In terms of cohesion, these models ask: Are shifts in the focal point of the discourse from one nominal referent to another adding to or subtracting from the cohesion of the discourse? Are changes in spotlight of the discourse smoothly made, or are they causing confusion about the focal point?

Table 8: Gundel, Hedberg, & Zacharski's Cross-Linguistic Comparison of Discourse
Forms (GHZ 1993:284)

<i>Language</i>	<i>In focus</i>	<i>Activated</i>	<i>Familiar</i>	<i>Uniquely identifiable</i>	<i>Referential</i>	<i>Type identifiable</i>
Chinese 17	Ø <i>ta</i> 's/he, it'	TA stressed 'he' <i>zhè</i> 'this' <i>nèi</i> 'that' <i>zhè</i> N		<i>nèi</i> N		<i>yi</i> N 'a N' Ø N
English	<i>it</i>	HE, <i>this, that,</i> <i>this</i> N	<i>that</i> N	<i>the</i> N	indefinite <i>this</i> N	<i>a</i> N
Japanese	Ø	<i>kare</i> 'he' <i>kore</i> 'this' <i>sore</i> 'that' medial <i>are</i> 'that' distal <i>kono</i> N 'this N' <i>sono</i> N 'that N' medial	<i>ano</i> N 'that N' distal	Ø N		
Russian	Ø <i>on</i> 'he'	ON stressed 'he' <i>èto</i> 'this' <i>to</i> 'that'	<i>èto</i> N <i>to</i> N	Ø N		
Spanish	Ø <i>él</i> 'he'	ÈL stressed 'he' <i>éste</i> 'this' <i>ése</i> 'that' medial <i>aquél</i> 'that' distal <i>este</i> N	<i>ese</i> N 'that N' medial <i>aquel</i> N 'that N' distal	<i>el</i> N 'the N'	Ø N <i>un</i> N 'a N'	

¹⁷ Note that Gundel, Hedberg, & Zacharski, in regards to Chinese, Japanese, and Russian, state that "a noun with no preceding determiner in these languages can be interpreted as either uniquely identifiable (definite) or merely referential or type identifiable (indefinite)." (GHZ 1993:284)

For example, the story in the next example does not have a smooth shift in the center of attention, and this accounts for why it is not fully cohesive. The last sentence is confusing because the referents *Mike* and *Tony* have switched grammatical positions, but also because which referent is referred to with a pronoun has changed.

Example 7: A Story that is Not Fully Cohesive

Mike called Tony at 6 AM. Tony was sick and furious about being woken up so early. He told Mike to get lost and hung up. *Of course, he hadn't intended to upset Tony.* (Grosz, Joshi, & Weinstein 1995:207)

The story above can still be understood, but it takes longer to process because conventions used to signal a change in which referent appears in subject position and is referred to by a pronoun have not been followed.

Researchers who are concerned with the introduction, shift, and maintenance of reference define what constitutes these categories in different ways. Agreement has not been reached about whether there can be more than one topic, center of attention, or focus at one time, and this complicates the process of defining maintenance and shift.¹⁸ The importance of the subject position and the distance between referring expressions is also disputed. Furthermore, some researchers trace transitions from sentence to sentence, while others examine transitions in chains of reference (meaning they look at all the representations of a referent and determine how these representations have shifted).

¹⁸ Note that this is not the same use of *focus* used by many other linguistic researchers. In most of these alternative models of discourse, *focus* is used to refer to given information that is the center of attention. This use should not be confused with the meanings of *topic* and *focus* used by many syntacticians.

Centering Theory (Grosz, Joshi, & Weinstein 1983,1986, 1995; Walker, Joshi, & Prince 1998), discussed below, is one of the most commonly discussed theories of transitions of centers of attention. There are other, much simpler ways of looking at transitions than Centering Theory. In fact, simple models are typically what is used in second language studies, as will be seen in a later section.

Centering Theory

Centering Theory (Grosz, Joshi, & Weinstein 1983, 1986, 1995; Walker, Joshi, & Prince 1998) is an algorithm for determining which nominal referent is the center of attention based on the grammatical functions of the noun phrases. The status of being in or out of the main spotlight of the discourse is important because it has been found that, universally, language items that are focal points or given can be represented by shorter or less full forms than those that are not in focus or new (Hendriks 2003:292). Centering Theory states that each language has a hierarchy of grammatical functions, and referents in the highest grammatical function are more likely to be the center of attention in the immediately following utterance. For example, the hierarchy for English would be: Subject > Object(s) > Other (Walker, Joshi, & Prince 1998:7—see also in this article their hierarchy for Japanese).

Although the hierarchies of grammatical functions are language specific, the links that make discourse connections between sentences are universal. Centering Theory proposes that every sentence has a list of forward-looking centers (C_f s), one backward-looking center (C_b), plus one preferred center (C_p). The list of forward-looking centers are ranked (at least partially) by the language-specific hierarchies of grammatical functions. These make up the list of possible referents that could be the focal point in the next utterance. The preferred center is the highest ranked referent on the list of forward-looking centers, and amounts to a prediction of the topic of the next sentence. The backward-looking center is loosely defined as the topic of the current sentence. The C_b is

‘backward-looking’ because this referent is the one that connects the current utterance to the previous one. If a discourse topic is new, then the first utterance under this topic has no C_b , or an undefined C_b . By definition, Centering Theory specifies that there can be only one backward-looking center per utterance.

Centering Theory defines four types of transitions: continue, retain, smooth-shift, and rough-shift. These are defined by comparing the backward-looking centers of the current and previous sentence, as well as the backward-looking center and the preferred center of the current sentence. For example, when the C_b is the same as in the last sentence, and it matches the preferred center, this is a continue type of transition. The rest of the transitions are explained in the table below.

Table 9: Centering Theory Transitions (Walker, Joshi, & Prince 1998:6)

	<i>Current $C_b = Past C_b$ or Past C_b undefined</i>	<i>Current $C_b \neq Past C_b$</i>
<i>Current $C_b = Current C_p$</i>	CONTINUE	SMOOTH-SHIFT
<i>Current $C_b \neq Current C_p$</i>	RETAIN	ROUGH-SHIFT

The example below shows how a piece of text would be coded for centers and transitions using this model. A piece of discourse with rough-shifts would be less cohesive and take longer to comprehend because there would be no connection between the backward-looking centers and the preferred center.

Example 8: Story Coded Using Centering Theory

John has been having a lot of trouble arranging his vacation. He cannot find anyone to take over his responsibilities. He called up Mike yesterday to work out a plan. Mike has annoyed him a lot recently. He called John at 5 a.m. on Friday last week. (Grosz, Joshi, & Weinstein 1995:217)

John has been having a lot of trouble arranging his vacation.
C_b = ; C_f = {John} (his = John)

He cannot find anyone to take over his responsibilities.
C_b = John; C_f = {John} (he = John) (CONTINUE)

He called up Mike yesterday to work out a plan.
C_b = John; C_f = {John, Mike} (CONTINUE)

Mike has annoyed him a lot recently.
C_b = John; C_f = {Mike, John} (RETAIN)

He called John at 5 a.m. on Friday last week.
C_b = Mike; C_f = {Mike, John} (SMOOTH-SHIFT)

One problem with Centering Theory is its claims about the distribution of pronouns. Centering Theory states:

If there are multiple pronouns in an utterance, then one of these pronouns must realize the C_b. In addition, if there is only one pronoun, then that pronoun must be the C_b. (Walker, Joshi, & Prince 1988:5)

By linking pronominalization to the C_b, Centering Theory tries to explain the fact that referents expressed as pronouns must be in focus. However, this restrictive rule is not always accurate. See the excerpt of conversation below from Gundel, Hedberg, & Zacharski (1998:192-3). Centering Theory cannot explain how *it* in the last phrase could be allowed to be realized as a pronoun. The referent of *it* should not be the backward-looking center since the referent 'golf ball' does not appear in the immediately preceding clause, (9b) (Gundel, Hedberg, & Zacharski 1998:194).

Example 9: Dialogue Showing Centering Theory's Explanation of Pronoun Use is Too Limited

And the guy wrote little marks on his golf club
 as to where to put his hands
 and he had marks as to where to put his feet
 and he did it all, uh, very scientifically
 and he got his golf score way down, you know
 and George played with him like, uh, once or twice
 and each time George'll just whack *it*.

Brennan (1998) has also criticized Centering Theory, stating that the Centering algorithm does not model how people center attention. This is because Centering is based on strict transitions, meaning the model describes one focal point being replaced by another and essentially forgotten. This does not factor in the effects of memory and activation on focal point retention. The sentences in the next example come from Brennan's work in which native speakers described the action of a basketball game to others who could not see what was occurring. Sentence (b) was never produced and sounds odd.

Example 10: Effects of Time and Distance on Sentence Topic (Brennan 1998:245)

- (a) Smith to Jones. Jones shoots.
- (b) ? Smith to Jones_i. He_i shoots.
- (c) Smith hands the ball over to Jones. Jones shoots.
- (d) Smith hands the ball over to Jones_i. He_i shoots.

Centering Theory accurately predicts that (b) should be somewhat odd because the transition from the first to the second sentence would be classified as a rough shift.

Sentence (d) should be a rough shift as well, but in fact does not sound odd because of

the distance between *Smith* and the pronoun in the second sentence. *Smith* should be the preferred center for the second sentence, but over the time taken to process the longer sentence, *Smith* is less salient in memory, and *Jones* becomes a better antecedent for the pronoun. Centering Theory cannot account for the difference between (b) and (d).

Transition models that are not as complicated as Centering Theory do not have the same problems with predictions of pronominal reference or the effects of memory and activation on reference (which is a logical outcome of ignoring certain complexities). These simpler models do not have algorithms for determining types of centers. Instead, transitions like continuation or maintenance of topic, for example, are defined by whether co-referring NPs appear in the immediately preceding utterance. A simpler transition model will be described in the next section.

A Simple Transition Model—Von Stutterheim, Mangold-

Allwinn, Barattelli, Kohlmann, and Kolbing (1993)

Simpler transition models have been frequently used in L2 studies examining discourse proficiency, and have been shown to provide useful generalizations and predictions. These models are simpler than the Centering Theory algorithm, and also simpler than the Givenness Hierarchy. The criteria for classifying given referents in simple transition models is firmly set by a definition, so there is less possibility of coding errors. Since it is unknown what level of abstraction learners actually use when selecting lexical forms to express discourse relations, it is unknown whether a simple model can be used to determine causes of L2 determiner errors.

The simple transition model that will be examined was developed by Von Stutterheim et al. (1993). This model has three categories to classify referents—referents could be new, maintained, or rementioned. New referents are those mentioned for the first time in the text. Maintained referents are any NPs that appear in both the current and previous clause. Rementioned referents are subsequent mentions of referents that appear

in the current clause, but not the previous one because other referents have intervened in intermediate sentences. See the example that follows. The second time the referent *an old woman* is used it is classified as a remention because the preceding clause did not contain any mention to the woman. In Centering Theory, *an old woman* would not be considered appropriate as a focal point in this sentence because it did not appear in the last sentence. Note that the simple transition model does not imply that this should be a problem. Transitions are not classified as rough or smooth shifts in center of attention.

As far as what this theory says about the appropriateness of any given NP, Von Steutterheim et al. (1993) did this by looking at native production in different contexts. For example, if a new referent is a noun, a maintained referent may be a pronoun, and a rementioned referent may be a repetition of the name. (See for instance Table 10 below.)

Example 11: L2 Learner Writing Coded with the Simple Transition Model (Female
#003)

Once upon a time, there was *an old man* and *an old woman* in village.
new *new*

One day *an old man* went to the mountain to work.
maintained

And *an old woman* went to the river to wash the cloths.
rementioned

When *she* is washing there the peach is streaming from the river.
maintained

The next type of model to be described is a semantics-focused model. Semantics-focused models concentrate on the distinction between given and new reference as well, but use formal logic, and examine truth values and entailments more in depth.

Semantics-focused models are more abstract in their ways of looking at word, sentence, and discourse meanings than the other discourse theories discussed.

Semantics-Focused Models

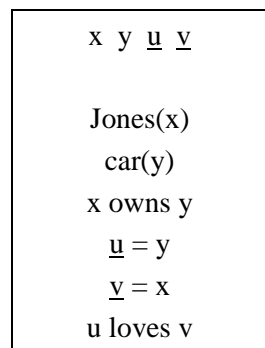
Discourse Representation Theory (DRT) (Kamp & Reyle 1993) is the main model of this type, and it is primarily different from the other theories considered in its formalism and use of predicate logic. DRT considers nominal reference, verb tenses, and aspect, using logic to evaluate and describe truth conditions and entailments.

Discourse Representation Theory—Kamp and Reyle (1993)

Discourse Representation Theory (Kamp & Reyle 1993) uses predicate logic to describe nominal reference and chains of reference. Discourse Representation Structures (DRSs) are built from the elements of sentences that need to be evaluated for truth conditions. Defined construction rules are used to transform syntactic structures into predicate logic. For each noun phrase, a new nominal referent is introduced into the DRS. Pronouns can either be referential or bound variables. Therefore, if an NP consists only of a pronoun, an equality statement is added to the DRS to show what the antecedent or chain of reference is. The example of a DRS on the following page contains both a proper noun and an indefinite noun phrase. The letters at the top of the DRS are meant to represent the number of referents in the model. The predicate logic below shows the relations of the predicates in the sentences. Although pronouns and full nouns are represented differently, notice that indefinite NPs and proper nouns are both represented the same way.

Kamp and Reyle propose a slightly different predicate logic for definite NPs.

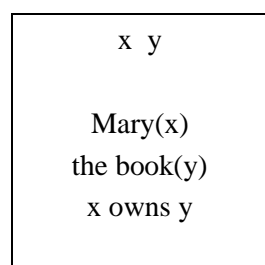
Figure 4: A DRS for the Sentences: Jones owns a car. He loves it. (see Kamp & Reyle 1993:77)



The DRS below has a definite NP in it. In it, the noun phrase preceded by *the* is simply represented with the article in the DRS. In Kamp and Reyle (1993), it is acknowledged that this way of representing definite NPs is inadequate, but the question of refining this is left to later works (Kamp & Reyle 1993:254-255).

Such later works have indeed described the semantics of definite and indefinite articles with more detailed predicate logic, but it remains unclear how useful DRT would be in examining the kinds of errors that L2 learners make in article or determiner use.

Figure 5: A DRS for the Sentence: Mary owns the book. (See Kamp & Reyle 1993:254.)



Kamp and Reyle (1993) focuses on articles, and does not discuss how definite noun phrases with demonstrative determiners can be differentiated in a DRS. Later, more detailed, descriptions of DRT have the same problem in that few forms of determiners can be distinguished in the formal meaning representation. In DRSs, certain types of noun phrases are represented by the same predicate logic. The fact that the meaning interpretation is more relevant to the DRS than the form makes it difficult to state and evaluate what might be first language transfer from a language that did not have articles. In a study of L2 determiner use, though, the form is not irrelevant.

Guerts and Beaver (2007) in particular discuss how the form of the actual language production is not always relevant in DRT. For example, they state that “there is essentially no difference between the DRT analyses of cross-sentential anaphora...and sentence-internal anaphora... In either case, the pronoun simply serves to pick up an accessible discourse referent.” (Guerts and Beaver 2007:Section 3.2) It is not just sentence boundaries that do not matter in the DRS, though. Guertz and Beaver state that at its heart DRT is not compositional:

Not only is DRT a representational theory of interpretation, it is a non-compositional theory as well. These two features are intertwined. Consider, for instance, the way pronouns are interpreted in basic DRT, by first setting up a referent marker, which is subsequently linked to another discourse referent. This is a non-deterministic process, but even if it were not, it is clear that the anaphoric link is not part of the meaning of the pronoun. In standard DRT, the pronoun does not, in and of itself, introduce something into the DRS that has a model-theoretic interpretation. A standard statement of compositionality would say that the meaning of compounds must be a function of the meaning of their components and their mode of combination. But if some of the components, like pronouns, do not introduce into the DRS any object that can naturally be described as the meaning of that object, then it is clear that we do not have a compositional system. (Guerts and Beaver 2007:Section 6)

DRT is a very abstract and highly formalized model for representing meaning. Although it is possible to use it to examine perceived L2 determiner errors, it is not clear that this level of philosophical detail is necessary to understand the problem of article

acquisition. The next model to be examined in fact incorporates DRT into a model of discourse with a different focus. The next type of model to be discussed concentrates on text structure. Text-type models are also compatible with other models of discourse. Smith (2003) uses her text model in conjunction with DRT, and also claims that her model could be used with anaphoric hierarchy models for a more complete picture of how discourse functions.

Text-type Models

Text models of discourse look at the effect of text type on how noun phrases and other elements are realized. They hypothesize that communicative goals affect the lexical selection of discourse elements. From a reader's perspective, this means that the purpose of a text can be deduced in part from the linguistic cues present. Of the two text-type models below, Smith (2003) is the more comprehensive, covering texts of more types and considering more linguistic features. The model of von Stutterheim, Mangold-Alwinn, Barattelli, Kohlmann, and Kolbing (1993), though simpler, has a slightly different perspective and is based on facts of German instead of English.

Modes of Discourse—Smith (2003)

Different Modes of Discourse are different text types. Smith (2003) defines five modes, which are differentiated by how time proceeds, how ideas progress, and what types of predicates are typical. The types of modes are divided into temporal modes, in which time is important for the progression of the text, and atemporal modes, in which time is not important.

The three temporal modes are Narrative, Report, and Description. In Narratives, time moves along a sequence of events, lead by either verbs or time adverbs. In addition, most of the predicates consist of either events or states. Reports also mainly have predicates that are events or states, but they can also contain General Statives. Smith defines General Statives as follows: "General Statives are expressed by generic and

generalizing sentences. They invoke patterns of Events and States rather than particular situations (Smith 2003:12).” Some examples of General Statives are below.

Example 12: Samples of General Statives (Smith 2003:24)

- (a) Dinosaurs are now extinct.
- (b) The lion has a bushy tail.
- (c) Mary speaks French.
- (d) John always fed the cats last year.

Time in Reports is related to the speech time, and the events do not progress in a sequence. The text progresses from one time viewpoint to another in what Smith states is a back and forth manner. Descriptions are the last temporal mode of discourse that Smith defines. Time is static in descriptions, and the text progresses around a view of a scene. The majority of predicates are events, particularly ongoing events, and states.

The two atemporal modes of discourse are Information and Argument. Information texts contain primarily General Statives. The movement in the text proceeds metaphorically, through an important noun phrase Smith describes as the Primary Referent. The Primary Referent is not the topic, but often is the NP with the patient theta-role, and it is important to the semantics of the sentence (See Smith 2003:17). The text excerpt below shows movement through Primary Referents, which are underlined. The Primary Referents in this passage are all important to the overall understanding of the excerpt, and are all related in some way. This is what causes a feeling of movement of ideas.

Example 13: Information Mode: Text Progression through Primary Referents (Smith 2003:17)

When people try to get a message from one individual to another in the party game “telephone,” they usually garble the words beyond recognition. It might seem surprising, then, that mere molecules inside our cells constantly enact their own version of telephone without distorting the relayed information in the least.

The criteria for selecting Primary Referents are given in the next example.

Example 14: Criteria for Determining the Primary Referent (Smith 2003:125, 244)

Events

- (a) Undergoes a change of state
The high school outsider becomes *the more successful adult*.
- (b) Causally affected by another participant
The national outpouring has forced *us* to confront the situation.
- (c) Does not exist independently of the event
High school students present and past have come forward with *stories about cliques and an artificial world*.
- (d) Moves or otherwise changes
Young people mature substantially earlier in the late 20th century than they did when high school was invented.

States

- (e) Literally or metaphorically located
Dragons are usually arranged almost heraldically round a conceptual center point.
- (f) Dependent on the situation for existence
The predominant output was *the white ware* with transparent ivory toned glaze which made the kilns famous.
- (g) Figure relative to a Ground
A group of kilns is northeast of Ch’ang-an, the capital city of the T’ang dynasty.
- (h) Has a property ascribed to it
The most important kilns are *those at Tao-chu in Shensi*.

From these criteria, it can be seen that Primary Referents are rarely subjects, but may often be objects. The last atemporal mode of discourse is Argument. Primary Referents are also responsible for metaphorical movement through Argument texts. Arguments consist of two main types of predicates: General Statives and Abstract Entities (facts and propositions). Some examples of Abstract Entities are given below. In each example, *Mary's refusal of the offer* is being reported as a fact, or else an opinion about it is being given. The communicative intent is not just to report an event, but rather comment on a metacognitive or abstract perception of an event.

Example 15: Samples of Abstract Entities (Smith 2003:25)

- (a) I know that Mary refused the offer.
- (b) Mary's refusal of the offer was significant.
- (c) I believe that Mary refused the offer.
- (d) Mary's refusing the offer was unlikely.

Aside from the mode, Smith also highlights the importance of topic and focus to understanding how discourse is expressed in texts. Focus is defined as new information or as the element that would receive focus stress if the sentence were spoken. This is different than how focus is defined in the other discourse models of this section. This definition is the same as that which would be used by many syntacticians. In English, the unmarked case is the one in which the last semantically important word receives the focus stress. In other languages, there are focus phrases and the focus may be preposed. It is important to note that Smith uses the term focus to mean something different than most of the discourse models surveyed so far. In many of the other theories, "focus" refers to what Smith defines as the topic. Topics are given information. Very often

topics appear as subjects, and in Centering Theory would be described as the backward-looking center. Smith's criteria for determining the topic of a sentence are below.

Example 16: The Criteria for Identifying the Topic Phrase (Smith 2003:198-9, 245)

An NP may be the topic phrase of a sentence if it:

- (a) is the subject of a sentence;
- (b) is a pronoun;
- (c) realizes the agent or experiencer argument of the main verb;
- (d) is coreferential with the topic phrase in the preceding sentence;
- (e) is coreferential with a topic phrase in the context;
- (f) is coreferential with a phrase in the context;
- (g) is lexically related to other material in the context;
- (h) is in a parallel grammatical position with an NP in the context.

This list is very useful and synthesizes many of the elements that are considered by various other models into a comprehensive set of criteria¹⁹.

The next text-type model, that of Von Stutterheim et al. (1993), looks at fewer text types than Smith. On the other hand, it also has a different perspective on why the forms selected for nominal reference differ depending on the communicative purpose of the text.

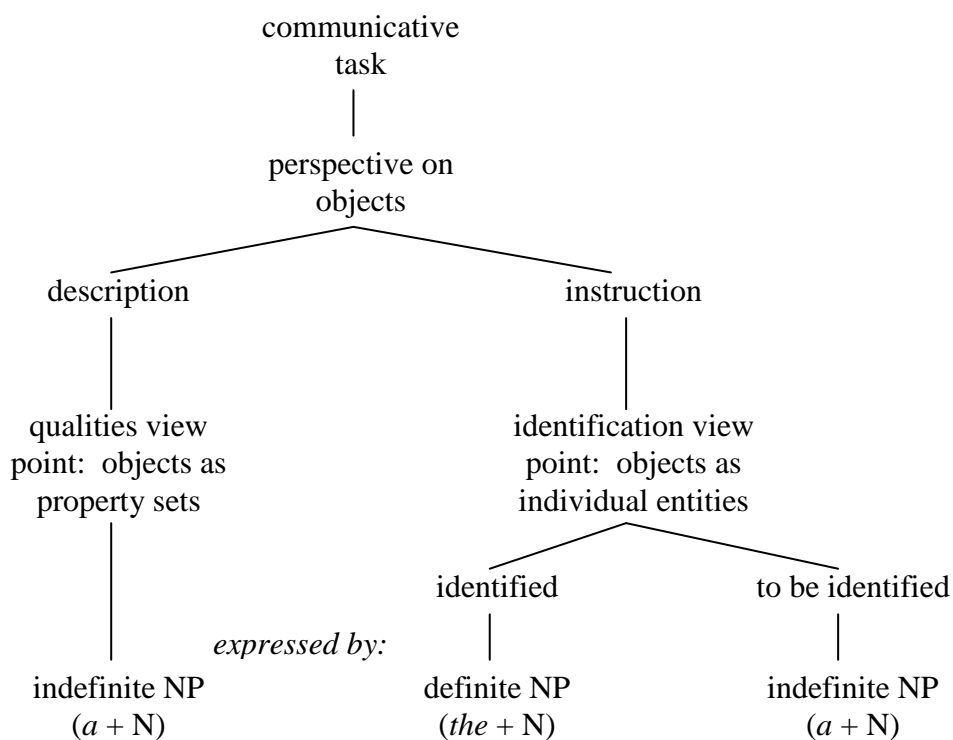
Von Stutterheim et al. (1993)

The model of Von Stutterheim et al. (1993) is relevant for only instruction or description, and furthermore only focuses on articles (not all determiners), but what is stated about perspectives on objects is very interesting. Von Stutterheim et al. assert that

¹⁹ In fact, Smith cites Prince (1992) and Birner and Ward (1998) as both demonstrating findings indicating that text-type models better predict the lexical forms of referents than discourse models that try to model how addressees perceive discourse (see Smith 2003:145). This is relevant to an examination of determiners because it indicates that the purpose and style of a text could significantly affect how determiners and noun phrases are chosen and positioned.

the communicative task provides a viewpoint on objects that appear in a text-type, and this viewpoint in turn affects what kind of lexical forms are available for nominal reference (Von Stutterheim et al. 1993:103). Their diagram that follows explains the actual forms chosen.

Figure 6: Text Type, Viewpoint, and Type of Article (Von Stutterheim et al. 1993:111)



For instruction-giving texts, objects are viewed as individual entities, and the given versus new distinction is important for selecting the lexical form of the nominal referent. In contrast, in descriptions, objects are not viewed as individual entities. They

are seen as representatives of a kind of object, and indefinite NPs are more often used. The example below from Smith (2003) shows how this model is accurate.²⁰

Example 17: Sample of a Description from Smith (2003:29)

We were in *an impressive and beautiful situation on a rocky plateau*. It was too high for *grass*, there was *very little earth* and the place was littered with *boulders*, but the whole plateau was covered with *a thick carpet of mauve primulas*.

In this example, there are only two definite noun phrases and two pronouns, but six indefinite noun phrases. Although descriptions do not solely contain indefinite NPs as Von Stutterheim et al.'s diagram may imply, they may be the most common type of NP in this text type. Just as Smith (2003) defined what kind of predicates are most likely with the different modes, there may be some modes that have an NP form that is more common than others.

Text-type models consider discourse from a different perspective than the other types of discourse models. They look at the overall structure of a piece of writing instead of how utterances are perceived by listeners. The next section will compare the models presented, and then look at different L2 studies using these kinds of alternative models.

Summary and Evaluation of Discourse Models

Each of the different flavors of discourse model—anaphoric hierarchies, transition models, semantic models, and text models—consider similar linguistic elements, but in different ways. All the types recognize that in nominal reference, the status of given

²⁰ Von Stutterheim et al. (1993:120) provide examples of descriptions from their study, but this example from Smith (2003) is meant to show that their conclusions are generally applicable, and do not simply apply to German and the specialized context of Von Stutterheim et al.'s study.

versus new is important. Anaphoric hierarchies specifically break down givenness into smaller categories, either based on lexical items or states of attention. For example, the Givenness Hierarchy correlates lexical forms with six different states of attention. While not all the theories are easily applicable cross-linguistically, the Givenness Hierarchy strives for a universal perspective on discourse meanings by comparing lexical forms of reference from many languages to the same state of mental focus or attention.

Anaphoric hierarchies, transition models, and text models also all share the idea that the topic or center of attention of the discourse has an important effect on the surface form of NPs. The meaning of *topic* differs in these theory-types, but factors that are seen as important by all three are, for example: the salience of the subject position, the effect of how close a co-referring element is, and the grammatical position of the current element relative to a co-referring NP in the adjacent sentences. Smith's (2003) Modes of Discourse text model provides very comprehensive criteria for determining the topic. Simple transition models consider much less when determining the topic.

The next section examines some L2 studies that use these types of discourse models as their theoretical foundation. These studies look at nominal reference, but do not principally examine L2 production for mistakes in articles and determiners.

L2 Studies Using Alternative Models of Discourse

Prior studies using the discourse models presented in the last section to examine L2 acquisition have been successful in finding a discourse pattern of error that results in observed variability. Since determiners express discourse relations, it is possible that L2 errors in their use may be related to a discourse problem, and not merely a general morphological access problem. The purpose of this section is to survey how these alternative discourse models are used in L2 studies. In addition, there will be some discussion of what parts of discourse are viewed as universal versus language-specific.

L2 Use of Null Subjects—Saunders (1999)

Saunders (1999) found that discourse proficiency affects the appropriate use of null subjects in the L2 Spanish of native English speakers. Her analysis employed an anaphoric hierarchy to compare the NPs of native Spanish speakers with those of English speakers learning Spanish. Both sets of speakers were given the task of relating a narrative from a picture story. The hierarchy that Saunders' study used is based on Lakoff (1976), Stoddard (1991), and Strömquist & Day (1993), and is shown below. Items higher on the scale are less accessible in the discourse, and those lower on the scale are more accessible.

Figure 7: Saunder's (1999) Anaphoric Hierarchy

indefinite NPs >> definite NPs >> proper nouns >> overt pronouns >>
null pronouns (Saunders 1999:51)

Saunders considered it universal that new referents be introduced with lexical items higher in the hierarchy, with subsequent mentions usually made with lower forms unless another referent intervenes, in which case the referent is reintroduced relatively higher up the scale (Saunders 1999:52). The language-specific aspect was whether or not null subjects were generally available for use. English restricts use of null pronouns to very few registers, while they are common in Spanish.

The goal of using the hierarchy was to determine whether the native English speaking students truly understood the meaning and use of null subjects. Saunders wanted to gauge whether learners were transferring referent use from their L1, being over-explicit in a way not acceptable in the L1 or L2, or using null pronouns in discourse appropriate situations. She discovered that although intermediate level learners used null subjects in at least 80% of the contexts in which a native Spanish speaker would, the

learners still used forms that were more full or more explicit than expected. In general, learners overused proper nouns and definite NPs to maintain reference, even at an advanced level (Saunders 1999:122, 125, 157). This is not simply transfer from English: In many situations when learners used proper nouns, a pronoun would have been preferred or required in English (Saunders 1999:124). In addition, Saunders found that the number of null subjects used by learners was correlated with the complexity of the narrative structure, and was influenced by the issues outlined below.

Example 18: Influences on L2 Learner Use of Null Subjects (Saunders 1999:172, 174)

- (a) the role, familiarity, and relative importance of the story character;
- (b) the function of whether the character was being introduced, maintained or reintroduced; and
- (c) the syntactic position in the sentence.

The fact that these discourse features were shown to affect the use of null subjects raises the question of whether these issues might affect the use of determiners. The next study demonstrates the use of an anaphoric hierarchy to study overall L2 discourse cohesion instead of null subjects.

Overall Discourse Cohesion—Strömqvist & Day (1993)

The goal of Strömqvist and Day (1993) was to determine how L2 acquisition of discourse proceeds using an overall cohesion score based on an anaphoric hierarchy. Strömqvist and Day hypothesized that overall cohesion is enhanced by the features below. These features were used to assign an overall cohesion score to learners based on the number of times one of these requirements was satisfied versus the total number of times it could have been. Method (c) (below) for increasing overall cohesion is based directly on an anaphoric hierarchy. If an indefinite common noun is first used to identify

the main character, then either a definite common noun, pronoun, or proper noun should be used for later mentions. Strömquist and Day's limited anaphoric hierarchy pairs NP lexical forms.

Example 19: Ways to Increase Overall Cohesion of a Narrative (Strömquist & Day 1993:147)

- (a) Using a personal pronoun only to refer to the main character,
- (b) Referring to the main character only in the first grammatical position when beginning to describe something new, and
- (c) Making subsequent mentions of referents less lexically specific than their first mentions.

The first mention of a character in a narrative should be the form on the left, and subsequent mentions should be the form on the right in the outline shown below.

Table 10: Strömquist & Day's Anaphoric Hierarchy (Strömquist & Day 1993:141)

<i>First mention</i>	<i>Subsequent mention</i>
indefinite article + common noun →	definite marker and common noun
indefinite article + common noun →	pronoun
proper noun →	proper noun
proper noun →	pronoun

The top two pairs of first and subsequent mentions are described as relatively specific to Swedish, but the last two pairs are universal (Strömquist & Day 1993:141).

Strömquist and Day's study found that using these criteria for evaluating/improving cohesion, the overall cohesion of five adults learning Swedish fell over time. After first achieving a median cohesion score of 10 out of 12, seven to nine months later the same adults only had a median score of seven out of 12.²¹ As adult learners already possess fully developed cognitive skills allowing them to relate narratives, Strömquist and Day conjectured that the decrease in overall cohesion resulted from "an increased experimentation with morphology and lexical units on the part of the adult learners as a strategy for expanding their linguistic resources (Strömquist & Day 1993:153)." If this is correct, it suggests that the overall cohesion scores calculated in this way could be correlated with changes in L2 patterns of determiner use. Because half of the overall cohesion score was determined by the rankings of the anaphoric hierarchy alone, it is possible that using only an anaphoric hierarchy for the analysis would also indicate the same changes in overall cohesion. The next study reviewed does not use an anaphoric hierarchy, but rather employs a transition model.

L1 Transfer of Nominal Reference—Hendriks (2003)

Hendriks (2003) used an extremely limited transition model to study whether learners studying a variety of L2s could be said to use overly explicit NPs. The only transition she classified was maintenance, which she defined as "all linguistic expressions referring to a protagonist after the first act of referring to that particular protagonist (Hendriks 2003:299)." This definition ignores distance between references as a factor in reference maintenance, but even with this limited view, Hendriks found significant differences in the production of L2 German, French, or English by adult native speakers of Chinese.

²¹ Native adult speakers of Swedish were found to have median cohesion score of 10/12 (Strömquist & Day 1993:149).

The Chinese learners' reference maintenance did not approximate native speaker production of Chinese, French, German, or English, but learners were more target-like in French, possibly because the French-learners took a different narrative perspective. In the L2 English and German narratives, the stories were told from the perspective of a narrator. However, in the L2 French narratives, the stories were told from the perspective of one of the protagonists (Hendriks 2003:311-312). This alternate story perspective made it easier for the Chinese learners writing the French narratives to maintain a coherent focus in the story. It was also found that the L2 learners used bare nouns in German and English in places where native speakers would not—a pattern not appearing in the L2 French. In L2 French, however, learners used topic-promoting left dislocations that would be unacceptable in native French.

It is interesting that unlike the findings of previous studies, only the German production was remarkable for using overly long or explicit noun phrases to maintain reference, and this was because significantly fewer pronouns were used in the German L2 narratives compared with the others (Hendriks 2003:231). Hendriks' final conclusion was that the over-explicitness noted in other research studies in actuality depends on the native-target language pair (Hendriks 2003:322). This conclusion may be debatable, but such results imply that issues with selecting lexical forms for reference may not be related to the L1 transfer of discourse rules, but rather to the target language. The final study that will be surveyed is another that uses a transition model. The study below again looks at the use of null subjects, and again attempts to ascertain if L2 learners use overly explicit forms of reference.

Null Pronouns and Over-Explicit Reference—Nistov
(2001)

Nistov's (2001) study used a transition model of discourse that categorizes only shift and maintenance of reference. Nistov's definitions have an extra stipulation not

usually found in other transition models, though. She defines a shift as when a referent is not mentioned in the preceding utterance (Nistov 2001:59). Referent maintenance is defined as occurring when the referent is cited in the previous utterance, but Nistov makes the following caveat:

Preceding does not necessarily imply the immediately preceding utterance (= clause) if this clause does not contain reference to a character that contributes to the thematic progress in the story (Nistov 2001:60).

She admits, though, that “thematic progress” is difficult to define and use.

Relying on these transition criteria, Nistov examined the reference expression of three native Turkish speaking teenagers learning Norwegian. Overall, she did not find a pattern of use of overly explicit noun phrases. In contrast, she found that the Turkish teenagers used more null and overt pronouns than native Norwegian speakers. The Turkish teenagers especially used more zero anaphora than native Norwegian speaking teenagers. This result may not be surprising since Turkish allows for more null elements than Norwegian, but some of these null pronouns appeared in contexts where they would not be acceptable in Turkish. This finding is important because it shows that even if reference and transitions are not exactly like they would be in a learner’s native language, there still may be some L1 effect.

Summary of Findings

The L2 studies in the last section tend to focus on exploring the concept of L1 transfer. In addition, all of these studies use general models of how discourse works, that did not rely on binary features. Furthermore, all found that looking at discourse was important to understanding the nominal reference patterns of the learners. The table that follows summarizes the findings of these studies.

Saunders (1999) found that L2 null subject accuracy correlated with many other linguistic features related to discourse proficiency. Both Saunders (1999) and Strömquist and Day (1993) found that proficiency in encoding discourse decreased from one

interlanguage stage to another, although Saunders found that advanced learners in her study rebounded. None of the studies looking for L1 patterns in L2 discourse production were able to find exact L1 transfer. Hendriks (2003), though, found that the target language lead to differences in discourse constructing ability by learners from the same L1 at the same L2 proficiency level. Nistov (2001) hypothesized that despite a lack of exact L1 patterns in L2 production, there may be some effect of the L1.

Table 11: L2 Studies Using General Discourse Models and their Findings

<i>Study</i>	<i>Discourse model</i>	<i>L1 transfer found</i>	<i>Over-explicit reference found</i>	<i>Overall discourse proficiency found to affect L2 nominal reference</i>
Saunders (1999)	anaphoric hierarchy	no	yes (proper Ns & definite NPs)	yes
Strömquist & Day (1993)	anaphoric hierarchy, overall cohesion	?	?	yes
Hendriks (2003)	transition model	no	in German only (not French or English)	?
Nistov (2001)	transition model	yes	no	yes

One problem with the studies presented in this section is that they do not define clearly what the learning tasks are. The models of discourse these studies rely on do not completely specify which discourse meanings are universal and which are language specific. Therefore, in these studies it is not clear what parts of describing nominal reference must be learned with the rest of the L2-specific linguistic features. However, this is not unsolvable. Natural Semantic Metalanguage has defined certain “semantic primes” that are lexical items or meanings that appear common to human language. The

lexical items relevant to noun phrases and that express these universal meanings in English are:

- Substantives: *I, you, someone, something/thing, people, body,*
- Determiners: *this, the same, other/else,* and
- Quantifiers: *one, two, much/many, some, all.* (Goddard 2008:62)

If there are “semantic primes,” then there are semantic meanings or functional morphemes that are universally used to build reference in discourse. These meanings should be transferable, while other discourse-relating meanings instantiated in a particular language should be potential areas of problems when learning it as an L2.

None of the studies in this section could have reached their conclusions using a binary features model of discourse, or a language-specific discourse theory. In addition, it is important to note that all three studies searching for the influence of general discourse proficiency on the forms of nominal reference found some effect. These alternative discourse theories were useful in discovering this effect. For these reasons, further research into L2 determiner and noun phrase use that is not based on a binary-features model is advisable. The methods used in this study to do so are described in the next section.

Research Goals

The Three Main Research Aims

The primary goals of this research are:

1. To describe some characteristics of co-reference and discourse construction in native and L2 essays in English,
2. To determine where there are significant differences in co-reference and discourse construction between (a) native and L2 writers, (b) L2 writers from China and Korea, and (c) L2 writers of different proficiency levels, and

3. To discover L2 patterns of error in the selection of articles, determiners, pronouns, and other lexical items that contribute to co-reference and discourse cohesion.

Each of these research goals entail different research questions. Relevant questions are listed below each goal in the sections that follow.

Research Questions

Questions Related to Goals One and Two

Goal 1: To describe some characteristics of co-reference and discourse construction in native and L2 essays in English.

Goal 2: To determine where there are significant differences in co-reference and discourse construction between (a) native and L2 writers, (b) L2 writers from China and Korea, and (c) L2 writers of different proficiency levels.

The first two goals listed above are linked, and will therefore be investigated together for the most part. Some of the questions that this investigation will seek to answer for these goals relate to looking for randomness and patterns, as well as if there is any indication that proficiency level, text-type, L1, or L2 status makes a difference in various features that are relevant to discourse construction. Since syntax choices, vocabulary breadth, frequency rates, and transitions are all discourse-relevant features, these will all be investigated, but the basic questions that this research goal will focus on can be summarized as:

- Are learners behaving randomly in selecting any important aspects of discourse construction and cohesion?²²

²² The inherent assumption here is that whether learners are behaving randomly or not can be detected. It is possible that usage that appears on the surface to be random may not be. This is another concern related to the use of frequency data. The exact meaning assigned to a discourse morpheme cannot be determined. However, in this study, if the *production* seems to be randomly distributed between forms in similar contexts, this will be assumed to be an indication that learners are behaving randomly.

- If learners are not behaving randomly, then what patterns are they demonstrating in their writing construction?
- Are these patterns more similar to those of native speakers, those of a similar L2 proficiency, or those from a similar L1 background?

Much of the investigation related to these goals will be more general than just looking at articles and determiners. This is to build a foundation upon which to better understand the meanings that learners are associating with determiners and articles.

Questions Related to Goal Three

Goal 3: To discover L2 patterns of error in the selection of articles, determiners, pronouns, and other lexical items that contribute to co-reference and discourse cohesion.

The questions for Goal 3 specifically focus on articles, determiners, and other functional morphemes with significance in discourse construction. For this part of the investigation, looking at errors will be very important, as will using alternative models of discourse to try to discern the meanings that learners are assigning to functional discourse morphemes that have no correlates in their native language. Some of the questions that this study will attempt to answer for this goal are:

- If L2 writers show patterns of error, which discourse meanings are they relating to which forms?
- Are the errors only made with articles or certain types of determiners?
- Are the errors only made in particular meaning contexts?
- Are the learners using overly explicit reference?
- Are the learners over-using particular determiners or articles?
- Are the learners omitting determiners?

As can be seen, many of these questions are based on the findings of other studies on articles, such as those discussed in the introduction and literature review. So many different kinds of errors have been found in prior studies that it is important to determine

what if any relationship the use of articles by the L2 students here have with these prior studies. Once it is clear what the learners are actually doing in regards to article use, explanations can be explored. The next two sub-sections present goals for future research.

Questions for Future Research: Goal Four

Goal 4: To assess which theoretical models of discourse or article use are most effective in categorizing or explaining L2 patterns of error.

This goal and its questions build on Goal 3. For Goal 3, the kinds of errors that the L2 students are making will be found and categorized. Once these have been examined, then it can be determined how they might best be explained, or what theories and models account for the errors and correct usage of the learners the best. Some of the questions that future study will try to answer to reach this goal will be:

- If student writers are not selecting articles, demonstratives, pronouns, and full NPs randomly, then do their patterns of use correlate well with a particular theoretical model of discourse or explanation of L2 errors?
- If learners are making systematic errors, then what explanation or theory appears to best describe the choices that the student writers are making as they construct their essays?

The hope for the outcome of investigating this goal is that by being open to other models of discourse, a possible new perspective on the cause of article and determiner errors will be found, or a new method for explaining errors may be identified that can be studied further in future investigations.

Questions for Future Research: Goal Five

Goal 5: To search for evidence supporting or disproving certain L2 acquisition theories such as access to UG, L1 transfer, and L2 initial state.

Aside from simply trying to answer theory-specific, determiner-specific questions, later studies will examine data to see if there are indications of how second language acquisition in general proceeds that can be gleaned from looking at frequency data. The questions that may be considered are:

- In examining the hierarchies and discourse meanings that are in UG or are universal, is there evidence for adult access to UG?
- Are there significant differences in how L2 learners from different L1 backgrounds use NPs, construct discourse, or select determiners that would suggest L1 transfer is occurring?
- Are there other indications of L1 transfer, such as learners selecting forms parallel to those used in their native language for a particular discourse meaning?
- Is there any evidence indicating what the L2 initial state might be?

All of these goals are ambitious, but still do not cover the full gamut of possibilities of what could be examined using this corpus data or when investigating discourse construction. There are many areas open for possible further research, such as looking more in depth at the syntactic construction of noun phrases in the learner and native groups, or using recordings to examine some of the phonological considerations like pitch changes associated with functional morphemes, focus, and topic in discourse.

CHAPTER II. METHODS

Data Collection

The data used in this analysis consist of noun phrases that come from a corpus of writings by native and non-native writers of English. In the sections below, the major details of the process of data collection and analysis used in this investigation will be outlined. The first major section describes the collection of both the learner and the native writer corpus. After that, the various coding systems used to analyze and compare the discourse production of the learners and the native writers are described. Lastly, the statistical methods used to evaluate the significance of patterns found in the data are explained.

Collection of the Corpora of Essays

The corpus that supplies the data for this study was collected in two comparable parts (two corpora). One-half consists of essays by adult non-native English speaking students (the learner corpus). The other half contains writings by adult native English writers (the native corpus). This section describes the collection of the writings, focusing first on the learner essay set, and then on the native writer one.

Collection of the Learner Corpus

Timed essays from 20 students studying abroad at the University of Iowa were collected from the English as a Second Language department. These essays were written as part of English proficiency tests that students are required to take before beginning their coursework for the semester.

Description of the Learner Essays and Proficiency Exams

The two proficiency exams for which the students wrote the essays collected were the Iowa Intensive English Program placement exam (IIEP exam) and the University of Iowa's English Proficiency Evaluation (EPE).

The EPE: For Admitted Students

The EPE is administered to students who have been admitted to The University of Iowa, but whose entering TOEFL scores are below 600 on the paper-based test, below 250 on the computer-based test, or below 100 on the internet-based test. Minimum TOEFL scores required for undergraduate admission to The University of Iowa are 530 on the paper-based test, 197 on the computer-based test, or 71 on the internet-based test (refer to The University of Iowa's (2005) "English Language Requirements for Admission"). For graduate admission, TOEFL scores must be at least 550 on the paper-based, 213 on the computer-based test, or 81 on the internet-based test. The EPE is therefore not used to evaluate students' English for admission purposes, but rather to identify which students may need to take English as a Second Language classes to help them further improve those English skills that they will need to succeed in their regular coursework in English at the University.

The IIEP Exam: For Un-Admitted Students

In contrast, the purpose of the IIEP exam is to facilitate instruction. It is used to sort the students who will be studying with the Iowa Intensive English Program by proficiency level. IIEP students are not necessarily seeking a degree and may only intend to study to improve their English, or they may have been conditionally admitted to study at The University of Iowa. Conditionally admitted students have taken the TOEFL, but their scores are not high enough to reach the minimum requirements for full admission. In order to be considered for conditional admission to The University of Iowa, students' TOEFL scores must fall between 450 and 530 on the paper-based exam, 133 and 197 on the computer-based exam, or between 45 and 70 on the internet-based test. Conditional admission additionally requires enrollment in the Intensive English Program for up to one year. Thus, all conditionally admitted students will take the IIEP exam to be placed in the appropriate intensive language classes.

Conditionally admitted students can advance to full admission in a couple ways:

- Those in the advanced level may advance to regular admission if they receive grades of A or B in all of their IIEP classes and receive a recommendation from the Intensive English Program.
- Students may also advance to regular admission by achieving the required TOEFL score for regular admission, but they also must have a grade point average of at least 2.5 in their IIEP coursework.

Once students are regularly admitted, they will take the EPE.²³

Similarities Between the EPE and the IIEP Exam

For both the EPE and the IIEP exam, students were required to take a 30 minute timed essay test, and these essays from 20 students who took these tests (10 EPE and 10 IIEP) serve as the basis of the learner corpus. Aside from the essay test, EPE students took a multiple-choice reading test, and also were tested in a 10-minute one-on-one oral interview. IIEP students also had a similar oral interview, and took a 50 minute reading comprehension and vocabulary test. The reading comprehension test is typically administered before the essay test, and the oral proficiency test is administered last. When taking the essay test, both EPE and IIEP students were presented with two test questions and were asked to select one of them to write about. They completed their essays in a quiet lecture hall, using paper and pencils provided for them by the test proctors. When time was called, learners had to stop writing, even if they were in the middle of a sentence. However, they had the option to stop writing before the thirty minutes was up if they felt their essay was complete.

²³ Some of the students' whose EPE writing was collected for this corpus may have taken the IIEP written exam previously. However, there is no overlap between the writing topics offered on the EPE and IIEP exams, nor are the same topics regularly repeated. Every student would have been presented a novel topic. Furthermore, no two essays in the corpus are from the same student.

Essay Topics

The students' essay topics and how many students selected each one are summarized in the table below.

Table 12: Essay Topics in the Learner Corpus

<i>Topic title</i>	<i>Description of topic</i>	<i>Number of learners who selected it</i>
globalization	advantages and disadvantages of culture borrowing and globalization	3
population	description of a graph of past and projected population growth in the developed and developing world	7
being someone else	explanation of how the student's life would be different if they were someone else	3
emotional event	description of an emotional event in the student's life	3
overcoming challenge	description of how the student faced a challenge or obstacle in his/her life	1
most important possession	description of the student's most important possession and explanation of why	1
ceremony or tradition	description of a special ceremony or important tradition in the student's country	2

The fact that learners could choose between two essay topics and that the essays were collected from different testing periods means that not all of the 20 essays in the learner half of the corpus were written on the same topic. In actuality, the essays in the learner corpus each focus on one of seven different topics. Due to the fact that the essay tests allow the learner to select one of two topics, record-keeping, and the need to select learners from the same language or country background, there were not enough essays

available to build the corpus using all the same topics. Some topics were selected by more than one student, and some were written about by only one student.

Advantages and Complications Stemming from the Method of Collecting the Learner Corpus

One positive of collecting the learner essays from the English as a Second Language department is that timed essays are “authentic” materials produced by students in an academic setting. Granger (2002) outlines the problems that can result from collecting and analyzing corpora built of writing that is not “authentic”. In this case, “authentic” simply means that the writing task was one that students actually had to excel at for something they wanted in their real lives. To reach their academic goals, they had to do well on this test, and would have to take similar tests in the future. The exercise (for the students) was not one created simply for the purpose of collecting a corpus. The fact that the learner essays were not all written on the same topic can also be viewed in one way as an advantage. The purpose of writing and the text-type influence the structure of discourse. On one hand, being able to examine essays on a variety of topics means that learner patterns of discourse construction in a variety of text-types can be examined for differences in determiner use and noun phrase construction. On the other hand, this same “advantage” can be viewed as a complication. More data types are available, but there may not be enough uniformity to discern reliable patterns. The choice of using essays from these particular tests made complete uniformity of topic impossible.

Another complication caused by using these ESL essays is that fully comprehensive background information about the learners could not be collected. Both student privacy concerns and the purpose for which the ESL department maintains the records limited the amount of personal details that could be gathered. Information on students’ countries of origin was available, but there was limited availability of details about their exact native languages, genders, ages, years spent studying English, ages of

first contact with English, or lengths of time living in the United States. The snags that might result from this have been handled, for example, by collecting essays from ten students from China or Taiwan, and ten students from Korea. This was done in order to make the assumption (based in part on trends in what kind of English as a Second Language students typically study at Iowa) that students from Korea spoke Korean, and that students from Taiwan or China spoke Mandarin Chinese, possibly in addition to other Chinese dialects such as Taiwanese, Cantonese, or other Sino-Tibetan languages. The focus was placed on these language backgrounds because the acquisition of languages with article systems by native writers of Korean or Chinese is very common in traditional studies examining the L2 use of articles. Neither Korean nor Chinese (or other related Sino-Tibetan languages) have exact correlates²⁴ for the English articles *a/an* and *the*. Seeing how learners interpret and use these target-language discourse markers is instructive in regards to how they construct and manage discourse relationships.

Another potentially minor issue resulting from using these essays is that occasionally a student will take either the EPE or IIEP exam more than once. Whether any of these students were re-taking the exam is not known. However, even if students were re-taking the exam, they would not be presented with the same essay questions, so the situation of constructing discourse on a novel topic would be the same for all test-takers. Therefore, considered as a whole, the advantages from collecting the learner essays from the ESL department in this manner outweigh the disadvantages.

²⁴ There is debate regarding whether the Chinese *yige* “one” is actually currently functioning or transitioning to functioning as an indefinite article. However, even if one assumes that this is true, Chinese still does not have a correlate for *the*, and therefore seeing how the discourse meaning of this target-language term is interpreted by learners is still relevant to evaluating discourse models.

Collection of the Native Corpus

To build a corpus for comparison with the learner corpus, writing from a community sample of 20 native English writers was also collected. The native writers were self-selected (by answering an advertisement—see Appendix A). They were not compensated for their time, and no personal information was collected or retained about them other than the fact that they are native writers of English and all college graduates or professionals. No information was collected about gender, age, or exact education level so that data about the native writers would match that collected about the English language learners.

The Native Writer Testing Situation

The test situation for the native writers was similar to that of the non-natives, but not exactly the same. The native writers were asked to choose one of two essay topics and write for half an hour. The differences between the native writer testing situation and the non-native writer testing situation can be summarized thusly:

- The native writers were able to complete the essay test using either paper and pencil or by typing their response on computer;
- The native writers were allowed to select a time to take the test that was convenient to them;
- The native writers were asked to monitor the 30 minute time limit for themselves instead of being proctored; and
- There was not the same pressure of having a testing situation that could affect placement in college coursework.

This last point means that the native corpus does not meet Granger (2002)'s criteria for an authentic corpus in the same way that the learner corpus does. However, this should not significantly affect the analysis since the main use of the native corpus is comparison

with the learner corpus. The learner corpus and the patterns of co-reference within it are the primary focus. These patterns were produced in an authentic type of student writing.

Essay Topics

All the native writers were presented with the same two essay topics. Although these topics are similar to some of the topics that were presented to the non-native writers, they were not the same questions used on the EPE or IIEP test. The essay topics and how many native writers selected each one are displayed in the table below. The full text of the questions and instructions as they were given to the native writers appear in Appendix A.

Table 13: Essay Topics Given to Native English Writers

<i>Topic title</i>	<i>Description of topic</i>	<i>Number of learners who selected it</i>
age to enter bars	explain whether the local government needs to make it illegal for citizens under age 21 to enter bars and why	3
general education requirements	explain whether public universities should eliminate general education requirements and why	17

A full description of details such as the average length of the learner and native writer essays, the number of noun phrases, and the complexity of the noun phrases, is given later. The immediately following section lists and explains the different criteria that were used to categorize the data from the essays in preparation for analysis.

Methods of Coding Data

The data used in this analysis consist of noun phrases that come from the corpus of writings by native and non-native writers of English whose collection was described in the previous chapter. In the sections below, the various coding systems used to analyze and compare the production of the learners and the native writers will be outlined. After that, the statistical methods used to evaluate the significance of patterns found in the data will be discussed.

After collecting the essays, they were typed and saved as plain text files. Then they were manipulated in various ways for analysis. In addition, noun phrases were classified in a variety of ways in order to examine features relevant for passing along discourse meanings. The main ways in which NPs were marked for important discourse features consist of the following methods, which will be described in more detail either in this section or a later section when the results are reported:

1. Part-of-speech tagging: All words in the essays were automatically tagged for parts-of-speech (e.g., noun, verb, adjective, determiner...) using MontyLingua 2.1 (Liu 2004). Then, a perl program was used to isolate the noun and determiner phrases²⁵. After that, the number of NPs and total words were calculated. This was done mainly to support the analysis for the first and second research goals.
2. Co-reference and co-reference chain tagging: The NPs that referred to the same real-world objects or ideas were identified as co-referring. When several of these phrases referred to the same element over the discourse of an essay, this subgroup was noted as a chain of reference. Chains of reference are mental constructs that help language users organize concepts and build on information about a topic. An example of a very short chain of reference would be a pronoun

²⁵ Depending on the perspective of syntax one takes, this could be either a noun phrase or a determiner phrase. For the sake of convenience, I will refer to them as noun phrases (NPs).

and its antecedent. The two are mentally connected. This study will analyze the number of chains of reference, the number of NPs in each chain, and the forms of their constituents, but defining the formal means by which such connections are made will be left for later investigations. (Centering Theory (Grosz, Joshi, & Weinstein 1983, 1986, 1995; Walker, Joshi, & Prince 1998) outlines an algorithm for how people know what noun phrases to link from sentence to sentence in a discourse, but evaluating such an algorithm is beyond the scope of this study.) The forms of co-reference chain constituents will be examined by ranking chain members on a modified anaphoric hierarchy scale.

3. Error tagging: Perceived errors in the use of NPs were marked and classified according to type. The three major categories into which the NPs were categorized were: (a) not informative enough reference, such as missing the definite article, (b) overly informative reference, such as the use of a possessive when a definite article would suffice, and (c) inaccurate NP structure, obscure expression, or ambiguous meaning.
4. Transition tagging: NPs were classified as new, maintained, or re-mentioned referents. New transitions are those when a concept or thing is referred to in the discourse for the first time. NPs for which there is a co-referring NP in the previous sentence are labeled as maintained transitions. Re-mentioned referents are those NPs which have a co-referring mate somewhere in the previous discourse, but not in the immediately preceding sentence.
5. Grammatical category tagging: The syntactic position or context of NPs was labeled as either: (a) subject, (b) verbal object, (c) object of a preposition, (d) genitive specifier, (e) complement of the copula, (f) apposition, (g) comparison, or (h) title. Categories such as apposition, comparison, or title are not in general use, but were created for the purpose of this study because learners often make mistakes in such constructions, the NPs in such constructions are difficult to place

in the other categories, and these constructions have special discourse meanings. For example, phrasal appositions are asides and usually co-refer with the immediately preceding NP. They may maintain reference to an argument central to the discourse topic, but they themselves are not central. The archetypal phrasal apposition is extra information with more explicit reference that may be needed by some in the audience, but not all. Even when appositions refer to the subject of a sentence, the implication that they are background information not needed by all is a very different discourse function than the actual subject, which is typically central to maintaining old information or linking ideas key to the communication.

6. Noun-type tagging: NPs were sorted as common nouns, proper nouns, or other kinds of NPs, such as pronouns or lone demonstratives. One reason for this is that these three generally rank differently on anaphoric hierarchies, with pronouns and solo demonstratives ranking high on mental accessibility, and proper nouns ranking low because they are generally easy to comprehend whether they have been mentioned in the discourse before it not. Common nouns typically fall somewhere in between these two for mental accessibility on anaphoric hierarchies. Another reason is that proper nouns and pronouns do not typically co-occur with articles as common nouns do. Finally, one of the learner essay topics centered on describing a population graph. Because of this, there were more proper nouns in essays on this topic, and this classification allowed that difference to be explored.
7. Word-type classification or concordancing: As mentioned above, the number of total words was counted, but the number of different types of words and how often each word was used was also evaluated for each essay. A perl program was used to list each word in each essay and automatically count how often it

appeared. Certain patterns regarding word frequency are well-known,²⁶ and looking at the concordances for the essays allowed the learner essays to be evaluated against these expectations. Furthermore, this allowed for judging the student essays for variety in vocabulary use compared to the native writers and by proficiency level.

The majority of these classification methods will be described at the beginning of the related results section. In the rest of this section, though, issues concerning part-of-speech tagging, co-reference tagging, and computational methods of linguistic analysis will be further examined. Specifically, the discussion to follow will cover:

- Theoretical concerns about defining parts of speech,
- The differences between human and machine part-of-speech tagging,
- The mechanisms by which automatic taggers such as MontyLingua-2.1 function,
- How MontyLingua-2.1 output was modified and corrected for this study,
- A brief overview of different co-reference tagging schemes, and
- How the MUC-7 co-reference tagging scheme used in this study operates.

At the end of this chapter, the statistical analysis methods will be outlined.

Part-of-speech Tagging

All words in the essays were marked for part-of-speech. The purpose of this was to facilitate the identification and isolation of noun phrases and determiners for closer analysis. As much as possible, the process of part-of-speech tagging was automated using the Python version²⁷ of the linguistic analysis program MontyLingua-2.1 (created by Hugo Liu of the MIT Media Lab, version copyright 2002-2004, and available at Mr.

²⁶ See, for example, Zipf's law.

²⁷ Python is an uninterpreted computer language. There is also a Java version of MontyLingua-2.1 that is available. Thanks to Bob Ahrens for help with some of my Python code. The code used to call MontyLingua to tag parts-of-speech appears in an appendix.

Liu's webpage). MontyLingua-2.1 relies on the PennTreebank definitions of parts-of-speech. The PennTreebank is a large corpus of articles tagged for syntactic structure to allow for the quick location of particular grammatical constructions. As one step toward parsing the syntax of sentences in the PennTreebank corpus, parts-of-speech were differentiated. However, defining parts-of-speech is still debated in linguistics. Before the decisions that were made by the PennTreebank are examined, some theoretical considerations about part-of-speech tagging will be outlined, followed by the contrasting of human and machine tagging in this area.

Defining Parts-of-Speech

Although people have intuitions about what parts-of-speech are, they are difficult to formally define. Linguists and language teachers typically apply reason and a variety of general criteria to classify words according to their parts-of-speech. In contrast, computational or automated algorithms, which cannot rely on intuition or higher reasoning, follow simplified rules to determine parts-of-speech. The results of machine categorization are not as accurate as human classification, but can be achieved more quickly.

Human Classification of Parts-of-Speech

Human coders use several criteria at once to determine part-of-speech because not all words can be easily categorized. For example, the English word *wound* could be either a past tense verb, or a noun. Without seeing its syntactic environment, the part-of-speech is unclear. For other words, not even the context is sufficient, and other criteria must be used to reason which part-of-speech is most appropriate. In some cases, human coders will not be able to agree on the part-of-speech. The criteria most commonly used by human coders to distinguish parts-of-speech are:

- Shared morphological characteristics, such as receiving the same inflectional or derivational affixes. *Examples:* Some nouns in English can be followed by the plural *-s* suffix. Other English nouns are formed by adding the suffix *-ation*.
- Similar syntactic positions or distribution in sentences. *Example:* Common nouns can occur after *the* in English.
- Some degree of semantic relatedness. *Example:* English nouns are sometimes described as being words that refer to real-world people, places, things, or ideas.
- Shared phonological characteristics. *Examples:* In English prosody, prepositions and pronouns do not typically receive sentence stress, but rather are reduced in pitch, duration, and loudness when compared with similarly positioned nouns, verbs, adjectives, or adverbs. In addition, there is a pattern in English in which certain two-syllable words stressed on the first syllable are nouns, while in contrast a certain two-syllable words stressed on the second syllable are verbs (e.g., nouns: record [¹ˌɹɛkɔːd], notebook [¹noʊt.bʊk], window [¹wɪn.dəʊ], hookup, promo; verbs: record [ɹɛkɔːd], look up [lʊkʊp], promote [pɹəˈmoʊt]).
- Comparison with words of similar meaning in other languages. *Example:* In English the word *tall* in a sentence like *He is tall* would be described as an adjective. In Mandarin Chinese, this sentence would be expressed as 他高 *ta1 gao1*, “*he tall*.” Many would classify 高 *gao1* “*tall*” as an adjective because the word of similar meaning in English is in a syntactic position like an adjective, and takes adjectival morphology. In Chinese, though, this is not as clear, and many question whether 高 *gao1* “*tall*” is truly a verb in Chinese because the lack of the copula does not give it the same syntactic position as an adjective in English, and because there is little or no real inflectional or derivational morphology in Chinese to use as assisting evidence. According to Baker (2003:6), in many languages, it is difficult to distinguish between intransitive stative verbs and

adjectives in this way, and some languages may not distinguish adjectives and nouns. This may increase the use of parallels with other languages in order to place words in different categories. Whether there is significant variation in regards to part-of-speech categories cross-linguistically has yet to be definitively determined.

- Based on theoretical considerations. *Example:* Principles & Parameters uses the binary features $\pm N$ and $\pm V$ to define four lexical categories—noun, verb, adjective, and adposition (Baker 2003:3). Although the theories have been developed based on words that have already been categorized for part-of-speech using other criteria, once the theories have been committed to, one might determine the parts-of-speech in a new example based on what the theory predicts.
- Intuition. *Example:* Not all words that might be placed in the same lexical category share common features in all the previously listed areas. Rarely is one criteria sufficient.

The way that these criteria for parts-of-speech apply to commonly and less-commonly used lexical and the functional categories are summarized in two separate tables that follow on the landscaped pages. The tables' information comes from Baker's (2003) book Lexical Categories and Haegeman and Gueron (1999). Haegeman and Gueron's explanation focuses on the characteristics of parts-of-speech in English, and how parts-of-speech are viewed under the Principles & Parameters (Chomsky) theory of syntax, and not the currently more commonly used theory of Minimalism.

Table 14: Criteria for Determining Parts-of-speech in English—Four Commonly Used Lexical Categories

	<i>Nouns</i>	<i>Verbs</i>	<i>Adjectives</i>	<i>Adverbs</i>
<i>Inflectional and derivational morphology</i>	plural morpheme, genitive morpheme (Haegeman & Gueron 1999:54-5)	inflectional morphology (number, tense, person) (Haegeman & Gueron 1999:56)	morphology like comparison or superlative suffixes (Haegeman & Gueron 1999:56-7)	derivational suffix <i>-ly</i>
<i>Syntactic position or words with which they commonly co-occur</i>	preceded by <i>the</i> or possessive pronoun (<i>his</i>) (Haegeman & Gueron 1999:54-5)	preceded by modals <i>will/ can/ must</i> , or <i>to</i> for the infinitival form (Haegeman & Gueron 1999:56) take complements (Baker 2003:14 from Hale & Keyser)	co-occurrence with degree words like <i>more</i> and <i>most</i> , <i>very</i> , <i>so</i> , <i>too</i> , <i>quite</i> , <i>rather</i> , <i>that</i> , <i>how</i> in this location: <i>the</i> —noun, possessive pronoun—noun (Haegeman & Gueron 1999:56-7) APs can be appended to transitive clauses to indicate goal or result of action modify nouns, and can be preceded by measure phrases (Baker 2003:2) form predicates, requiring a subject (Baker 2003:2 from Hale & Keyser)	can occur after degree words like <i>very</i> , <i>more</i> , <i>so</i> , <i>too</i> , <i>quite</i> , <i>rather</i> , <i>that</i> , <i>how</i> not same distribution as adjectives (Haegeman & Gueron 1999:57-8)

Table 14—Continued

	<i>Nouns</i>	<i>Verbs</i>	<i>Adjectives</i>	<i>Adverbs</i>
<i>Semantic characteristics</i>	concrete entities, objects, people, abstract entities (Haegeman & Gueron 1999:54-5) denote things, “long-term states of affairs” (Baker 2003:2 from Hopper & Thompson, Givon) “words typically used to refer” (Baker 2003:2 from Croft, Hengeveld, Bhat)	“denote events, which are dynamic, short-term states of affairs” (Baker 2003:14 from Hopper & Thompson, Givon) used to predicate (Baker 2003:14 from Croft, Hengeveld, Bhat)	states or properties, “typically medium-length states of affairs” (Baker 2003:2 from Hopper & Thompson, Givon) used to modify (Baker 2003:2 from Croft, Hengeveld, Bhat)	manner of completing an action—how, when, or where the event took place
<i>Theoretical considerations</i>	+N, -V / +subj, -obj (Baker 2003:2) +N = “has a referential index” (Baker 2003:2)	-N, +V / +subj, +obj (Baker 2003:14) +V = “has a specifier” (Baker 2003:21)	+N, +V / -subj, -obj not nouns and not verbs, -N, -V (Baker 2003:2)	---
<i>Cross-linguistic comparison</i>	never seem to allow noun incorporation (Baker 2003:2)	may allow noun incorporation in some languages (Baker 2003:14)	difficult to distinguish from stative verbs in many languages	---
<i>Class status</i>	open class	open class	open class	open class
<i>Phonological characteristics</i>	receive sentence stress	receive sentence stress	receive sentence stress	receive sentence stress

Table 15: Criteria for Differentiating Parts-of-speech in English—Other Frequently Assumed Categories

	<i>Prepositions (or adpositions)</i>	<i>Determiners</i>	<i>Auxiliary verbs</i>
<i>Inflectional and derivational morphology</i>	morphologically invariant (Haegeman & Gueron 1999:58)	some invariant, some only partly invariant (vary by number only: possessives) (Haegeman & Gueron 1999:59-60)	may take inflectional morphology (number, tense, person)
<i>Syntactic position or words with which they commonly co-occur</i>	followed by <i>the</i> + noun some take clausal complements (Haegeman & Gueron 1999:58)	before nouns (but not always for demonstratives) in complementary distribution (Haegeman & Gueron 1999:59-60)	before verb
<i>Semantic characteristics</i>	they “have a lot of lexical content” (not further specified) (Haegeman & Gueron 1999:58)	generic, specific, definite, indefinite	inflectional meaning, similar semantic functions (time, mood, aspect) (Haegeman & Gueron 1999:62-3)
<i>Theoretical considerations</i>	-N, -V / -subj, +obj (Baker 2003:2 from Jackendoff)	---	---
<i>Cross-linguistic comparison</i>	may not exist in all languages	may not exist in all languages	may not exist in all languages
<i>Class status</i>	open class (Haegeman & Gueron 1999:58) functional, not a lexical category (Baker 2003:2)	closed class (Haegeman & Gueron 1999:59-60)	closed class (Haegeman & Gueron 1999:62-3)
<i>Phonological characteristics</i>	usually phonetically reduced, not receiving sentence stress	often phonetically reduced (Haegeman & Gueron 1999:59-60)	phonetically reduced (Haegeman & Gueron 1999:62-3)

How Machine Classification of Parts-of-Speech is Different from Human Coders' Classification

MontyLingua-2.1 and its component MontyTagger use the PennTreebank tagset for marking parts-of-speech (Santorini 1990—see the full criteria in Appendix B or the PennTreebank website). Unlike human coders, who typically differentiate five to eight categories, the PennTreebank tagging system categorizes words into 45 different word classes (Marcus et al. 1993). This may seem large, but the PennTreebank in fact has a small set of categories compared to some other computational projects (see Jurafsky & Martin 2000:288). For example, the Brown Corpus, one of the most commonly used English corpora, classifies words into 87 different categories (Francis 1979, Francis & Kucera 1982), while the C7 Tagset, a part-of-speech tagging method commonly used in computational linguistics, has 146 classifications (Leech et al. 1994).

This wide variation in the number of classifications highlights a fundamental difference between theoretical categorizations of parts-of-speech, which typically strive for finding the fewest number of categories possible, and computational linguistic classification systems, in which creating as many classifications as necessary to meet the goal is acceptable. Small differences that demonstrate limitedly used patterns may be given new groupings in order to improve the accuracy of the application which employs the part-of-speech tagging or the accuracy of the tagging program itself. The number of classifications is not a burden because computers can search and apply a large number of rules in seconds. Minimalism is also not valued in computational algorithms because they are focused on accuracy in application and not determining the commonalities in structure of all human languages.

There are many different part-of-speech tags in the PennTreebank system that would be considered one category in a simpler system. For example, are eight different tags for verbs in the PennTreebank tagset. See the following table.

Table 16: The PennTreebank's Eight Categories that Correspond to the Category "Verb"

<i>Specific verb type</i>	<i>PennTreebank tag</i>
gerund or present participle	VBG
modal verb	MD
past participle	VBN
particle	RP
past tense verb	VBD
present tense verb, other than 3 rd person singular	VBP
present tense verb, 3 rd person singular	VBZ
base form of a verb	VB

These tags separate different types of verbs based on their tense, agreement, or morphology in order to facilitate searching for information or other uses of corpora. For example, if researchers want to look at sequences of past events, they can easily isolate past tense verbs by searching for the tag VBD.

In computational linguistics, there is a split between the methods used by humans to tag parts-of-speech, and that used by computer algorithms. Manuals like the Penn Treebank manual are used by people to create the machine rules for tagging so that methods can be shared and used consistently, but they are also to tag "gold standard" test corpora against which the performance of machine taggers is evaluated. Each "gold standard" test corpus is considered perfectly accurately tagged, and the tags are determined by one or more human raters. Human raters use the tagging manuals to make test corpora, but also to help make the list of instructions used by the computer taggers and the dictionaries that list words with their possible parts of speech already specified.

The PennTreebank classifies words as different parts-of-speech or word types using many of the same techniques discussed in the last section on theoretical perspectives on parts-of-speech. Santorini (1990) is a coding manual that enumerates all

the categories and provides specific details on how words should be categorized. In some cases, categories are defined by looking at the distributional context of the items in question. Distribution rules can usually be easily transferred to instructions in a computer algorithm. Since the PennTreebank coding manual was developed using actual texts from the Brown corpus, distribution is often described by using excerpts from real texts showing the position. This may be more helpful for the human coder, by building the ability to use intuition to recognize the part-of-speech. The distribution is also indicated by giving a formula for its location, which can help with the computer rule. One example of how the distribution is indicated in the Penn Treebank tagging guide is “compounds of the form *n-th X-est*, like *fourth largest*” are tagged as adjectives with the tag JJ (Santorini 1990:3).”

For human coders, the feature of defining parts-of-speech that the Penn Treebank manual relies on the most to distinguish the word classes is probably intuition (because humans can use intuition). The description of how the tag for “Foreign Word (FW)” relies completely on a tagger’s intuition or opinion:

Example 20: The Category “Foreign Word” is Tagged by Intuition

Use your judgment as to what is a foreign word. For me, *yoga* is an NN, while *bête noire* and *persona non grata* should be tagged *bête/FW noire/FW* and *persona/FW non/FW grata/FW*, respectively. (Santorini 1990:5)

For the computer rule that might be created to tag foreign words, a dictionary or fairly exhaustive list of what these are would probably be created, unless there were similarities in spelling or morphology that could be productively used to locate the foreign words.

Most of the tagging manual is taken up by explaining cases in which human raters seem to have trouble agreeing on the part-of-speech tags. When there could potentially

be strong disagreement between raters, or when intuition is not enough to allow people to consistently agree in the tagging, other methods are defined. Sometimes this means exhaustive lists for use even by human raters, which clearly works better for closed classes. Modal verbs are one category that is defined mainly by list.

Example 21: Penn Treebank List for the Part-of-Speech Category *Modal Verb* (Santorini 1990:5)

“This category includes all verbs that don’t take an –s ending in the third person singular present: *can, could, (dare), may, might, must, ought, shall, should, will, would.*”

In the modal verb example above, though, morphological similarity is used to distinguish this class in addition to the list given. It is noted that modal verbs never appear with the third person singular present tense marker -s. When exhaustive lists cannot be created, or when they would be too long for the tagging manual, then distributional contexts, semantic meanings, or morphology are heavily relied on, even in situations in which there might be disagreement between human taggers. In this example that talks about adjectives, the position or distribution is included: “compounds of the form *n-th X-est*, like *fourth largest*” are tagged as adjectives with the tag JJ (Santorini 1990:3). Semantic meaning, though, is key in the example from the definition of the tag WRB for Wh-adverb shown on the next page. This definition tries to distinguish two meanings associated with the word *when* so that it can be tagged in two different ways. The distribution in this case would not be useful because the two differentiated types appear before a subordinate clause, and the clause can appear in different positions in relation to the independent clause. Also, there is no morphology that can be used, because neither one would take any affixes in these positions.

Example 22: Semantic Meaning in Tagging: Wh-Adverbs (Santorini 1990:8)

When in a temporal sense is tagged WRB. In the sense of “if,” on the other hand, it is a subordinating conjunction (IN).

Examples:

When/WRB he finally arrived, I was on my way out.

I like it when/IN you make dinner for me.

To summarize, computational linguists use more categories of parts-of-speech than theoretical linguists because they are focused on application rather than understanding the common structure of human language. Furthermore, they write and use rules for human coders to use in addition to rules used by computers to find parts-of-speech so there is a perfect model against which the accuracy and efficiency of computer tagging algorithms can be evaluated.

MontyLingua-2.1, the automatic, computerized part-of-speech tagger used to find noun phrases in the essays examined in this study, uses the PennTreebank tagset described above as its basis. The next section will look in detail at how the computer algorithm in MontyLingua actually works.

The Computer Algorithm Used by MontyLingua-2.1 to Tag

Parts-of-Speech

MontyLingua uses the PennTreebank tagset to define how many categories of words it will mark, and also to make training sets of “perfectly” tagged corpora for the algorithm in the automatic tagger to learn from. There are three key components of the tagger: (a) default training files, (b) a set of dictionaries or lexicons, and (c) the list of steps or instructions in its algorithm, which incorporates probabilistic decision-making and is based on the Brill94 algorithm. MontyLingua-2.1’s tagging component, MontyTagger, was initially built to tag English non-fiction writing, and the author of the

program Liu estimates that the version in the Python computer language can tag approximately 200 words per second with about a 96% level of accuracy (more about whether this is actually a good level of accuracy will be discussed in sections to come). The Python program used to initiate MontyLingua-2.1 to tag data for this study is given in Appendix C. MontyTagger is a Transformation Based Tagger, which means that it relies on distribution rules applied in series. The difference between Transformation Based Taggers and other taggers, as well as how the Brill algorithm works will be discussed in the next section. The purpose of these upcoming descriptions is to explain exactly how the basic part of speech tagging and phrasal parsing was accomplished automatically. Part-of-speech tagging was a necessary first step in the analysis of the data of this study.

How Automatic Part-of-Speech Taggers Function

Three methods of automatically identifying parts-of-speech with a computer program are commonly used: (a) Rule-Based Tagging, (b) Stochastic Tagging, and (c) Transformation-Based Tagging, which is what is utilized by the MontyTagger. Since the speed of computer processing allows for taggers to use large lists, dictionaries, or arrays of the most commonly encountered words paired with their possible parts-of-speech, the most difficult actions that an automatic tagger must undertake are:

- Determining the part-of-speech of words in the lexicons that could be used as more than one possible part of speech, and
- Deciding how to tag novel words—words that do not appear in the dictionaries or the rules of the tagger (see Jurafsky & Martin 2000:300).

To deal with these two issues, algorithms summarize and rely on distribution, associated morphemes, word order, and surrounding words. For example, the word *the* could only possibly be one part-of-speech—a determiner. The tagger's lexicon will list

that *the* is a determiner, and the tagger will not have to perform any action other than looking up the lexical item.

The part-of-speech of a subset of words, though, cannot be determined when the word is taken out of context. Some words can be used as a noun, verb, and/or adjective without affixing different morphemes. For example, the word *frame* out of context could be either a noun or a verb. The fact that its part of speech is ambiguous is a problem for the automatic tagger. At this point, one possibility is to cause the tagger to look at adjacent words and compare them against rules describing the distribution of words, such as nouns commonly follow articles or adjectives. Another possibility would be to use probability, and choose the part-of-speech solely on the basis of whether *frame* is most often used as a noun or a verb. Dictionaries, distribution rules, and probability are central to almost all automatic tagging programs.

Three Common Approaches to Automatic Tagging

Three common methods used to automatically tag parts-of-speech are:

- Rule-Based Tagging,
- Stochastic Tagging, and
- Transformation-Based Tagging.

Rule-Based Taggers

Rule-based taggers are called rule-based because they use large sets of rules to disambiguate parts-of-speech. Typically these rules refer to the syntactic distribution of parts-of-speech, and they are written by a human programmer, not a computer program. The programmer looks at a sample corpus and tags it by hand, then creates rules that could be used to produce the tags in the sample. The rules are then tested against different corpora and refined by the programmer. Again, because computers can quickly read and apply long lists of instructions, accuracy of the final tags is valued over minimizing the number of rules, so some of the sets can be quite large. For example, the

ENGCG (English Constraint Grammar) tagger, “the most comprehensive system” (Barnett et al. 1996), has 1200 grammar-based constraints.

Stochastic Taggers, Including Hidden Markov Models

In contrast to Rule-based taggers, Stochastic taggers rely on probability over distribution rules. The probabilities used by a Stochastic tagger come from a model (“gold-standard”) corpus that has been tagged by human raters. The Stochastic tagger then reads the corpus and stores the words linked to their corresponding tags in a dictionary for lookup. If all the words in this dictionary could be associated with only one part-of-speech tag, then the probabilistic part of the tagger would be unnecessary. However, because a large portion of words with the same spelling are associated with more than one, the next step in the computer program is the computation of the probabilities. These are calculated by looking at strings of two words (bigrams) or three words (trigrams) in the model corpus. Looking at the example sentences that follow, it can be seen that for the word *frame* two of the trigrams that would be considered are ‘to *frame* him’ (to/TO *frame*/VB him/PRP²⁸) and ‘picture *frame* around’ (picture/NN *frame*/NN around/IN). Two probabilities that the Stochastic tagger would calculate would be the probability that *frame* is a verb when it follows ‘to’, and the probability that *frame* is a noun when it precedes a preposition or the word ‘around.’

The Hidden Markov Model (HMM) tagger is one of the most common kinds of Stochastic part-of-speech taggers. Markov models are finite state machines in which there are a certain number of states, and the probability of the transition from one state to the next is calculated.

²⁸ See Appendix A for a complete description of the Penn Treebank part-of-speech tagset. In the example sentences, the tags mean the following: DT = determiner, IN = preposition or subordinating conjunction, NN = singular or mass noun, PRP = personal pronoun, TO = the word ‘to’, VB = base form verb, VBD = past tense verb, VBG = gerund or present participle verb, VBN = past participle verb (Santorini 1990:7).

Example 23: Two Sentences Illustrating Example Trigrams in a Part-of-speech Tagged

Test or Model Corpus

(1) The gangster said that the police were trying to frame him for murder.

the/DT gangster/NN said/VBD that/IN the/DT police/NN were/VBD trying/VBG to/TO frame/VB him/PRP for/IN murder/NN (see Appendix B for PennTreebank tagset)

(2) The picture frame around the painting came unglued yesterday.

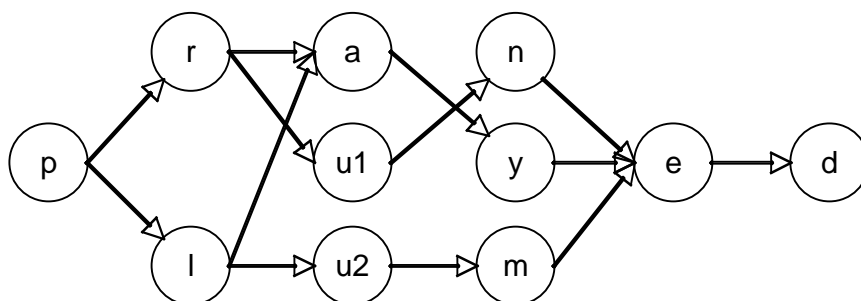
the/DT picture/NN frame/NN around/IN the/DT painting/NN came/VBD unglued/VBN

If you compare a Markov model to a train line, the stations would be the states, and the transitions would be the tracks. There are a finite number of stations where the train could be, just as there are a finite number of states that the machine can be in. If the train starts at the beginning station, and there are two rail lines going to the next station, the probability that the train can take either set of tracks can be calculated. If there are two possible transitions to two different states, the probability that the machine will use each one of the transitions can be calculated. Once the train or the machine reaches the final station or state, the probability that all the transitions actually taken would have been followed can be calculated.

There are many ways to use finite state machines to manipulate and study language. Aside from part-of-speech tagging, finite state machines are commonly used to model problems like speech recognition. In speech recognition, the states are different phonemes, and the probabilities of the transitions reflect the likelihood that a particular phoneme would follow one that has been identified. The speech recognition program can use the probabilities to guess or estimate what the word might be before it is finished. Psycholinguists have studied finite state machines to try to determine how adequate a

model it is for human speech perception. Because some sequences do not occur in any human language, or in particular languages, this kind of finite state machine to a certain extent models syllable structure constraints. In the following figure, a finite state machine that can result in the words *prayed*, *played*, *pruned*, and *plumed*, but not **pluned*, is shown. **Pluned* is not an English word, and may only occur in the case of a speech error, so there is a “u1” and a “u2” to make this outcome impossible. However, there is no particular reason why **pluned* cannot be a word in English. The onset and the coda consonant combinations are allowed in other words in English.

Figure 8: Finite State Machine that Produces *prayed*, *played*, *pruned*, and *plumed*, but not **pluned*

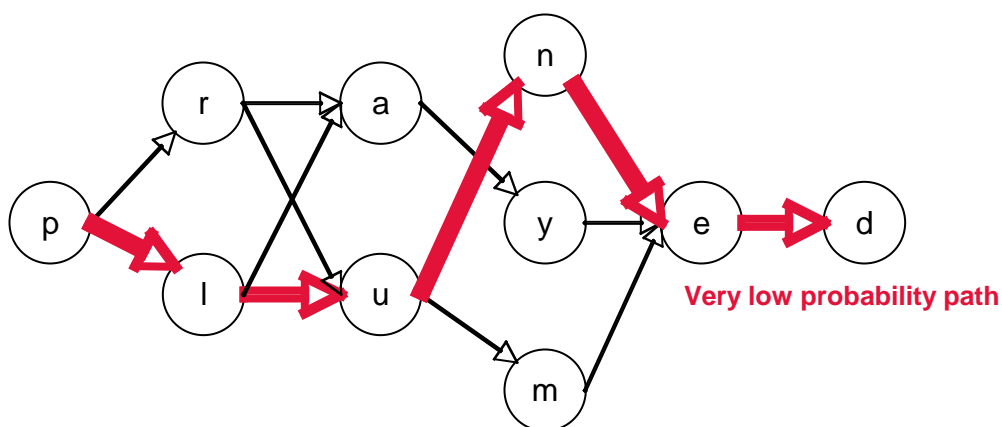


To handle this in a finite state machine, there could be only one “u”, but the probability of moving along the “p-l-u-n-e-d” path would be very low. See the next figure.

Hidden Markov models differ from traditional Markov models in that the states of the finite state machine are hidden. To describe it in a slightly different way, the states are not as important as the input and the outcome. For example, a word and its context are input into a Hidden Markov model used to automatically tag parts-of-speech. The model’s finite state machine’s states and the transitions are used to determine the most

probable part of speech for that word in that syntactic context. Then the part-of-speech is output. For the user, there is no advantage to knowing the actual states and transitions of the finite state machine used to arrive at this answer. The Markov model behind the selection of the outcome is hidden.

Figure 9: Finite State Machine that Produces *prayed*, *played*, *pruned*, *plumed*, and
**pluned*



One reason for the Hidden Markov Model's popularity is the availability of computers. It would be complicated and time-consuming for a human to calculate the conditional probabilities in Hidden Markov models used for automatic taggers, but the computational power and speed of computers make their use widespread. In fact, most HMM taggers will output tag sequences for entire sentences instead of doing so word by word (Jurafsky & Martin 2000:303), which would be an almost impossible calculation for a human tagger given a long sentence. The probabilities in Hidden Markov models use Bayesian statistics, or conditional probabilities. In basic terms, the statistics maximize the probability that a given word will have a particular part-of-speech also given the probability that a preceding number of words have particular given parts-of-

speech. Jurafsky and Martin (2000:303) summarize this by saying that the probability of this needs to be maximized:

- $P(\text{word}|\text{tag}) * P(\text{tag}|\text{previous } n \text{ tags})$.

Spelled out this means that taggers seek the highest probability that there would be a particular tag given the word in question *and* the probability that you would get the previous so many tags given that tag. If using trigrams, this would mean looking at the tags of the two words before the word in question. Bayes' Rule is shown in the figure below.

In order to use Bayes' Rule and Hidden Markov Models to infer tags, some data must be input as an estimation of the distribution of probabilities in previous texts. In general, hand-tagged (human-tagged) training corpora are used for this. If trying to tag parts-of-speech in news articles, for example, a training corpus of similar articles or a set of news articles would be used for training in order to increase the accuracy of tagging novel articles. Training corpora are not strictly necessary when using these kinds of stochastic taggers, but they help produce higher quality results.

Figure 10: Bayes' Rule (as cited in Thornberg 2006)²⁹

$$P(x|y) = \frac{P(y|x) P(x)}{P(y)} = \frac{P(y|x) P(x)}{\int P(y|x) P(x) dx}$$

Transformation-Based Taggers

Transformation-based taggers are seen as combining aspects of rule-based and stochastic taggers. As in rule-based taggers, constraint grammar rules are used to select

²⁹ An alternative version of this equation replaces the integral sign with the summation sign (Σ) for cases in which the probability is discrete rather than continuous (Parzen 1960:119).

parts-of-speech for words that could possibly be associated with more than one. Similar to stochastic taggers, training corpora are used to infer how tagging should proceed in novel corpora. Transformation-based taggers include dictionaries of words correlated with all of their possible parts-of-speech to look up tags. If there is only one possible tag for a particular word, then the application of rules to determine which tag is most likely is unnecessary. If more than one tag is possible, then syntactic distribution rules created from the pattern of tagging in the training corpus are applied. Therefore, the main difference between rule-based and transformation-based taggers is that constraint grammar rules are written by human programmers in rule-based taggers. In contrast, Transformation-based taggers use training corpora to automatically create constraint grammar rules. The assumption is that other corpora of a similar genre or text-type will have probably have similar distribution patterns.

The most commonly used style of transformation-based tagger is based on the algorithm developed by Eric Brill in 1994. In fact, this is what the automatic part-of-speech tagger MontyTagger uses. The next section will describe the Brill algorithm in more detail. In the table on the next page, there is a comparison of the three major types of part-of-speech taggers.

The Brill (1994) Transformation-Based Tagging Method

The goal of Brill's (1994) algorithm is to select the most likely part-of-speech by learning and using patterns from already tagged material. Programs that employ Brill's process use dictionaries, machine learning, probability, and the iterative application of rules to tag parts-of-speech in unencountered texts. The dictionaries consist of a list of all the word tokens in the training corpus, arrayed with their parts-of-speech and the percentage of time each instance of the word appeared with a particular tag. These percentages are then used as probabilities that an example of a word in a new corpus will be a particular part-of-speech.

Part of the machine learning involved then, is the “learning” of these probabilities through the creation of the lexical array. The rules are generally syntactic distribution, word order, or word co-occurrence rules that are also “learned” from the pre-tagged training corpus, and are created by the learning algorithm by filling in the blanks in rule templates.

Table 17: Comparison of the Main Features of Three Part-of-Speech Tagging Methods
(information from Jurafsky & Martin 2000:300-303)

<i>Feature</i>	<i>Rule-based taggers</i>	<i>Stochastic taggers</i>	<i>Transformation-based taggers</i>
<i>theoretical foundation</i>	Constraint Grammar architecture of Karlsson et al. (1995)	Bayesian statistics, Hidden Markov Models	Brill (1994)
<i>training corpus used?</i>	yes, by humans	yes, by machines	yes, by machines
<i>probability used?</i>	no	yes	no
<i>syntactic distribution rules used?</i>	yes, hand-written	no	yes, machine inferred
<i>dictionaries or lexicons used?</i>	yes	yes	yes
<i>commonly cited tagger names</i>	ENGTWOL, ENGCG	HMM tagger, Maximum Likelihood tagger, Markov Model tagger	Brill tagger, MontyTagger

These rules are described as “transformations” because they are written to instruct the changing of part-of-speech tags already assigned by probability or another rule to a different part-of-speech. One of the key difference’s between Brill’s transformation-based learning and how constraint grammar or stochastic tagging processes work is this repeated reading and re-labeling of the words in the corpus. An example of one kind of rule template filled in by the machine learning by interpreting patterns from a “Gold-

Standard” tagged corpus is given in the next example. Rules such as these are designed to capture generalizations such as “nouns follow determiners.” It is clear that such rules are not always correct. An adjective, for example, can follow a determiner, and be followed by a noun.

Example 24: A Brill Transformation Rule Template (Jurafsky & Martin 2000:310)

*Change tag **a** to tag **b** when:*

The preceding/ following word is tagged **z**.

The word before/ after is tagged **z**.

One of the two preceding/ following words is tagged **z**.

One of the three preceding/ following words is tagged **z**.

The preceding word is tagged **z** and the following word is tagged **w**.

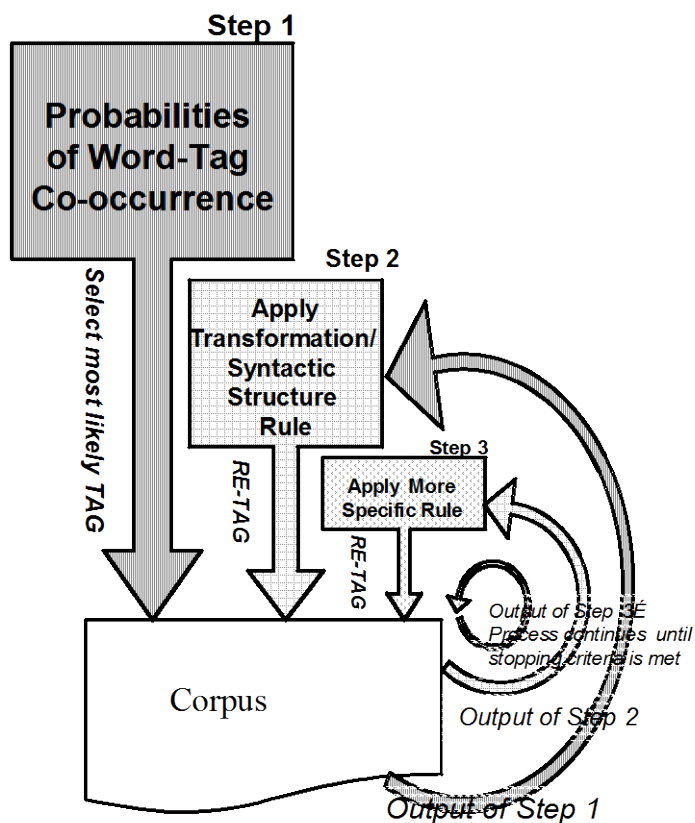
The preceding/ following word is tagged **z** and the word two before/ after is tagged **w**.

However, the idea is that on each round of rule applications, there is the possibility of getting more and more accurate tagging. Jurafsky and Martin (2000:309-310) summarize the iterative nature of Brill’s transformation-based learning (TBL) process in this way:

Brill’s TBL algorithm has three major stages. It first labels every word with its most-likely tag. It then examines every possible transformation, and selects the one that results in the most improved tagging. Finally, it then re-tags the data according to this rule. These three stages are repeated until some stopping criterion is reached, such as insufficient improvement over the previous pass. Note that stage two requires that TBL knows the correct tag of each word; that is, TBL is a supervised learning algorithm. The output of the TBL process is an ordered list of transformations; these then constitute a ‘tagging procedure’ that can be applied to a new corpus.

Graphically, this process could be represented graphically as in the figure below.

Figure 11: Iteration in Brill's (1994) Transformation-Based Learning Tagging Algorithm



As mentioned, the dictionary, the machine-created distribution rules, and the probabilities in a transformation-based tagger all come from a corpus that has already been part-of-speech tagged and determined to be accurate. If the word *race* is taken as an example, exactly how this works can be examined in a different way. If a part-of-speech tagged version of the Brown corpus³⁰ is used to train a certain transformation-based

³⁰ The Brown Corpus contains 1 million words, and was compiled by W.N. Francis and H. Kucera of Brown University incorporating 500 texts of American English written in 1961. These writings come from 15 different genres, such as mystery, romance, journalism, government reports, and more types of fiction and non-fiction. See Kucera and Francis (1964) for more information.

tagger, then when the dictionary is created, it would contain a listing for the word *race* that shows it could be either a noun or a verb.

In fact, in the Brown corpus *race* is used as a noun for 98% of its tokens, and as a verb for only 2%. These percentages will be recorded and used as probabilities, so that when a tagger trained on this corpus is set to tag a new corpus, initially every token of the word *race* will be tagged as a noun, since that is the most likely tag (Jurafsky & Martin 2000:309). Then the distribution rules will be used to change the tags and try to get a more accurate tagging. For example, if in the Brown corpus every time *race* followed *to* it was tagged as a verb, a rule such as the one in the next example might be developed and applied to the new corpus.

Example 25: Jurafsky & Martin (2000)'s Example of a Machine-learned Rule Used by a Transformation-Based Tagging Program

Change NN to VB when the previous tag is TO.

NN = singular count or mass noun

VB = base form of a verb

TO = the word *to*, used as an infinitive marker or preposition

This rule does not specifically mention the word *race*, but it is possible that some rules refer to specific tokens. More general rules referring to categories would be applied first, and later more specific rules that take into account longer sequences of words or else specific tokens. It depends on how specific the programmer chooses to allow the transformation rules to be. It could be specified in the way that the rule templates are written that only categories and not tokens may be referred to by final rules. This would be limiting the set of possible transformations the tagger will come up with, which “is

done by designing a small set of **templates** [emphasis theirs], abstracted transformations. Every allowable transformation is an instantiation of one of the templates (Jurafsky & Martin 2000:309).” Another possibility for limiting or stopping the process is that the programmer could state that the program end after 10 passes of the corpus changing tags.

Summary

Computer programs are now commonly used to do such tasks as tagging parts-of-speech. The three most common ways to do this are by using syntactic distribution rules to label parts-of-speech as in the Constraint-grammar taggers, using probabilities learned from another tagged corpus as in Stochastic taggers, or using a combination of both in the Transformation-based taggers based on Brill’s (1994) algorithm. MontyTagger, the part-of-speech tagger used in this study is a Transformation-based Brill tagger. The main difference between these different automatic methods is whether they use probability and/or distribution rules to estimate the tags in new corpora. Whether or not machine learning is used to extract the rules and probabilities is another difference. Constraint-grammar taggers are the least automated, since not only do these use human-tagged test corpora, but they also use rules written by programmers. Although these tagging systems work faster than humans, they are not as accurate. How accurate they can be, how this is evaluated, and how the tagging in this study was corrected will be described in the next section.

Tag Correction and the Accuracy of Automatic Taggers

Although computational linguists do not strive to limit the number of distribution rules or the size of lookup dictionaries, having the most accurate results that can come from the simplest process is definitely amongst their goals. Liu (2002) reports that the MontyTagger can tag 200 words per second and attain up to 96% word-level accuracy when tagging non-fiction writing in English (in the Python version of the program). This means that 96% of the time, the tagger chose the same tag as that which appeared in the

gold-standard human-tagged and corrected corpus. At first glance, 96% accuracy seems good, but is it really, and how can this be decided? Would a tagger that was 85% accurate also be good? Computational linguists evaluate these possibilities in different ways.

One way to evaluate accuracy other than comparing machine output to “perfectly tagged” corpora, is to compare people to other people, and use this as the upper standard. In fact, when humans tag parts-of-speech in a corpus, they disagree about 3-4% of the time (Marcus et al. 1993, Jurafsky & Martin 2000:308). This may make the reported 96% accuracy for the MontyTagger seem even better. If humans agree about 96% of the time, then 96% accuracy may be the best that is possible. This is typically referred to as the “human ceiling.”

Another way that computational linguists try to evaluate the effectiveness of their programs is by trying to establish a baseline percentage of accuracy to use for comparison. This is done by looking at how accurate the simplest method of automatically determining parts-of-speech is, and comparing it to more complicated methods. For example, if every word were just tagged as a noun, how accurate would the tagging be? If this method was about 50% accurate, then a complicated program with 60% accuracy would not be that much more effective than using this one rule. In fact, Charniak et al. (1993) found that simply choosing the most probable tag for every word (the part-of-speech which a word most often appears as) can lead to 90-91% accuracy (see Jurafsky & Martin 2000:308). So, how good is MontyTagger’s 96% accuracy? MontyTagger is more complicated, but a very simple system can achieve 91% accuracy.

The problem is that these percentages by themselves do not indicate how good a tagger is. First of all, these are the upper limits of accuracy. Secondly, corpora vary in writing style and the complexity of the language. This means that a better way of determining a tagger’s accuracy could be comparing the output from tagging new texts to that from old texts. It is better to evaluate a tagger against itself than against another

tagger developed using a different corpus. Jurafsky & Martin (2000:315) even go so far as to say that “it’s really impossible to compare taggers which are being run on different test sets or different tasks.”

There are many other ways in which the effectiveness of automatic taggers are evaluated. Another commonly used method is Kappa, or the K statistic, which evaluates inter-rater reliability. Contingency rows or confusion matrices are also extensively used. These tables highlight which parts-of-speech are incorrectly tagged, and what incorrect tag is most often chosen instead. There are particular parts-of-speech with which automatic taggers (and people) consistently have decision problems. For automatic taggers, the most issues are found in differentiating:

- Common nouns, proper nouns, and adjectives;
- Verb particles, negative words, and prepositions; and
- Past tense verbs, past participles, and adjectives (Franz 1996, Kupiec 1992, and Ratnaparkhi 1996 as cited in Jurafsky & Martin 2000:313).

This is because the morphemes are similarly spelled but have different meanings, or because what parts-of-speech typically surround them are similar.

Given that mistakes in tagging are expected, the output of the MontyTagger had to be corrected to a certain extent after the automatic tagging was completed. The next section describes what was done.

Adapting and Correcting MontyLingua-2.1 Output for This

Study

Practically, this is how MontyTagger works:

- As the part-of-speech of a word is identified, a slash (/) is appended to the individual word followed by the code for the presumed part-of-speech (for example, VBD, NN, or JJ).

- MontyLingua-2.1 parses sentence chunks into noun, verb, and adjective phrases. Elements included in these phrases are contained within parentheses and marked with NX, VX, or AX, respectively.

In this study, the tags marking parts-of-speech and phrase boundaries were altered, corrected, or deleted for these three reasons:

- Verb phrase and adverb tags were not relevant to the main focus of the investigation. Removing them made it easier to focus on the analysis of the noun phrases and parts-of-speech within noun phrases.
- Syntax and spelling errors in (especially learner) production caused an increase in the number of errors in tagging and parsing phrases. These mistakes needed to be corrected prior to analysis. As discussed, some manual correction is typical after automatic tagging, depending on the intended use of the output.
- MontyLingua does not include all complements and adjuncts of a noun phrase within its tagging of a noun phrase. In order to tag co-reference, encapsulating the entire noun phrase was necessary because the complements and adjuncts alter the semantic interpretation of the referent of the noun phrase. The purpose of the phrasal parsing in this project is not the same as that for which the MontyLingua parser was developed.

Tag Removal

Those tags that were not relevant to noun phrases and the analysis of this study were automatically removed with a Perl program³¹. This program is given in Appendix D. The specific tags removed were verb phrase and adjective phrase markers, verbal tags, and adverb tags, along with tags marking coordinating and subordinating

³¹ Perl is another uninterpreted computer language. The program used to remove the tags has been included in an appendix.

conjunctions and prepositions. Below is a sentence from the text of Student 1's essay, with all the part of speech tags included, and explanations of what the tags mean below the example.

Example 26: Sample of Part-of-Speech Tagging, Student 1

(NX First/NNP ./, Friendship/NNP NX) (VX is/VBZ VX) (NX an/DT good/JJ example/NN NX) of/IN (NX culture/NN borrowing/NN NX) ./.

(NX NX) = noun phrase

NNP = singular proper noun

(VX VX) = verb phrase

VBZ = present tense verb, 3rd person singular

DT = article or determiner

JJ = adjective or ordinal number

NN = singular or mass common noun

IN = preposition or subordinating conjunction

This same passage after the tags that were not necessary for the analysis had been removed, looks like the next example.

Example 27: The Same Sample with Tags Removed

(NX First/NNP ./, Friendship/NNP NX) is (NX an/DT good/JJ example/NN NX) of (NX culture/NN borrowing/NN NX) ./.

As can be seen, removing the tags extraneous to this investigation highlights the constructions relevant to the purpose of the study. The three NX phrases are easier to identify in the second example above. This also highlights an error in the tagging that

needs to be corrected. The first two words have been tagged as proper nouns in the same noun phrase instead of as a connecting word and a noun. So, corrections to the tagging still obviously need to be made.

Tag Correction

Other tags in this sample were deleted, moved, or changed because the effectiveness of the automatic tagger was limited by mistakes in spelling and grammar, and the fact that it was not developed to be used on the exact kind of corpus used in this study. One example of the way in which student errors affect tagging relates to the problematic tagging of “First, Friendship” in the previous example. Student 1 uses a capital letter for “Friendship,” which would be interpreted as an error in capitalization by most native English writers since it is the second word in the sentence. The part-of-speech algorithm has construed this capital letter as indicating that *friendship* is a proper noun (NNP) and not a singular or mass common noun (NN). This has further lead to *first* and *friendship* being placed within a single noun phrase, when *first* here is functioning as a sentence connector, and is not modifying *friendship*. After correcting these related errors, the tagged sample appears as follows in the example on the next page.

Example 28: The Sample from Student 1 Again after Tag Correction

First ./, (NX Friendship/NN NX) is (NX an/DT good/JJ
example/NN NX) of (NX culture/NN borrowing/NN NX) ./.

Friendship is left capitalized in order to retain the writing of the learner as it was produced, but the noun phrase tagging and parsing has been altered.

In the next example, which is from Student 2, a similar mistake is shown. The gerund *wearing* was tagged as a verb. However, its position following the genitive noun *people's* seems to indicate that the gerund should be read as a noun or noun-like. It is even possible that the learner intended to use an entirely different lexical item (such as *clothing*). It is impossible to know what the learner truly intended, but based on the syntactic position of the gerund, this will be interpreted as a tagging error which also must be altered. The fully tagged and corrected sample are given below.

Example 29: A Fully Tagged Sample from Student 2

(NX People/NNS NX) 's/ POS (VX wearing/VBG is/VBZ almost/RB VX) (AX same/JJ AX) to/TO (NX western/JJ culture/NN NX) not/RB only/RB on/IN outside/IN ./.

(NX NX) = noun phrase

NNS = plural common noun

POS = possessive ending

(VX VX) = verb phrase

VBG = gerund or present participle

VBZ = present tense verb, 3rd person singular

RB = adverb or negation

(AX AX) = adjective phrase

JJ = adjective or ordinal number

TO = to—preposition or infinitival marker

NN = singular or mass common noun

IN = preposition or subordinating conjunction

Example 30: The Same Sample with Corrected Tags

(NX People/NNS 's/ POS wearing NX) is almost same/JJ to (NX western/JJ culture/NN NX) not only on outside ./.

In the corrected sample, *wearing* is left without a tag because as it deviates from typical native writer production and as mentioned there is no access to the learner to clarify their intended meaning. It is not clear what tag would be appropriate if another one were substituted. Similar corrections had to be made occasionally in the native corpus due to errors in spelling or punctuation.

Phrase Correction

As mentioned, aside from removing irrelevant tags and correcting tagging mistakes, phrase boundaries were also altered so that they would include all the adjuncts and complements of the noun phrase. This was also done to make the phrase boundaries consistent with the syntactic restrictions on noun phrases outlined in the rules for coding co-reference. Looking at the sample of tagged text from Student 1 again, it can be seen that the second noun phrase *an good example* has not been included in the same phrase as its prepositional modifier *of culture borrowing*.

Example 31: The Sample from Student 1 after Tag Correction

First ./, (NX Friendship/NN NX) is (NX an/DT good/JJ
example/NN NX) of (NX culture/NN borrowing/NN NX) ./.

Since the PP modifies the referent and specifies what kind of *example*, it is relevant for calculating the semantics of the NP referent. Once the phrase boundaries have been corrected, the tagging of this sample looks like:

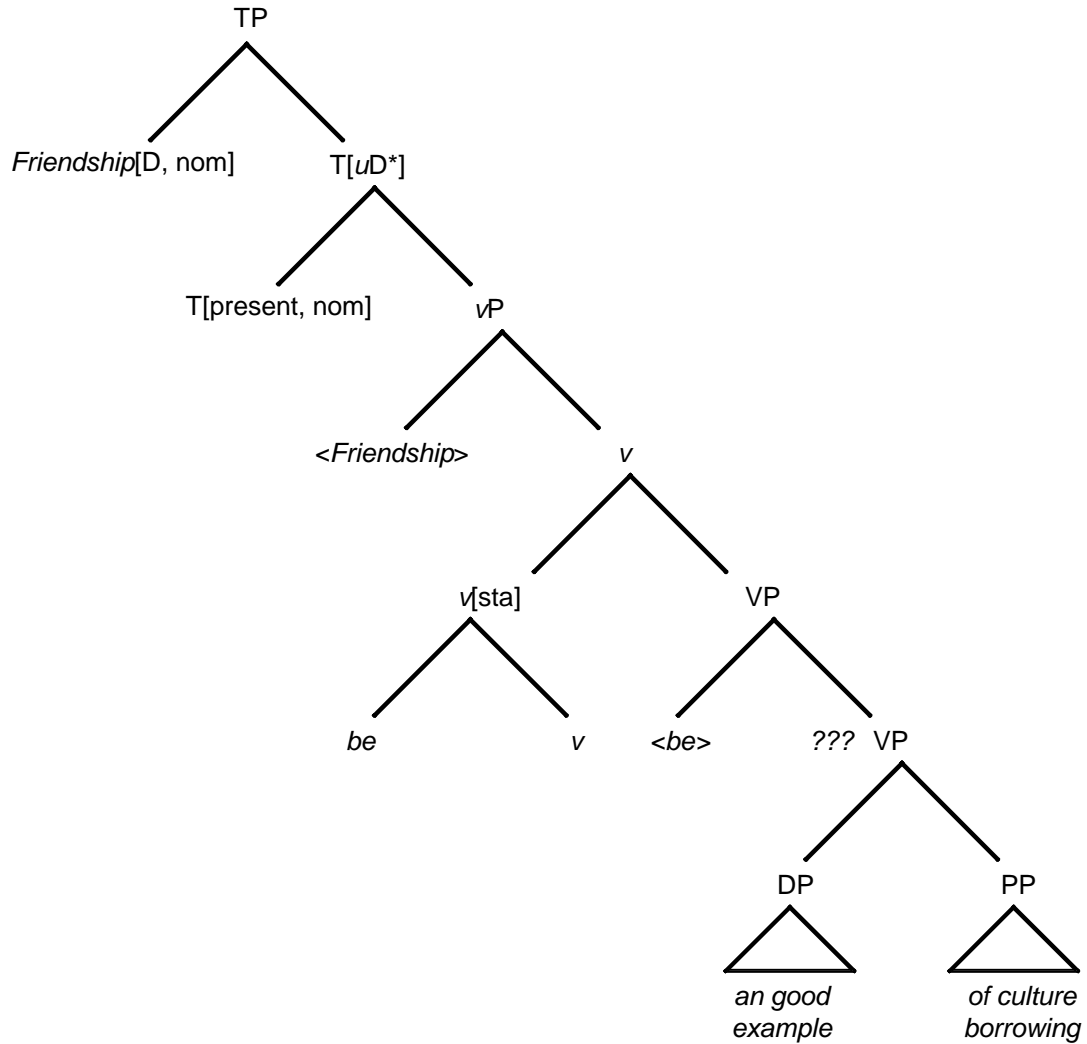
Example 32: The Same Sample from Student 1 after Phrase Boundary Correction

First ./, (NX Friendship/NN NX) is (NX an/DT good/JJ
example/NN of (NX culture/NN borrowing/NN NX) NX) ./.

With this alteration of the phrase boundaries, the PP containing the NP *culture borrowing* is now embedded in the larger NP *an good example of culture borrowing*. This is meant to indicate that *culture borrowing* can co-refer with another instance of an NP that refers only to *culture borrowing*, but *an good example* can only now be marked as co-referring with other *good examples of culture borrowing*. The NP head cannot be tagged as co-referring with other “good examples,” such as perhaps “a good example of native culture.” Basically, the tree structure is being changed from that of the first figure that follows on the next page, to that of the second. Instead of being connected with the verb phrase (which leads to an odd meaning in this case), the PP is connected to the NP.

In other excerpts, it is not as clear how and where to link the components. For example, this NP from one of the native writers has two possible meanings and two possible structures: “the elimination of general education requirements in today’s universities.” The meanings are very similar, but not exactly the same. The first meaning links the PP “in today’s universities” to the *elimination*, with the meaning that the eliminations are happening right now.

Figure 12: Tree Structure Representation of the Initial Parsing of the Last Example



The second meaning links “in today’s universities” to “general education requirements.” This meaning focuses the reader’s attention on the fact that the general education requirements are those currently in use. The “now” focus is on the requirements and not on the elimination. The structures for these two meanings appear

after the alternate structure for the last example. In cases such as this where two meanings and two structures were possible, the longer NP was chosen unless there was a clear indication in the context that the PP should be attached to the verb phrase.

Figure 13: Tree Structure Representation of the Changed Parsing of the Last Example

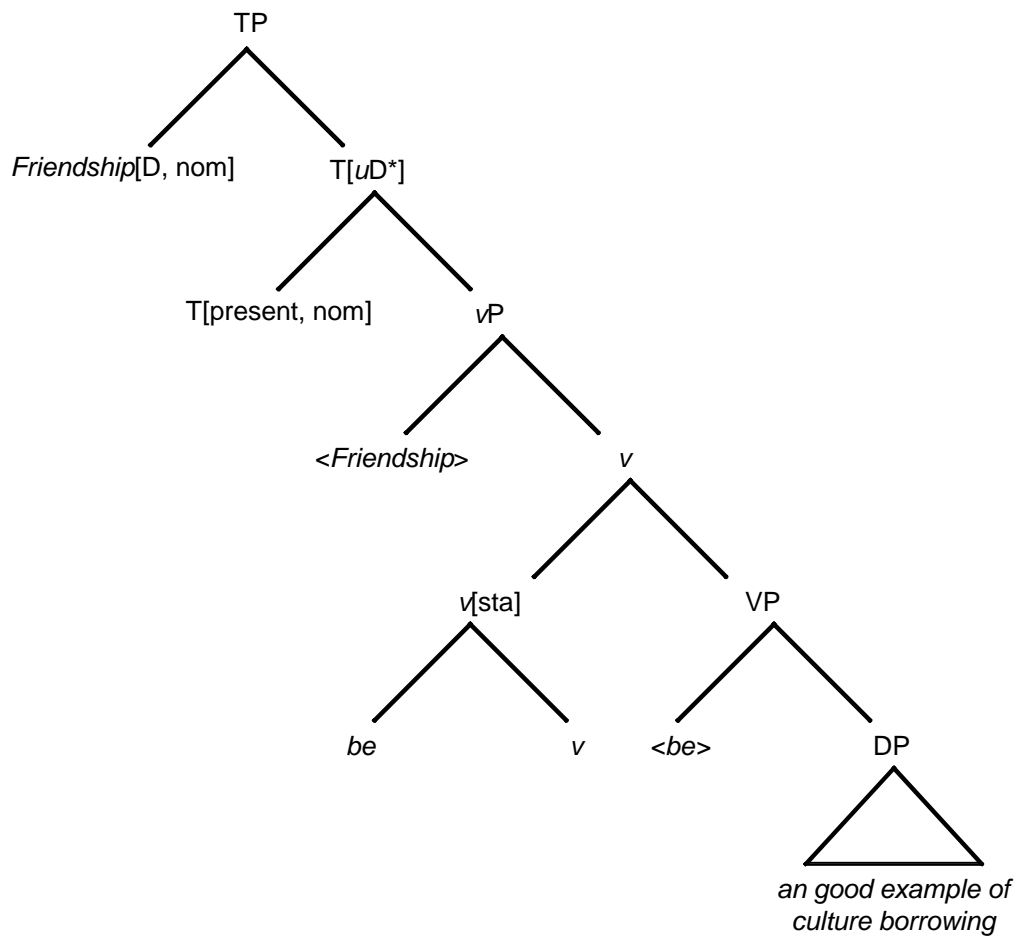
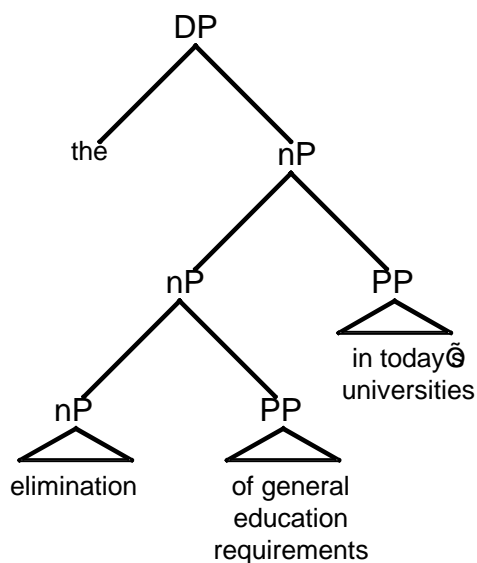
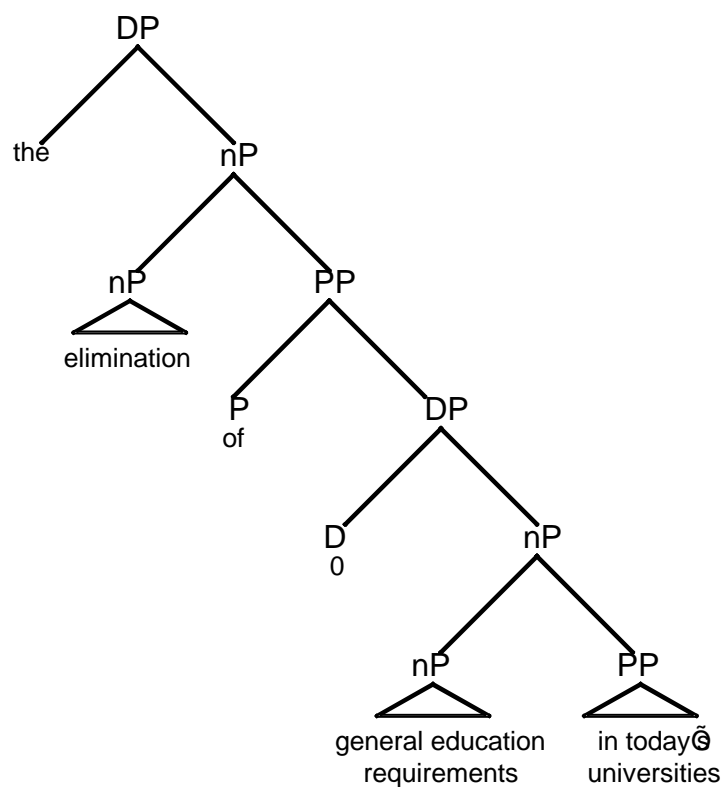


Figure 14: Tree Structure for the Meaning Highlighting the Timing of the *Elimination* in the Phrase “the elimination of general education requirements in today’s universities”



Under the co-reference tagging rules that will be described next, the head noun of a noun phrase cannot participate in co-reference as a separate entity. In contrast, genitive nouns and possessive pronouns are considered to co-refer to real-world elements separately from the rest of the noun phrase being considered. The MontyTagger includes these in larger NPs, but does not tag genitives as part of their own noun phrases. For this reason, the tagging of genitive noun phrases is another common type of phrase boundary alteration that was made in the essays. The example that illustrates this also shows how odd or non-native word choices by the writer (using *same to* instead of *same as* and omitting *the* or using a null determiner) can also lead to inaccurate phrase boundaries in the tagging. Therefore, both the intended and unintended behavior of the MontyTagger was at times altered. See the uncorrected and corrected samples from Student 2 following the syntax tree on the next page.

Figure 15: Tree Structure for the Meaning Highlighting the *Today-ness* of the General Education Requirements in the Phrase “the elimination of general education requirements in today’s universities”



Example 33: The Uncorrected Sample from Student 2

(NX People/NNS 's/ POS wearing NX) is almost same/JJ to (NX western/JJ culture/NN NX) not only on outside ./.

In this example above, “people’s” is not tagged as its own NP, and “same to” is not linked with “western culture.” The has been corrected in the example that follows. From

only two tagged NPs in the example above, there are five in the final tagging. This aligns the phrase parsing with that needed for the co-reference tagging. The system of co-reference tagging will be described in the next section.

Example 34: The Same Sample from Student 2 with Corrected Phrase Boundaries

(NX₁ (NX₂ People/NNS 's/ POS NX₂) wearing NX₁) is almost
(NX₃ same/JJ to (NX₄ western/JJ culture/NN NX₄) NX₃) not only
on (NX₅ outside NX₅) ./.

In the sections that follow, a full explanation of how co-reference was tagged and how the appropriate syntactic boundaries for noun phrases will be given.

Co-reference Tagging

Once the part-of-speech tagging was corrected, co-reference between the NPs was marked. The MUC-7 guidelines for tagging co-reference (Hirschman 1997) were used as the basis of this tagging scheme. The full text of the coding rules with examples appear in Appendix E. In this section, first some general considerations about co-reference tagging will be discussed, and then some of the trickier aspects of the coding scheme will be examined.

Characteristics Shared by Commonly Used Co-reference

Tagging Schemes

Massimo Poesio has a description of different co-reference tagging schemes (see Poesio 1999a) in which he explains:

When designing a scheme for annotating anaphoric relations it is then necessary to identify the anaphoric expressions and relations more relevant for one's needs. Narrowing the scope of the scheme may also be necessary in order to achieve good agreement among subjects. This can be done by specifying syntactic constraints on markables, which are the text spans that enter into coreference

relationships, by specifying constraints on the sorts of objects in the world for which coreference will be marked up, or by restricting the kinds of coreferential relations which will be considered (for instance, by deliberately failing to mark bridging references³²).

This is a helpful way of outlining the major features of any co-reference tagging scheme:

- Each scheme will have syntactic constraints on markables,
- Specifications regarding what kinds of real-world entities are important to tag, and
- Restrictions on what anaphoric relationships are important.

For example, when determining the syntactic structures which will be considered, it is possible to limit the tagging to only noun phrases, while excluding entire sentences—although sentences may indeed participate in anaphoric relationships. Even if only noun phrases are tagged, though, there are still certain problematic syntactic structures and lexical items for which decisions about tagging co-reference must be made. Some of these more difficult to handle structures include appositions, first and second person pronouns, noun phrases that follow the copula, empty elements, and nominal clitics on verbs (which is not so relevant for English, but would have to be dealt with when tagging Spanish, Bulgarian, and many other languages).

Unlike syntactic structure constraints, which almost every co-reference tagging scheme will specify, very few schemes limit the kinds of real-world objects that may be tagged. Typically if such limitations are made, it is in response to the goal of the research project or the set-up of the task given to those who contributed their speech or writing to the corpus being examined. For example, a project based on using Anderson et al.'s (1991) MapTask corpus (collected at the University of Edinburgh) may choose to limit

³² A bridging relationship is an anaphoric or coreferential relationship in which the two syntactic elements do not refer to exactly the same semantic element. An example of this would be a situation in which one noun phrase refers to a set, such as *the couple*, and another noun phrase in an anaphoric relationship with it refers to only part of the set, such as *the wife*.

the tagging of co-reference to only the landmarks on the map writers had to use to give or receive directions. Similarly, research using the TRAINS task-oriented corpus (Gross et al. 1993) may limit a tagging scheme to marking only the phrases that refer to the towns or box cars and materials which participants were asked to manipulate. This narrowing of the scope of the tagging task makes it proceed more quickly and helps to focus the data on the questions central to the investigation.

Some limitations on tagging schemes can also be made in order to increase inter-rater reliability. Distinguishing between different kinds of bridging relationships can be tricky for annotators, so some co-reference tagging schemes opt to only mark identity, or referents that refer to exactly the same real-world element. Bridging inferences such as set-subset, function-value, whole-part, item-attribute, event-cause, class-example, or item-possessor are more complicated because there can be very fine distinctions between some of the types. Furthermore, there is no agreed-upon, complete list of bridging relationships, so it is unclear how detailed the distinctions between categories should be, or how many bridging relationships is sufficient to capture useful generalizations. Therefore, because marking bridging relationships can make a co-reference tagging system too intricate to manage, these are often not included in tagging schemes unless they are central to the purpose of the investigation.

After describing these three basic characteristics of co-reference tagging schemes, Poesio goes on to contrast the specific details of several existing co-reference tagging systems, including MUC-7 (Hirschman 1997), DRAMA (Passoneau 1996), Lancaster University's UCREL (Fligelstone 1992), Bruneseaux and Romary (1997), and the MapTask (Anderson et al. 1991) already mentioned. Of these existing schemes, the MUC-7 scheme was used as the basis for the co-reference tagging scheme in this project because it was originally designed to mark text, is based on the syntax of English, does not greatly limit the range of referent types to be marked, and also tags only one central

anaphoric relationship—identity. More about this scheme and its specific details are explained in the next section.

The MUC-7 Co-reference Tagging Scheme

MUC-7 is the abbreviated name of the seventh (and final) Message Understanding Conference that was held in 1997. The purpose of the conference was to compare and evaluate automated computer programs' ability to accurately and fully extract information about named entities in newswire texts (for example, people, places, corporations, and dates). In order to summarize and catalog related information, tagging co-reference became a necessary task for that project (see Chinchor's (1998) "Overview of MUC-7/MET-2" for more information), which led to the development of Hirschman's (1997) co-reference tagging scheme. This scheme was purposely developed to support information extraction, reach high inter-rater reliability (95%), and allow for fast tagging so that a relatively large quantity of text could be tagged.

The syntactic constraints that are placed on text spans to be tagged in MUC-7 are enumerated below. To summarize the main features of this scheme, though, one could state that it only focuses on noun phrases (clauses are not marked), and that all complements, adjuncts, and modifiers of the head noun are included in the text span. Furthermore, MUC-7 does not place restrictions on the kind of semantic elements that may be tagged, and only marks the anaphoric relationship of identity. When two referents indicate the same object in the real world, they are tagged as co-referring. Identity is seen to be symmetrical and transitive, and is the most simple relationship.

Minor alterations to the MUC-7 scheme have been made in order to better support the goals of this project, and also to slightly increase the reliability of the tagging. Some examples of simplifications related to the task of the project are that non-referential pronouns (Existential-*it* and *there*) were not tagged, and neither were certain set expressions such as "a lot of." Expressions like "a lot of" have articles and determiners

and the form of noun phrases, but there is no real reference for “lot,” and such phrases are not productively formed. Learners memorize them as a group, so these set expressions do not truly indicate how they use noun phrase structure and determiners in order to indicate chains of reference. Also excluded from the tagging in this project are unmodified or simple dates, numbers and currency expressions. This simplification of the system could potentially increase the reliability of the coding system since dates, numbers, and currency expressions often are parts of function-like chains of reference in which the referent may be seen to change over time. These types of situations are difficult to tag for co-reference.

Attempting to increase the reliability of the tagging is important for this project because unlike in the MUC-7 task, half of the writing to be examined was written by non-native English writers. Variations from the norm in terms of word choice and grammatical structure decrease the reliability in the tagging, so seeking a gain in reliability by simplifying the coding system is desirable. Furthermore, these types of numeric or date expressions were important for Hirschman’s scheme because the focus of MUC-7 was on retrieving information from newspaper articles to create summaries, but they are not important for this investigation and have therefore been excluded as extraneous. As in the MUC-7 scheme, only the identity relationship has been linked between noun phrases, and bridging inferences are not classified. Identity is the most easily classified anaphoric relationship.

In the discussion below, some aspects of the syntactic constraints and the mechanics of marking co-reference are outlined. The majority of these constraints and considerations come directly from Hirschman’s (1997) work (taking into account the variations simplifications already mentioned). As will be seen, change over time, conjunction, and ambiguity of the text are major difficulties for both identifying the syntactic units to tag and marking identity of reference. In the examples that follow, subscripts have been added to clarify the perceived noun phrase embeddings, while

identification (ID) numbers indicate which entities co-refer through identity. These examples come either directly from the student or native essays, or Hirschman's examples, as indicated.

Tag the Largest Phrasal Unit Possible

As mentioned, one of the most important aspects of delineating the syntactic units in this scheme is that the entire noun phrase should be tagged, including determiners and adjective phrases, along with the adjunct or complementary relative clauses or prepositional phrases and any individual nouns embedded within those structures. The largest phrase possible is considered the main syntactic unit for co-reference, and then the NPs in the adjuncts and complements are like center embeddings, so that the tree structure of the syntactic parsing is in part maintained. However, in this scheme there is really no difference made between how an adjunct versus a complement is represented, and they are treated as equally important in the semantic calculation of the meaning and reference.

Embedded NPs and Co-reference in PPs and Relative Clauses

An example of the coding of a prepositional phrase is given in the excerpt of the student essay that follows. The tricky part about tagging the syntactic boundaries and co-reference of the PPs is that they have phrase-internal NPs. These may participate in co-reference separately from the phrase as a whole. On the other hand, the head noun cannot participate in co-reference without the altered meaning calculation of the PP under this tagging scheme. The phrase from Student 1 that appears on the next page exemplifies this. NX_1 is the full syntactic unit with *relationships* as the head noun. Within the PP, the tagging system identifies both the possessive pronoun *their* (NX_3) and the NP *their friends* (NX_2) as text spans that can individually participate in co-reference relationships

with other noun phrases within the same sentence, or else within the discourse of the writing as a whole.

Example 35: NP + PP with Possessive Pronoun from Student 1

“the relationships of their friends”

DET + HEAD N + PP

(NX₁ the relationships of (NX₂ (NX₃ their NX₃) friends NX₂) NX₁)

In regards to the head noun, the phrase boundaries are meant to indicate that NX₁ can only co-refer with other syntactic text spans that also refer to *relationships of their friends*, and NX₁ would not be tagged as co-referring with a noun phrase referring to other kinds of relationships.

Relative clauses are also included within the largest syntactic unit that is considered for co-reference tagging, whether they are complements or adjuncts. Noun phrases containing adjective clauses are difficult to tag in both the native writer and the learner essays. This is because the native writers tend to use very long, extended noun phrases when they use relative clauses, so it is difficult to tag the embeddings. In contrast, the learner noun phrases that contain relative clauses are often constructed in ways that are ambiguous in meaning or which could be considered ungrammatical. In this case, the reader must again attempt to infer what the writer intended based on the context, which may decrease inter-rater reliability.

This next example from a native writer exemplifies an extended noun phrase in that it contains three clauses—a non-finite clause, relative clause, and noun clause. In this particular case, none of the noun phrases in these clauses co-refer with each other, but if they did, they would be marked as such.

Example 36: An Extended Description by Native 28

“the temptation to endorse producing graduates who have not been exposed (forcibly, if necessary) to what is called a liberal education”

(NX₁ the temptation to endorse producing (NX₂ graduates who have not been exposed [forcibly, if necessary] to what is called (NX₃ a liberal education NX₃) NX₂) NX₁)

Notice also that the *wh*-words have not been tagged as noun phrases. This means that they cannot be marked as participating in co-reference. This is consistent with Hirschman (1997). She gives the examples “Which engine would you like to use?” and “Who is your boss?” In these questions, neither *which engine* nor *who* are considered markable. The syntax of relative clauses and full questions is clearly not the same, but the semantics of the question words is comparable in that if one chose to mark co-reference, it would be difficult to create consistent rules to use to determine what the *wh*-operator refers to.

Learner relative clauses are tagged in a similar way, although more inference is required in determining the structure and meaning since the majority of the relative clauses are not constructed in fully native-like ways. Those clauses which can be reasonably assumed to be relative clause-like are still included within the full syntactic unit of the noun phrase that may be tagged for co-reference. In the example below, the learner is introducing the topic of actions of his father which he might or might not want to emulate when he is a father. The description *what everyone does* is relative clause-like because (1) it follows and further specifies the conjoined noun heads “very good things and bad things,” and (2) has a complementizer, subject, and finite verb. Two of the ways in which the questionable grammar of this phrase could be improved would be if (a) the described conjoined noun phrase were preceded by *the* and the complementizer *that* was used instead of *what*, or (b) if the word *like* were inserted before *what everyone does*.

There are multiple ways to alter this phrase to improve its grammaticality, but as these two possibilities show, the altered, grammatical version may typically end up as some form of a relative clause. Grice's Maxims of Relevance and Quality lead to the hypothesis that readers will try to match an ungrammatical structure to a grammatical possibility with minor changes when trying to understand a particular phrase. This example, then, may be read as a relative clause. So, it will be analyzed as such for the tagging of co-reference and the syntactic units open to co-reference tagging. This same assumption is made with reading other problematic sentences written by learners. Given this hypothesis, there are five different syntactic units that could be tagged for co-reference in this example. See below.

Example 37: NP + Malformed Relative Clause from Student 12

“My father did very good things and bad things what everyone does.”

ADJ + NP + Rel Cl

(NX₁ very (NX₂ (NX₃ good things NX₃) and (NX₄ bad things NX₄) NX₂) what (NX₅ everyone NX₅) does NX₁)³³

Note again that the complementizer *what* is not tagged separately, since wh-words in general are not considered in regards to their co-reference.

³³ The parsing of this phrase is ambiguous. *Very* could modify solely *good things*, or it could modify *good things* and *bad things*. In addition, the relative clause *what everyone does* could modify just *bad things*, or else the whole conjunct *very good things and bad things*. In such situations, the assumption was that the relative clause or the adjective modified as much as possible. In other words, the parsing algorithm was “greedy” in the way that Perl regular expression matching is.

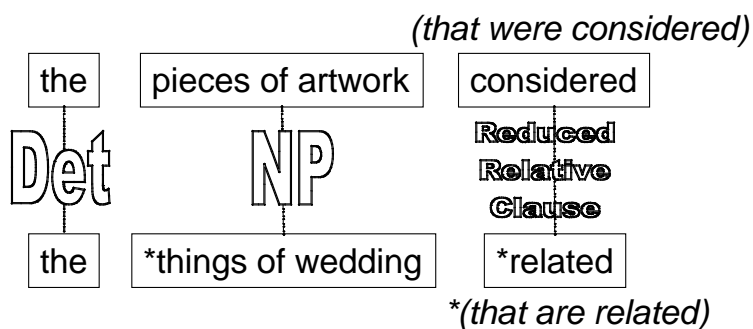
A similar process of inferring the structure or meaning can be seen in the next example, which is a more difficult case than the above. The meaning of the phrase below is something akin to *wedding-related things* or *the things that are related to the wedding*.

Example 38: Malformed Structure from Student 19

“the things of wedding related such as wedding rings, formal suits, a pair of watch, and so on”

Although the phrase is actually structured most similarly to an NP + descriptive PP-type description, the structure is not unrelated to that of a reduced relative clause. See the figure below.

Figure 16: Comparison of Grammatical Reduced Relative Clause Structure and Learner Phrase Structure



In this instance, interpreting the meaning as *the things that are related to the wedding* and interpreting the structure as that of an NP + PP followed by a reduced

relative clause, there are two main syntactic units that can be considered for co-reference, after excluding the apposition: (a) *the things of wedding related*, and (b) *wedding*. See the parsing given in the following example.

Example 39: Malformed Structure from Learner 19

“the things of wedding related”

DET + (NP + PP) + Reduced Relative Clause + Apposition

(NX₁ the things of (NX₂ wedding NX₂) related NX₁)

Although in many cases the grammatical structure and calculated meaning of students' phrases was assumed to be comparable to certain native or fully grammatical structures, in some cases it was impossible to do so. In those instances, the NPs and their related structures were simply excluded from consideration.

Conjoined NPs

Conjoined noun phrases have appeared in many of the examples given above, but their tagging has not been specifically mentioned. For the tagging scheme used here, the entire conjunction is tagged as one syntactic unit open to co-reference marking, but each NP making up the conjoined unit is also tagged individually. This is slightly different than Hirschman's (1997) MUC-7 scheme. Hirschman recommends tagging either the entire conjoined unit or the individual NPs only based on which units are referred to elsewhere by other NPs. This process was modified for this study so that the whole and the parts are always tagged. This change makes the tagging faster and more consistent because neither the coder's memory (of how items were already referred to) nor the coder's guessing skills (for how the writer will refer to the phrases in the subsequent text) must be relied upon when tagging the phrase. This alteration changes the number of

noun phrases that will be identified and considered, but it does not affect the ultimate number of co-reference chains that would be identified. The example of a native writer's noun phrase which is repeated below shows the tagging of the conjoined NP "art, geology, and philosophy."

Example 40: An Extended Description from Native 23

"Other subjects which I have studied for general education requirements, such as art, geology, and philosophy"

(NX₁ (NX₂ art NX₂), (NX₃ geology NX₃), and (NX₄ philosophy NX₄) NX₁)

This tagging allows the co-reference of either the individual subjects, or the combination of all the subjects to be identified as related.

One tricky concern when tagging conjoined NPs in the learner essays relates to the use of determiners. Some conjoined NPs will have a determiner only before the first conjunct. When reading native writer writing, the assumption is that the determiner can be interpreted as co-occurring with each conjoined element. For example, in the excerpt below, assuming that the author intended the audience to read the phrase as *the Constitution and THE Bill of Rights* is fairly safe. Even though *the* does not appear before *Bill of Rights* in the actual essay, the phrase would be ungrammatical without it, and the conjunction allows this reading.

Example 41: Conjoined NP with Determiner by Native 26

"the Constitution and Bill of Rights"

(NX₁ the (NX₂ Constitution NX₂) and (NX₃ Bill of Rights NX₃) NX₁)

In other words, it is acceptable to assume that an adult, educated native writer is writing in formal, standard English and in full prose style (not just in notes or fragments) on an essay exam.

It is difficult to come to the same conclusion when considering learner examples, especially when (a) the learner *does* omit or use zero forms of determiners in positions in which a native writer would not, or (b) the learner repeats the determiner in front of only selected conjuncts, such as in the excerpt from Learner 001 below.

Example 42: Three Conjoined NPs from Student 1

“their hobby, foods, and their lifestyle”

The question here is whether this phrase should be interpreted as: (1) *their hobby, THEIR foods, and their lifestyle* or (2) *their hobby, ∅ foods, and their lifestyle*. Without being able to question the author, it is impossible to distinguish which interpretation was intended. Therefore, the choice has been made to assume that the learner intended the more native-like grammatical structure. This means that option (1) above *their hobby, THEIR foods, and their lifestyle* is assumed to be the structure the author desired. The tagging would then be as shown in the next example.

The lack of *their* before *foods* is not considered to be use of a null determiner or bare noun, but no empty element is marked as co-referring with other instantiations of *their* as it would be if *their* were overtly stated, as it is before *lifestyle*.

Example 43: Three Conjoined NPs from Student 1

“their hobby, foods, and their lifestyle”

NP, NP, and NP

(NX₁ (NX₃ their NX₃)/ID001 (NX₂ hobby NX₂) (NX₄ foods NX₄)
and (NX₅ (NX₆ their NX₆)/ID001 lifestyle NX₅) NX₁)

With this interpretation, the tagging errs on the side of assuming that the learners are using more native-like grammatical structures. The general pattern of not tagging null elements is also maintained.

Appositions

Appositions are tagged in a manner similar to relative clauses. The apposition is tagged as a syntactic unit that can be marked for co-reference, and is also marked as contained within the larger NP structure. In most cases, the apposition is in addition identified as co-referring with the entire NP.

In the example below, the apposition “my country” is shown as being a separate NP (NX₂), and also is tagged as co-referring with the whole syntactic unit “Korea, my country.”

Example 44: An Apposition from Student 2

“Korea, my country, is not an exception.”

(NX₁ Korea (NX₂ (NX₃ my NX₃)/ID001 country NX₂)/ID002
NX₁/ID002

In this example, it is also the case that the possessive pronoun *my* is a separate noun phrase, and would co-refer with uses of the pronoun *I* in other parts of the essay.

Meaning Connections that Cannot be Tagged

In the last example, one of the less-intuitive parts of the tagging scheme is illustrated. Instead of the apposition *my country* being marked as co-referring with just the head *Korea*, it is tagged as sharing identity of reference with the entire phrase that also contains itself: “Korea, my country.” Hirschman (1997) explains that noun phrase heads cannot be marked as noun phrases separate from the entire syntactic unit, nor marked as co-referential independently. The next example (from Hirschman) demonstrates this.

Example 45: Noun Phrase Heads Cannot be Tagged Independently (Hirschman 1997)

“okay then I’ll take engine E two”

(NX₁ engine E two NX₁)/ID001

*(NX₁ engine NX₁)/ID001 (NX₂ E two NX₂)/ID001

The NP *engine E two* cannot be split into its component common noun type and its name or identifying number (*engine* and *E two*, respectively). It must be treated as one unit, so the two parts cannot be marked as co-referring. This specification leads to many cases in which part of an NP syntactic unit is tagged as co-referring with the entire piece that contains it.

The fact that the head cannot be split from the entire compound phrase in regards to co-reference does not mean that the head is re-defined as the entire phrase. The head of the NP remains the syntactic head. Therefore, in the examples on the next page, the heads are *taxes* and *newspaper*.

Just as NP heads cannot be split from the entire phrase and marked as co-referring with other elements independently, long names cannot be split up into separate parts.

Example 46: Heads of Compound Nouns and Collocations

(a) “income taxes” (Hirschman 1997)

(NX₁ income taxes NX₁), *head = taxes*

(b) “the college newspaper” (Learner 015)

(NX₁ the college newspaper NX₁), *head = newspaper*

In the example from Hirschman (1997) below, the two uses of the word *Iowa* cannot be marked as co-referring. This is because in the first case, *Iowa* is embedded in the name of the insurance company. It is part of the name, and although the lexical item is the same and refers to the same real-world object, it is a puzzle piece that cannot be considered separately.

Example 47: Names Cannot be Split when Tagging Co-reference

“Equitable of Iowa Cos. ... located in Iowa.” (Hirschman 1997)

(NX₁ Equitable of Iowa Cos. NX₁) ... located in (NX₂ Iowa NX₂)

*(NX₁ Equitable of (NX₃ Iowa NX₃)/ID001 Cos. NX₁) ... located in (NX₂ Iowa NX₂)/ID001

There is in fact no way under this tagging system to mark the relationship between these two uses of *Iowa*. This is similar to the situation with bridging references, or co-reference relationships other than identical reference. The relationship between the elements is simply not taken into consideration.

Modifying nouns in compound nouns are similar, but handled slightly differently than names and head nouns. Nouns that are used as modifiers before other nouns can be tagged as separate syntactic units participating in co-reference on their own if the entire

NP is not a compound noun or a name. Hirschman (1997) determines this based on whether or not the modifying noun is used separately in another noun phrase. If it always co-occurs with the head noun, then it is not tagged as a separate unit. This is somewhat difficult to tag because the tagger must look at the text before and after the phrase in question to see if it is used separately or refers to the head of another noun phrase. Hirschman sets up the tagging this way because if the modifying noun co-refers with a noun head or phrase elsewhere, then the phrase as a whole is not really working like a compound noun, but rather as two separate entities. In the example that follows, *aluminum* is initially a pre-nominal modifier of *siding*. However, in the subordinate phrase in the sentence, it is referred to separately from *siding* and just as the material alone.

Example 48: A Pre-Nominal Modifying NP that Refers Separately, from Hirschman (1997)

“The price of **aluminum siding** has steadily increased, as the market for **aluminum** reacts to the strike in Chile.”

(NX₁ The price of (NX₂ (NX₃ **aluminum** NX₃)/ID001 **siding** NX₂)/ID002 NX₁)/ID003 has steadily increased, as (NX₄ the market for (NX₅ **aluminum** NX₅)/ID001 NX₄)/ID004 reacts to the strike in Chile.

The two instances of *aluminum* are thus tagged as co-referring, and the overall sense of the sentence is that it is important that the siding is made of aluminum—it is not just part of a name, but the substance is key to the meaning of the sentence as well.

The example below from an English Language Learner does not have the same relationship between the two complete noun phrases of raw material vs. manufactured product as the sentence from Hirschman does, but it is still a case in which the modifying

noun in one NP appears as a head noun in another NP. These two instances can then be marked as co-referring, and considered as part of how the essay writer may have been trying to build a related discourse.

Example 49: A Modifying Noun Used Elsewhere as a Head Noun by Student 15

“In conclusion, the challenging spirit and real activity are pivotal to **the life**. From the event I got the active character and brave. **My life decree** is that ‘I just do it, not watch it.’”

In conclusion, the challenging spirit and real activity are pivotal to (NX₁ **the life** NX₁)/ID001. From the event (NX₂ I NX₂)/ID002 got the active character and brave. (NX₃ (NX₄ **My** NX₄)/ID002 (NX₅ (NX₆ **life** NX₆)/ID001 **decree** NX₅)/ID003 NX₃/ID004 is that ‘I just do it, not watch it.’”

In *life decree*, the noun *life* modifies the noun *decree*. This *life* though, is co-referential with the generic use of *the life* earlier on.

The learner’s example, aside from not displaying the same semantic relationship between the modifying and the head noun that co-refer, is also different from Hirschman’s example in two other ways: (1) Hirschman’s example is one sentence. The learner’s example is three sentences, so there are more clauses and more thoughts intervening between the two connected noun phrases. (2) In Hirschman’s example, the NP is first used as a modifier, and second used as a head. In the learner’s example, this is reversed. The order and distance do not matter except in how difficult it makes it to tag the relationship between the two NPs. As mentioned, Hirschman would only tag a modifying NP as its own syntactic unit if it were referred to separately elsewhere in the writing. Due to the difficulty of maintaining reliable tagging if a coder has to go through an entire essay in order to know if the modifying NP is referred to separately, though, Hirschman’s method was slightly altered for the tagging in this study.

In the essays here, in general the pre-nominal modifying NPs were tagged as separate syntactic units in case they occurred separately elsewhere in the essays. This does not substantially affect the reliability of the tagging, because if they are solely functioning as modifiers in some kind of compound noun situation, then they will not end up co-referring to anything other than themselves in the same position later on. A new co-reference chain will be identified this way, even when the compound noun is repeated, but this will not affect the analysis in a significant way.

NPs after Copula Verbs

Assertions and predicate nominals that follow the copula or the verb *have* are usually indefinite, and the sense is that the assertions are being introduced. Following the copula, the predicate NP is equated with the subject, but the subject may be definite while the predicate is indefinite in reference. This makes it somewhat awkward to link them as relating through identity, but this is what is done under this tagging system. In the example from Learner 012 below, the subject NP *he* (meaning the learner's father) is marked as co-referring with the predicate NP *very strict person*.

Example 50: Predicate Nominal Following the Copula from Student 12

“Moreover, he was very strict person...”

(NX₁ he NX₁)/ID001 was (NX₂ very strict person NX₂)/ID001

In this case, the indefinite predicate NP does not have an overt indefinite determiner, but this is not the pattern followed by all of the learners. In the following example from Learner 014, who is another Korean IIEP student, there is an overt indefinite determiner following the copula *am*.

Example 51: Post-Copula Nominal with Indefinite Determiner from Student 14

“So if I am a very rich person for a one day...”

So if (NX₁ I NX₁)/IID001 am (NX₂ a very rich person NX₂)/ID001
for (NX₃ a one day NX₃)/ID002

In this example, the subject *I* is tagged as co-referring with the predicate NP *a very rich person*, which has the determiner *a*.

When the predicate nominal follows *have* instead of the copula, it does not usually co-refer with the subject, but would be marked as co-referring with a subsequent, definite reference to the same real-world entity. In the following example from Learner 005, the post-verbal NP that starts with *a strong belief* is delineated as a syntactic unit that can be marked for co-reference.

Example 52: Predicate NP Following *Have*, from Student 5

“The people have a strong belief on the economic growth which is followed by increasing population”

(NX₁ the people NX₁)/ID001 have (NX₂ a strong belief on (NX₃
the economic growth which is followed by (NX₄ increasing
population NX₄)/ID002 NX₃)/ID003 NX₂)/ID004

None of the NPs or NP sub-units co-refer in this example, but the predicate NP following *have* is indefinite. This case is fairly straightforward, but negation in the predicate makes it more difficult to tag similar sentences for co-reference.

Negation is a tricky situation for evaluating co-reference because typically the predicate nominals are in a syntactic position where they would be evaluated as co-referring with the subject or with a previously or subsequently mentioned NP, but the

negation may make it so that in the final evaluation of the meaning, the two NPs do not seem to be referring to the same real-world element. In example (a) below, there are three NPs that should co-refer: (1) *Korea, my country*, (2) *my country*, and (3) *an exception*.

Example 53: Negated Predicates from Two Students

- a) "...and Korea, my country is not an exception" (Student 2)
- b) "They don't need more children" (Student 6)

If "Korea, my country" and "an exception" are marked as co-referring, it is strange in terms of meaning because the negation gives the understanding that these are in separate groups and do not co-refer. However, "an exception" is in a syntactic position after the copula which equates the two noun phrases. In these cases, the NPs were marked as co-referring so that the syntactic situation of two NPs linked by the copula could be compared between native and student writers to determine what differences in article or determiner use appear after the copula.

Meaning Functions

One other tricky issue regarding co-reference tagging to be discussed is the idea of "functions." Hirschman describes functions as chains of reference "that can be collapsed." Basically, this means that there is a common noun, typically related to the topic, with which several proper nouns or other discourse referents can be linked, even those this set of other NPs do not themselves co-refer. For example, an essay may discuss a particular individual who has held two jobs at two different points in time. The main guidance for resolving this is that preference is given to names or individuals over kinds or types. Therefore, all three NPs can be linked in the same co-reference chain—

even though the two jobs are not the same and would not be included in the same chain of reference if the person were not tied to both.

It is clear that deciding how to mark the co-reference in such a chain is complicated. For a contrasting example, consider an essay describing one job, which was held by two different individuals at different times. These cannot be marked as all co-referring or being part of the same chain of reference. Three NPs that could participate in co-reference chains with other NPs are identified, but they cannot be linked because the people cannot be conflated in the same way that the jobs can. Their identity as individuals takes preference over the job. The tagging that Hirschman suggests for this is presented in the example that follows.

Example 54: Co-reference Tagging of a Function from Hirschman (1997)

“Henry Higgins, who was formerly sales director for Sudsy Soaps, became president of Dreamy Detergents”

(NX₁ Henry Higgins who was formerly (NX₂ sales director for (NX₃ Sudsy Soaps NX₃)/ID001 NX₂)/ID002 NX₁)/ID002 became (NX₄ president of (NX₅ Dreamy Detergents NX₅)/ID003 NX₄)/ID002

Another guideline for marking the co-reference of a function is that if there are two NPs whose reference cannot be conflated, but whose meaning is related to a third common noun, mark co-reference between the items in the same clause or else link the common noun to the most recent item to fill the function. This type of function was seen quite often in the EPE students' essays about the population graph. See the example on the next page. In this example, the function is which country has the largest population at which time: Pxy = x has the largest population in year y.

Example 55: A Function in the Essay of Student 6

“In 2000, China has the most population and in 2025 China still has the largest population among these countries. But in 2050 the population of India will become the most one and it surpasses the population of China.”

The variables in this function are filled with three different sets of referents:

- The population of China in 2000, $x = \text{China}$, $y = 2000$;
- The population of China in 2025, $x = \text{China}$, $y = 2025$; and
- The population of India in 2050, $x = \text{India}$, $y = 2050$.

These NPs are not all co-referential, and should not be conflated in the same chain of reference in order to maintain other chains developed over a larger patch of the essay. In this case, the names of the countries get preference over the function, followed by the years. So, the system in this case will not allow a way to mark the relationship perceived between “the greatest/largest population” and the countries that fill that role because it is more important to retain the distinctions between these elements.

Summary

These are just a few of the complicated situations that arise in the task of coding co-reference. More considerations are discussed and more examples are given in Appendix E. Notice that this scheme does not even seek to define how humans recognize co-reference. Co-reference and co-reference chains are seen as a given, and real mental construct used in understanding and creating communication. The implication is that they are universal concepts, and that individuals will recognize connections between noun phrases, more connections in fact than can be adequately coded by tagging schemes such as that used for MUC-7. The majority of the co-reference guidelines focus on limiting or decreasing the number of meaning connections that are tagged given the wide array that are identified by people. This method of tagging, the idea that co-reference

chains are universal features of human language, and the fact that connections will be actively sought in the student essays mean that they will most likely be evaluated as doing well in expressing co-reference in the second language.

Statistical Analysis Methods

Most of the data collected for this investigation will be frequency data. When possible, statistical analysis will be completed in order to evaluate the significance of any differences uncovered. The group comparisons most important for this study are:

- Native English writers and L2 English writers,
- L2 English students of lower proficiency (IIEP students) and L2 English students of higher proficiency (EPE students),
- L2 English students from Korea, and L2 English students from China/ Taiwan (different first language backgrounds).

The statistical test that will be used is the Wilcoxon Rank Sums Test for comparing the means of two populations, also known as or comparable to the Mann-Whitney U Test. This is a non-parametric test suitable for small, dependent samples. Quantitative data is ranked and the sum of the ranks is calculated for one of the groups. This sum is then used with the formula on the next page, which gives a z score whose significance can be looked up.

The resulting z score provides a probability that indicates how whether the difference between the means of the two groups is significant enough that they are likely to be different. The level of significance used in the study is $\alpha = .10$, so any probability above 10% is taken as an indication that the means are most likely different. Any resulting probability less than 10% is evaluated as indicating there is no evidence to suggest that the means of the two groups being compared are different.

Figure 17: Test Statistic for the Wilcoxon Rank Sums Test (Lapin 1983:431-436)

$$Z = \frac{W - \left[n_A \times \frac{n_A + n_B + 1}{2} \right]}{\sqrt{\frac{n_A + n_B + 1}{12}}}$$

n_A = the number of participants in group A (for example, the number of students)

n_B = the number of participants in group B (for example, the number of native writers)

W = the sum of the rankings for the values in group A (for example, if there are two in group A with the values 5 and 15, and there are two in group B with the values 1 and 30, the ranks for group A would be second and third, and the rank sum, or W , would be five)

CHAPTER III. RESULTS, GOALS ONE AND TWO

In this section, data and conclusions analyzed to satisfy the first two research goals will be examined. These research goals are:

Goal 1: To describe some characteristics of co-reference and discourse construction in native and L2 essays in English.

Goal 2: To determine where there are significant differences in co-reference and discourse construction between (a) native and L2 writers, (b) L2 writers from China and Korea, and (c) L2 writers of different proficiency levels.

The native and L2 English essays were compared by looking at the number of total words, number of NPs, number of co-reference chains, and more.

Total Number of Words

The total number of words in an essay may be important in an analysis for a number of reasons. First, the length of the essay may indicate the proficiency of the writer to a certain extent. Second, the longer the essay and the more total number of words, the more noun phrases there are, and the more opportunities there are to evaluate how the discourse is constructed using noun phrases. In this section, the total number of words written by the native speakers will be compared to the total number in the student essays. Then, the students will be compared against each other in order to see if there is any significant difference between the EPE and IIEP students, or between the Korean and the Chinese/Taiwanese students.

Native Speakers vs. Student Writers, Number of Words

When considering the total numbers of words in these essays, it is important to remember the conditions under which the essays were collected. All the participants had only 30 minutes in which to read two essay questions presented to them, select one, and then compose and finalize their essay. The ESL students were monitored as they completed the test, and used paper and pencil to write their essay. The native speakers

chose when and where to complete their essays, and were able to type their essays if they so desired. They were also not as closely supervised.

Table 18: Total Number of Words per Essay

<i>ESL Students</i>		<i>Native English Speakers</i>	
<i>participant #</i>	<i># words</i>	<i>participant #</i>	<i># words</i>
<i>1</i>	170	<i>21</i>	93
<i>2</i>	328	<i>22</i>	311
<i>3</i>	220	<i>23</i>	159
<i>4</i>	171	<i>24</i>	231
<i>5</i>	193	<i>25</i>	365
<i>6</i>	233	<i>26</i>	319
<i>7</i>	190	<i>27</i>	199
<i>8</i>	171	<i>28</i>	296
<i>9</i>	312	<i>29</i>	412
<i>10</i>	205	<i>30</i>	428
<i>11</i>	249	<i>31</i>	398
<i>12</i>	265	<i>32</i>	662
<i>13</i>	195	<i>33</i>	343
<i>14</i>	216	<i>34</i>	241
<i>15</i>	295	<i>35</i>	279
<i>16</i>	243	<i>36</i>	560
<i>17</i>	158	<i>37</i>	380
<i>18</i>	125	<i>38</i>	469
<i>19</i>	177	<i>39</i>	359
<i>20</i>	171	<i>40</i>	146
<i>total</i>	4,287	<i>total</i>	6,650
<i>average</i>	214.35	<i>average</i>	332.5

The questions for both groups were similar, and fit the kind of variation that was seen in test bank of questions for the EPE and IIEP tests. The preceding table presents the number of words written by each participant—both native speakers and students.

As can be seen, both the shortest (93 words, Native 21) and the longest (662 words, Native 32) essays were written by one of the native English speakers. The shortest *student* essay is 125 words and was written by Student 18 (IIEP, Taiwan). Only Native 21 has a shorter essay than Student 18. The longest student essay is 328 words long and was written by Student 2 (EPE, Korea). This is about half as long as the longest native speaker essay. When comparing the longest student essay to the length of all the native speaker essays, it can be seen that ten native speakers wrote essays longer than 328 words, and ten wrote essays shorter than 328 words. The shortest native speaker essay, that of Native 21, is also not markedly shorter than the other native speaker essays. Native 40's essay was only 53 words longer, and Native 23's was only 66 words longer³⁴. Thus, native speakers were somewhat more likely to write longer essays, and the lengths of the native speaker essays are more varied or widely distributed than those of the students essays.

In the aggregate, it also seems that native writers tended to write longer essays. When the lengths of all the student essays are averaged and compared with the average of the native writers' essays, the student essays are shorter on average by about 115 words. When the medians of the lengths are compared, a similar conclusion can be made. The

³⁴ It is interesting to note that even though Native 21's test length is not unusual, Native 21 felt it necessary to report a reason for the short length. She stated that she felt fatigue in the middle of writing the test and therefore rapidly (and possibly poorly) put an ending on her essay, thereby accounting for the short length. Her need to report this when such feedback was not requested shows that she took the test seriously and wanted to do well. Although all the students and the native speakers were informed that they could end their essay at any time, no one else reported ending abruptly. It is true, though, that both EPE and IIEP test proctors have observed students finishing their essays quickly, and some have afterwards reported that this was due to fatigue. So in a larger corpus, potentially other students' essays would be shorter than expected due to fatigue.

median length of the native essays is about 115 words longer than the median length of the student essays. This is shown in the table that follows.

Table 19: Average and Median Total Number of Words per Essay

	<i>Students</i>	<i>Natives</i>
<i>Shortest essay</i>	125 words, Student 18	93 words, Native 21
<i>Longest essay</i>	328 words, Student 2	662 words, Native 32
<i>Average # words</i>	214	333
<i>Median # words</i>	200	331
<i>Wilcoxon Rank Sums Test</i>	$z = -3.03, p = .0024$	

Using the Wilcoxon Rank-Sum test to compare the distribution of the lengths of the student essays with the distribution of the lengths of the native speaker essays, this difference is statistically significant ($z = -3.03, p = .0024$). This result supports what one might expect: that more proficient language or more native-like language users have the capacity to produce more words than those less proficient when given the same amount of time or similar tasks. The next section examines this conclusion further by comparing the lower level students to the higher level students, and by comparing students based on their L1.

IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese

Students, Number of Words

In the last section, it was seen that the native writers were more likely to write longer essays than the students. This section will examine what differences, if any, exist between the length of the essays of the students when they are divided into smaller groups based on their proficiency or primary language/country of origin.

Comparing the EPE to the IIEP students, as can be seen in the next table, at first glance, the EPE students appear to write longer essays when comparing the longest and shortest essays and when looking at the average lengths. The longest EPE essay was longer than the longest IIEP essay, at 328 words (Student 2) and 295 words (Student 15), respectively. The shortest EPE essay was also longer than the shortest IIEP essay. When comparing the aggregate data of the two student groups, this conclusion seems less obvious. The average number of words in an EPE essay (219.3) is higher than the average number in an IIEP essay (209.4). However, the median length of the IIEP essays (205.5) is slightly longer than the median length of the EPE essays (199). In fact, there is no significant difference between the distributions of the number of words when comparing either the IIEP students with the EPE students ($z = -.08$, NS [$p = .9362$]). So, although when comparing the length of essays by individuals it seems that the EPE students write slightly longer essays, this difference on the whole is not noteworthy.

Table 20: Average and Median Total Number of Words per Essay, IIEP vs. EPE and Korea vs. China/Taiwan

	<i>IIEP students</i>	<i>EPE students</i>	<i>Korean Students</i>	<i>Chinese/Taiwanese</i>
<i>Longest essay</i>	295 words, Student 15	328 words, Student 2	328 words, Student 2	312 words, Student 9
<i>Shortest essay</i>	125 words, Student 18	170 words, Student 1	170 words, Student 1	125 words, Student 18
<i>Average # words</i>	209.4	219.3	227.8	200.9
<i>Median # words</i>	205.5	199	218	192.5
<i>Wilcoxon Rank Sum</i>	$z = -.08$, NS [$p = .9362$]		$z = .98$, NS [$p = .3720$]	

Breaking down the groups by nationality, it appears that the Korean students wrote longer essays than the Chinese or Taiwanese students. The longest Korean student essay (328 words, Student 2) was longer than the longest Chinese or Taiwanese student essay (312 words, Student 9). The shortest Korean student essay (170 words, Student 1) is also longer than the shortest Chinese or Taiwanese student essay (125 words, Student 18). The aggregate comparisons also support that there may be a trend of longer essays by the Korean students. The average length of the Korean essays (227.8 words) is about 30 words longer than the average length of the other student group (200.9 words). The median length of the Korean essays (218 words) is also longer than the median length of the group of Chinese and Taiwanese essays (192.2 words). However, as with the comparison of the IIEP to the EPE students, despite this, the difference between the distributions of the number of words when comparing the Korean students to the Chinese/Taiwanese students is not significant ($z = .98$, NS [$p = .3720$]).

So it can be seen that there is no great difference between the length of the writing of the students, either when compared by proficiency level or by nationality. As also shown, though, there is a significant difference between the distribution of the lengths of the native speaker essays and the distribution of the lengths of the student essays. The native essay lengths cover a wider range of lengths, and in general tend to be longer than the student essay lengths. The greater variability could be a reflection of the fact that while the students were somewhat controlled for writing proficiency, the native speakers were not. The greater variation also could be a sign of greater ability in manipulating writing style, though, as skilled writers can choose to present an argument succinctly, or in a more verbose fashion with more explanations and examples incorporated.

Total Number of Noun Phrases

Native Speakers vs. Students, Noun Phrases

The table below shows the total number of noun phrases in each essay.

Table 21: Total Number of Noun Phrases per Essay

<i>ESL students</i>		<i>Native English speakers</i>	
<i>participant #</i>	<i># NPs</i>	<i>participant #</i>	<i># NPs</i>
1	65	21	32
2	111	22	101
3	73	23	51
4	74	24	63
5	63	25	107
6	89	26	101
7	61	27	61
8	58	28	73
9	80	29	131
10	59	30	116
11	84	31	114
12	101	32	197
13	73	33	103
14	76	34	76
15	98	35	85
16	62	36	148
17	54	37	82
18	40	38	129
19	66	39	82
20	64	40	45
<i>total</i>	1,451	<i>total</i>	1,898

The number of noun phrases is important for looking at the use of determiners because if students use a smaller or larger percentage compared to native speakers, it could point to a different pattern in forming and manipulating discourse coherence.

Both the essays with the smallest (32 NPs, Native 21) and the largest (197 NPs, Native 32) numbers of noun phrases were written by native English speakers. These two

essays also contain the smallest number of words and the largest number, respectively. This demonstrates that the more words in general a writer writes, the more noun phrases they will also compose. The smallest number of noun phrases in a student essay is 40 NPs and was written by Student 18 (IIEP, Taiwan), who also wrote the shortest student essay. The largest number of noun phrases in a student essay is 111 NPs, and this essay was written by Student 2 (EPE, Korea), who also wrote the longest student essay.

So, for both the student essays and the native speaker essays, generally the number of noun phrases and the number of words vary directly with each other. But is there a significant difference between how this plays out for the native writers versus the learners? In fact, when the distributions of the total number of noun phrases in the student essays and the total number of noun phrases in the native speaker essays are compared, they are significantly different ($z = -2.22$ [$p = .0264$]). However, given that the distributions of the total number of words are significantly different, this is somewhat expected. As the number of words increases, the number of nouns should increase as well, but there are significant differences in the number of each when comparing the native writers' essays and the students' essays.

But what can be said about the proportion of the number of noun phrases to the total number of words? In fact, when you normalize the number of noun phrases by dividing the total number of noun phrases by the total number of words in each particular essay, the distributions of the native and learner NPs are significantly different ($z = 3.81$ [$p \leq .0001$]). As can be seen in the table on the following page, the individual students' essays tend to have a higher concentration of noun phrases.

Table 22: Ranks of Total Number of Noun Phrases per Total Number of Words

<i>participant #</i>	<i>participant category</i>	<i>total # words</i>	<i>total # noun phrases</i>	<i>total NPs / total words</i>	<i>rank</i>
37	native	380	82	.215	1
39	native	359	82	.228	2
28	native	296	73	.247	3
16	student	243	62	.255	4
9	student	312	80	.256	5
36	native	560	148	.264	6
30	native	428	116	.271	7
24	native	231	63	.273	8
38	native	469	129	.275	9
31	native	398	114	.286	10
10	student	205	59	.288	11
25	native	365	107	.293	12
32	native	662	197	.298	13
33	native	343	103	.300	14
35	native	279	85	.305	15
27	native	199	61	.307	16
40	native	146	45	.308	17
34	native	241	76	.315	18
26	native	319	101	.317	19
29	native	412	131	.318	20
18	student	125	40	.320	21
23	native	159	51	.3207	22
7	student	190	61	.321	23
5	student	193	63	.326	24
22	native	311	102	.328	25
3	student	220	73	.3318	26
15	student	295	98	.332	27

Table 22—Continued

<i>participant #</i>	<i>participant category</i>	<i>total # words</i>	<i>total # noun phrases</i>	<i>total NPs / total words</i>	<i>rank</i>
11	student	249	84	.337	28
2	student	328	111	.338	29
8	student	171	58	.339	30
17	student	158	54	.342	31
21	native	93	32	.344	32
14	student	216	76	.352	33
19	student	177	66	.373	34
20	student	171	64	.3742	35
13	student	195	73	.3743	36
12	student	265	101	.3811	37
6	student	233	89	.3819	38
1	student	170	65	.3823	39
4	student	171	74	.433	40

There is a higher likelihood that any randomly selected word in one of the students' essays will be part of a noun phrase. So, students on average use more noun phrases than native speakers in this corpus. As can be seen in the table, Student 4 (EPE, Korea) has the highest percentage of noun phrases per word at .433, but Student 4 does not have the largest number of words or noun phrases compared to the other student essays. The native speaker with the highest ratio of noun phrases to words is Native 21 (ratio: .344), who ranks thirty-second out of forty. Again, Native 21 has the shortest essay of all the writers, and also the least number of noun phrases. Native 37 has the lowest percentage of noun phrases per word at .215. The student with the lowest percentage is Student 16 (IIEP, Taiwan), who has a ratio of noun phrases to words of .255 and ranks fourth. In fact, if you look at the rankings, out of the 20 essays with the

lowest percentages of noun phrases per total number of words, 17 were written by native speakers, and only three by students.

IIEP vs. EPE Students, Noun Phrases

As mentioned previously, there was no significant difference in the distribution of the number of words in the IIEP and EPE students' essays ($z = -.08$, NS [$p = .9362$]).

When the number of noun phrases in the IIEP and EPE essays are compared, there is no significant difference either ($z = .08$, NS [$p = .9362$]). Moreover, when the *ratio* of the number of noun phrases to the number of total words is examined, there is again no significant difference between the distributions of the EPE students and the IIEP students ($z = .30$, NS [$p = .7642$]). This information is summarized in the table the follows on the next page.

Table 23: IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums

Test

	z	p	<i>significant difference?</i>
<i>total # words</i>	-.08	NS [.9362]	no
<i>total # NPs</i>	.08	NS [.9362]	no
<i>ratio of #NPs per total # words</i>	.30	NS [.7642]	no

This data shows that the proficiency level of the students in this study may not have an effect on how they use noun phrases and determiners in discourse.

Korean vs. Taiwanese/Chinese Students, Noun Phrases

When looking at just the total number of words per essay, there was no significant difference found between the Korean and the Taiwanese or Chinese students (using the

Wilcoxon Rank Sums Test, $z = .98$, NS [$p = .3720$]). However, there was a significant difference between the number of NPs in the essays of the two language/country groups ($z = 2.27$ [$p = .0232$]). When the ratios of the number of total noun phrases to the total number of words are compared, there is again no significant difference. See the summary of these details in the table that follows.

Table 24: Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Wilcoxon Rank Sums Test

	z	p	<i>significant difference?</i>
<i>total # words</i>	.98	NS [.3720]	no
<i>total # NPs</i>	2.27	.0232	yes
<i>ratio of #NPs per total # words</i>	1.44	NS [.1498]	no

These results suggest that there is some difference in the use of noun phrases by the Chinese/Taiwanese students and the Korean students. The immediately preceding table shows the ranks of the individual Korean, Chinese, and Taiwanese students.

Student 18 (IIEP, Taiwan) has the lowest number of noun phrases, with only 40 total. In fact, the six essays with the fewest number of NPs were all written by Chinese or Taiwanese students. The Korean student with the lowest number of NPs (Student 5) ranks seventh, with 63 total NPs. This implies that the Korean students in general included more NPs in their essays than the Chinese/Taiwanese students. This belief is strengthened when looking at the essays with the most NPs. The essay with the highest number of noun phrases was Korean (Student 2, EPE), with 111 total NPs. In fact, the three essays with the most NPs were all written by Korean students.

Table 25: Ranks of the Total Number of Noun Phrases, Chinese/Taiwanese Students vs. Korean Students

<i>participant #</i>	<i>participant category</i>	<i>total # NPs</i>	<i>rank</i>
18	IIEP, Taiwan	40	1
17	IIEP, Taiwan	54	2
8	EPE, Taiwan	58	3
10	EPE, China	59	4
7	EPE, China	61	5
16	IIEP, Taiwan	62	6
5	EPE, Korea	63	7
20	IIEP, Korea	64	8
1	EPE, Korea	65	9
19	IIEP, Taiwan	66	10
13	EPE, Korea	73	11
3	IIEP, Taiwan	73	12
4	EPE, Korea	74	13
14	IIEP, Korea	76	14
9	EPE, China	80	15
11	IIEP, Korea	84	16
6	EPE, China	89	17
15	IIEP, Korea	98	18
12	IIEP, Korea	101	19
2	EPE, Korea	111	20

The Chinese or Taiwanese student with the most NPs is Student 6 (EPE, China), who wrote 89 NPs. Out of the top ten essays with the most NPs, seven were written by Korean students, and only three written by Chinese or Taiwanese students. The fact that there was no significant difference between the Chinese/Taiwanese and the Korean

students in terms of the total number of words indicates that there is something specific to noun phrase usage that differs between the two groups.

This is different from the significant difference between the number of NPs in the native writers' and L2 students' essays, which was somewhat expected in that the number of words were significantly different: The natives tended to write more words and more NPs. In this case, it is not just that the Korean students are tending to write more words, because they are not. Neither are they using more NPs per word, because that ratio was not found to be a significant difference between the two groups. For this group of students, there is some effect of L1 on the L2 usage of noun phrases specifically.

Co-Reference Chains

As mentioned previously, a co-reference chain consists of the noun phrases, determiners, pronouns, and other nominals that refer to the same real-world idea or object. Co-reference chains are universal mental constructs. They are connections of ideas or things repeated over the course of the communication. Therefore, students would be expected to do well in creating discourse chains. The issue this investigation is concerned with, is what differences are there in the functional morphology or the full NP structure of NPs participating in a chain of reference. The theoretical consideration of what defines co-reference chains formally is beyond the scope of the current study.

Since what is considered old or background information in the discourse is important to the selection of nominals and determiners, the connections between each stepping stone in the chain of reference and the surface morphological/ lexical form is important for determining how the writer is building their connected message. This is potentially one of the key elements in looking at the meanings learners and native speakers are assigning to determiners.

Table 26: Number and Length of Co-Reference Chains by Participant

<i>ESL Students' Chains</i>					<i>Native English Writers' Chains</i>				
<i>student #</i>	<i>total #</i>	<i>mean length</i>	<i>median length</i>	<i>mode length</i>	<i>native #</i>	<i>total #</i>	<i>mean length</i>	<i>median length</i>	<i>mode length</i>
1	11	3.27	3	2	21	5	3.2	2	2
2	15	3.2	2	2	22	16	3.69	2.5	2
3	9	2.78	3	2	23	8	4	2.5	2
4	18	2.83	2	2	24	13	2.92	2	2
5	10	2.8	3	3	25	16	3.25	2	2
6	25	3	2	2	26	11	4.82	2	2
7	11	3.09	2	2	27	6	4	3	2
8	9	2.67	2	2	28	9	3	2	2
9	16	2.94	2.5	2	29	22	2.86	2	2
10	8	3.5	3.5	2	30	13	5	3	2
11	8	5.75	3.5	3	31	16	3.31	3	2
12	8	10.25	4	2	32	24	3.79	2.5	2
13	6	8.17	2.5	2	33	13	4.08	4	2
14	11	5.27	2	2	34	12	3.17	2	2
15	13	4.92	3	2	35	11	3.09	2	2
16	11	4.18	2	2	36	19	3.89	2	2
17	5	7.2	4	2	37	18	2.94	2	2
18	7	3.57	2	2	38	19	3.68	3	2
19	8	4.38	2.5	2	39	17	2.94	2	2
20	10	4.6	4	7	40	8	3	2	2
<i>avg. of all stud.s</i>	10.95	4.42	2.73	2.35	<i>avg. of all natives</i>	13.8	3.53	2.4	2

This section first compares the number of co-reference chains in each piece of writing, and the number of nominals the writers connected to the same referent on an individual level. After that, group data is compared by looking at the performance of

students vs. native writers, IIEP vs. EPE students, and Korean vs. Chinese or Taiwanese students. The average length of the chains is examined, as well as the number of chains when normalized by considering them per the total number of words and per the total number of NPs in the essays.

Number and Length of Co-Reference Chains for Individual Writers

The number and length of the co-reference chains in the individual essays varied widely. In some cases, the total count and the length of the chains seemed to be affected by the length of the essay, the essay topic, or the language proficiency of the writer, but in some cases it was not immediately clear what might account for the variation. The preceding table summarizes information about the number and length of co-reference chains by participant.

Student 6 (EPE, China) had the largest number of co-reference chains of all the students and native speakers, with 25 co-reference chains. Native 32 had the most co-reference chains per essay for the native speakers, with 24 total co-reference chains of noun phrases. Student 17 (IIEP, Taiwan) and Native 21 tied for the least number of co-reference chains per essay, with only five chains each. Averaging all the native speakers together, though, they tended to have more co-reference chains than the ESL students, with an average of 13.8 and 10.95, respectively.

When comparing the length of these co-reference chains (length meaning the number of referring nominals per chain), it can be seen that Student 3 (EPE, Korea) has the shortest co-reference chains on average, with a mean of 2.78 noun phrases per chain. Four of Student 3's nine chains in fact only have one link, or two noun phrases chained together. In fact, most of the participants, native or student, had a high number of two-item chains. Despite being small chains, these two-item chains are not trivial. Often, the

link appears in a following sentence, giving the opportunity to look at how the noun phrase form changes over discourse.

Native 29, the native speaker with the shortest average chains, is not far off from this, with a mean of 2.86 noun phrases per chain. Fourteen of Native 29's twenty-two chains contain only two linked noun phrases, so like Student 3, Native 29 has a high percentage of small chains. However, a learner's essay also has the longest chain. Student 12 (IIEP, Korea) has the longest mean co-reference chain, averaging 10.25 noun phrases per chain. In contrast, the longest average co-reference chain length for a native speaker is about half as long: Native 30 has the longest average, but this is still only five noun phrases per chain.

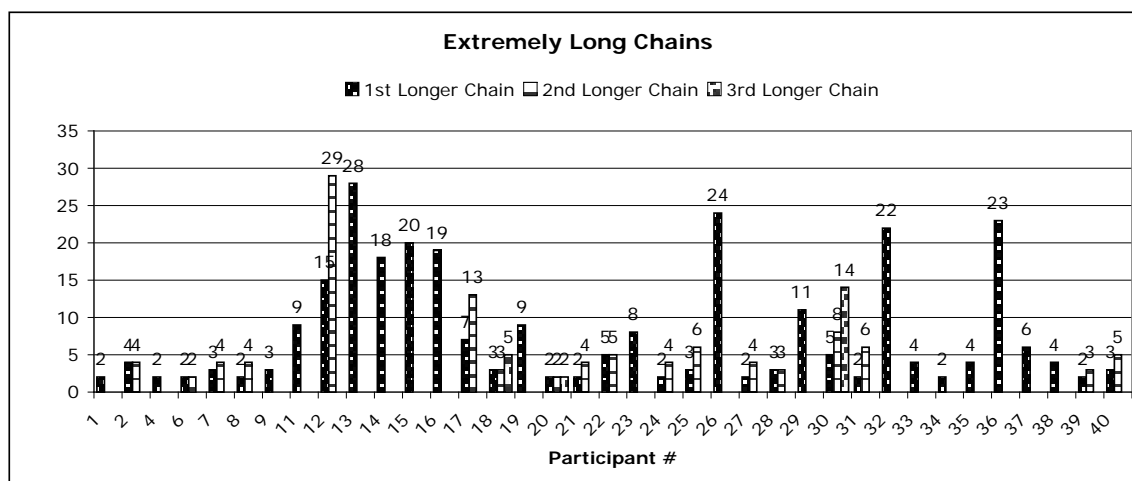
This seems like a somewhat unexpected result. In fact, when averaging the mean lengths of all of the students' chains, it can be seen that they are typically longer than the averages of the native speakers' chain lengths. The overall average for students was 4.42 noun phrases per chain, and for native speakers it was 3.53 noun phrases per chain. So, not only is Student 12 constructing chains that are unusually long, but the students in general have longer chains than the native speakers. In the case of Student 12, the long chains may be an artifact of the topic of his essay. As a narrative about his father's life, Student 12's essay is focused on one, real individual and real past events instead of presenting a variety of examples focusing on different real-world referents to support a general point. One possible reason why *in general* the student chains are longer may actually be their less proficient use of written English and the way that co-reference was marked based on the MUC-7 coding specifications. The MUC co-reference coding rules are limiting because, for example, part-whole relationships cannot usually be linked because they don't refer to exactly the same set of referents (a phrase listing all elements joined by *and* could be linked to a nominal referring to the whole set, but this is not common). For example, Native 25's essay includes this series of noun phrases:

Example 56: Sequence of Related Noun Phrases by Native 25 which are Not Coded as Co-referring According to the MUC-7 Rules (given in order of appearance)

8. “the public universities’ proposed plan to eliminate the general education requirements...”
9. “short-sighted thinking on the part of the universities”
10. “More progressive universities”
11. “Yale Medical School”
12. “The University of Virginia”
13. “These two examples”

The first two of these example noun phrases are marked as co-referring under the MUC system. The third noun phrase refers to a different real-world element, but the word *more* harkens back to the referent of the first two. In this way, they are connected in the discourse, but this connection is not noted in the MUC system. In addition, although the last four examples are related in a part-whole way, none of them are linked as co-referring. *Yale Medical School* is one of the *more progressive universities*, so there is a connection there, but they are not the exact same set. *The University of Virginia* is a member of the set of *more progressive universities* and the set of *these two examples*, but does not refer to the exact same real-world entity. None of these four phrases use the same words and the connection may be more subtle than a less than advanced learner could make. This will be examined again later and in more detail in the Discussion.

Figure 18: Comparison of Extra Length of Extremely Long Chains in the Essays of Most Participants



Other Patterns in Co-reference Chains

Because the length of the individual chains varied greatly, the median and the mode chain length was also calculated in order to give a fuller picture of the chain lengths and their variation. Students 12 (IIEP, Korea), 17 (IIEP, Taiwan), and 20 (IIEP, Korea) had the longest median chain lengths, with a median of four. Native 33 also had a median chain length of four, and this was the longest median chain length for the native speakers as well.

When the length of the individual chains in the essays are compared, it can be seen that all the writers except for three (Students 003, 005, and 010), had one to three chains that were longer than the others by at least two members. Four native writers (Natives 26, 29, 32, and 36) and five student writers (Students 12, 13, 14, 15, and 16) all

had one chain that had over ten members more than the other co-reference chains in their essays. These much longer chains skew the averages.

Table 27: Data Table for the Previous Figure: Extra Lengths on Extremely Long Co-reference Chains

<i>Participant #</i>	<i>1st Long Chain's Excess</i>	<i>2nd Long Chain</i>	<i>3rd Long Chain</i>	<i>Participant #</i>	<i>1st Long Chain's Excess</i>	<i>2nd Long Chain</i>	<i>3rd Long Chain</i>
1	+2	--	--	21	+2	+4	--
2	+4	+4	--	22	+5	+5	--
3	--	--	--	23	+8	--	--
4	+2	--	--	24	+2	+4	--
5	--	--	--	25	+3	+6	--
6	+2	+2	--	26	+24	--	--
7	+3	+4	--	27	+2	+4	--
8	+2	+4	--	28	+3	+3	--
9	+3	--	--	29	+11	--	--
10	--	--	--	30	+5	+8	+14
11	+9	--	--	31	+2	+6	--
12	+15	+29	--	32	+22	--	--
13	+28	--	--	33	+4	--	--
14	+18	--	--	34	+2	--	--
15	+20	--	--	35	+4	--	--
16	+19	--	--	36	+23	--	--
17	+7	+13	--	37	+6	--	--
18	+3	+3	+5	38	+4	--	--
19	+9	--	--	39	+3	+2	--
20	+2	+2	+2	40	+3	+5	--

Not only is the longest median chain length the same for the native speakers and students, but the shortest median chain length is as well. Eight different students³⁵ had median chain lengths of just two noun phrases (the least number possible in a chain), and twelve native writers out of the twenty total³⁶ had median chain lengths of only two. In fact, in the native English essays, 54% of the co-reference chains only included two references (149 out of a total 277 chains). In the student essays 48% of the co-reference chains only consisted of two nominals (105 out of 218 chains). Looking more in depth at the eight students that had median chain lengths of two, three of these were IIEP students (Students 14, 16, and 18), and five of these were EPE students (Students 2, 4, 6, 7, and 8). Therefore, program or proficiency level do not seem to be relevant for this feature. Distinguishing the students with median chain lengths of two by country of origin, three of the eight were from Korea (Students 2, 4, and 14), two were from China (Students 6 and 7), and three were from Taiwan (Students 8, 16, and 18). From this, on the surface it appears that first language does not affect the use of mainly short co-reference chains either. Also, it is not the case that all the IIEP or EPE students with median chain lengths of two units are from the same country. For example, two of the EPE students are Korean (Students 2 and 4), two are Chinese (Students 6 and 7), and one is Taiwanese (Student 8).

In regards to the mode length, only three out of the entire 40 participants had mode lengths longer than two noun phrases per chain. Student 20 (IIEP, Korea) had a mode length of seven noun phrases per chain, while Students 5 (EPE, Korea) and 11 (IIEP, Korea) each had mode lengths of three noun phrases per chain. No native speakers had a mode chain length other than two. Overall, short co-reference chains are most

³⁵ Students 2, 4, 6, 7, 8, 14, 16, and 18.

³⁶ Natives 21, 24, 25, 26, 28, 29, 34, 35, 36, 37, 39, and 40.

common, with only two noun phrases per chain as the norm for the majority of participants. These short chains co-exist in the same piece of writing with a few chains of substantial length.

The figures that follow demonstrate this in graphical form. In these graphical representations, the bars show the span of the co-reference chains across the noun phrases used in the essays. Each vertical mark (where bars change colors) signifies a noun phrase in the co-reference chain. So if a bar is only one color, such as the bar for Chain A, then there are only two noun phrases in that chain. A bar with multiple colors therefore represents a chain with multiple members. For example, Chain E, which is 6 different colors, shows the discourse span of the eight noun phrases that refer to the same real-world referent in this co-reference chain. If a bar of a particular color is very long, such as the light gray bar second in Chain D, it means that there was a relatively long section of text where that idea was not referred to before it was re-mentioned. In this case, 40 nouns phrases linked to other referents intervened before this referent was discussed again.

In the case of Student 7, then, the following figure shows that two long co-reference chains co-exist with several shorter chains. Chains E and F (with several bar color changes) are the two long chains in this piece of writing. All of the other chains are very short, most with only two noun phrases. In fact, except for Chain D, the short chains are also fairly localized. The referents are briefly discussed and then not mentioned again. The next figure shows a similar pattern in a native speaker essay. As in Student 7's essay, most of the chains in Native 25's essay are only two to three noun phrases long, and are relatively localized. Chains D, F, and L are exceptions.

Figure 19: Student 7, A Mixture of Short and Long Co-reference Chains in the Same Writing

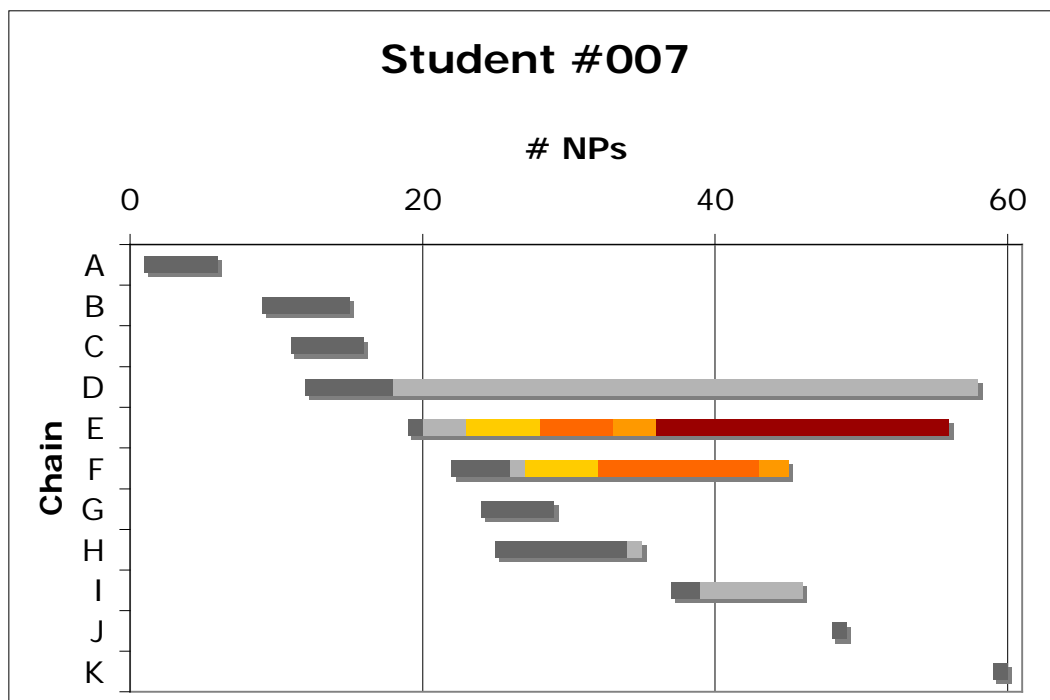
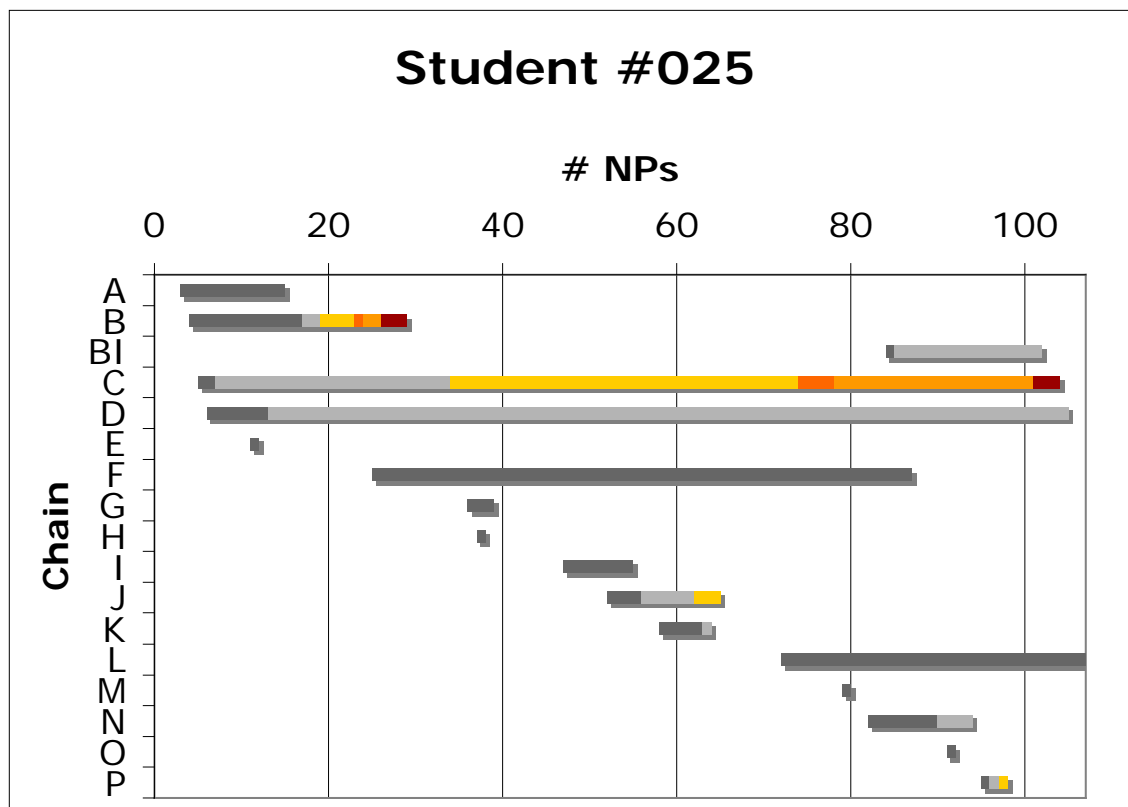


Figure 20: Native 25, A Mixture of Short and Long Co-reference Chains in the Same Writing³⁷



Despite having few chain members, the final member in the chain is re-mentioned only after quite a stretch of intervening text (after 92, 62, and 35 intervening noun phrases, respectively). Also similar to Student 7, Native 25 has two long chains—Chains B and C. The many color changes in the bars representing the stretch of the chains show that there are more noun phrases in each of these chains of reference.

³⁷ Please note that Chains B and BI are actually two parts of the same chain. Due to limitations in the chart function of Microsoft Excel, splitting some of the longer chains into separate bar representations was necessary. This figure therefore does not show a bar representing the span of this co-reference chain across the text from noun phrase 29 to noun phrase 84.

Just as in Student 7's essay, this essay contains two long chains, and then several short chains in which a referent is briefly discussed and then not gone back to. A couple of ideas link or cohere the essay as a whole, with examples and brief references to other concepts sprinkled throughout. Taken all together, this data shows that in general, these writers structure their discourse with one or two long chains of reference linked over the whole of their essays, and supported by many two item chains.

Another common pattern in the discourse structure is that some of the referents of these two-item chains act as book-ends—appearing in the introduction and the conclusion of the essay only. This can be seen in the chain pattern of Student 9 (see the next figure). Student 9's essay has three phases that can be distinguished by the numbers assigned to the NP referents:

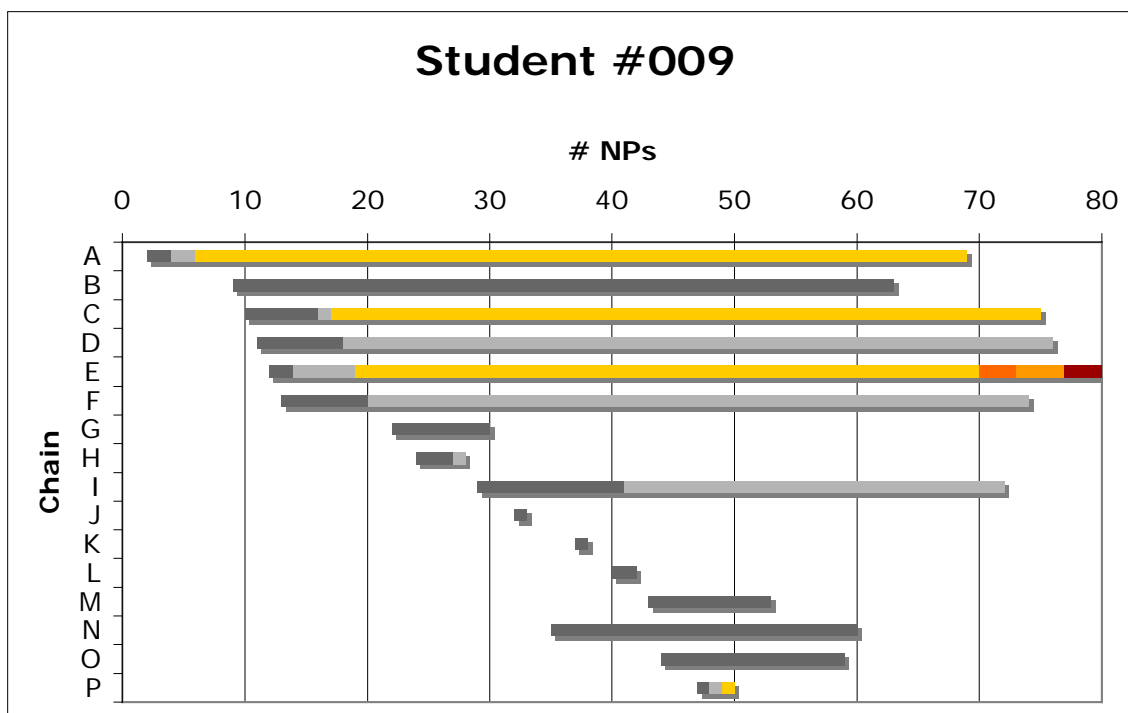
- Referents zero to 20,
- Referents 20 to 60, and
- Referents 60 to 80.

The referents that participate in chains A through F are all introduced or maintained in the set of the first 20 NPs. Then they are not re-mentioned until after the 60th NP. This is typical of the common US discourse structure of a five-paragraph essay.

So, there is a section of text in which these referents are not discussed. They are referred to before NP #20 and after NP #60. The real-world ideas or objects referred to by these chains therefore act as bookends sandwiching the other referents that are discussed in the 40 intervening noun phrases. Of these 40 intervening noun phrases, only the referent of Chain I is re-mentioned in the final section when the referents from the first 20 noun phrases are repeated. This is similar to a discourse or essay pattern that has been attested in native English writings—a general-specific-general pattern (see for example Swales & Feak 2004). In this pattern, a general topic is introduced, specific examples or explanations supporting the main point of the writing are presented, and then

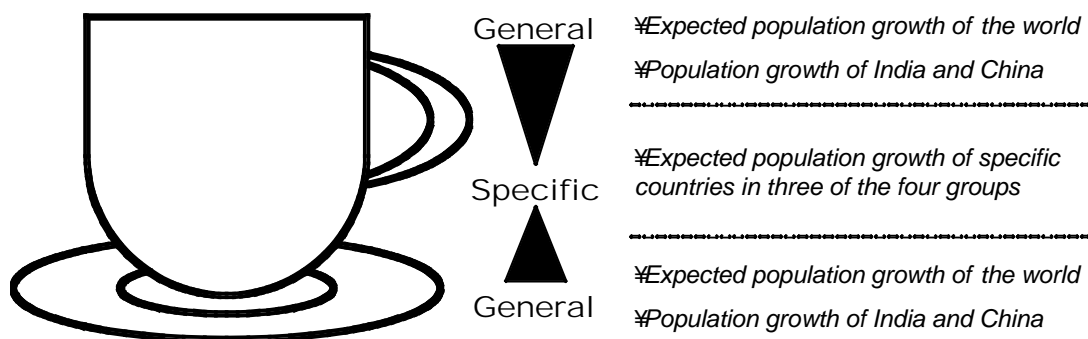
the writing is concluded with a return to a more general perspective through a summary or some conclusion linked to wider concepts.

Figure 21: Student 9, Book-Ended Referents in the Same Writing



The structure of the ideas can be metaphorically described as a teacup. See the next figure. On the right hand side of the figure, the topics of the three sections of the essay are shown. Note how the topics of the first and the third section are the same. In fact, NP number 62 is immediately preceded by the phrase “to sum up.” This phrase clearly indicates the author’s intent to go back to a more general perspective relating to the overall topic of the essay. This pattern is common in other student and native speaker essays.

Figure 22: The ‘Teacup’ Pattern of Discourse in Student 9’s Essay—General-Specific-General Idea Presentation (Swales & Feak 2004)



Group Comparisons of Co-reference Chains

When comparing groups by looking at their average co-reference length and using the Wilcoxon Rank Sums test, it can be seen that there is only a significant difference between the average chain lengths of the IIEP and EPE students. No significant difference was found between the native writers and the learners taken all together. Neither was there seen to be a significant difference between the Chinese/Taiwanese writers and the Korean writers. This information is summarized in the table that is following.

These findings support what was tentatively concluded earlier—that the greatest difference in the length of co-reference chains can be seen when comparing the less proficient IIEP students to the more proficient EPE students rather than when comparing the native languages of the learners. It is interesting, though, that when the IIEP and EPE students are considered together, the length of their chains are not significantly different from the native speakers.

Table 28: Average Length of Co-Reference Chains in Essays, Wilcoxon Rank Sums
Test Results

	<i>z</i>	<i>p</i>	<i>significant difference?</i>
<i>Students vs. native writers</i>	.68	NS [p = .4966]	no
<i>IIEP vs. EPE students</i>	3.78	.0002	yes
<i>Korean vs. Chinese/Taiwanese students</i>	.23	NS [p = .8180]	no

When looking at the individual means and ranks of the EPE and IIEP students, the difference between the two proficiency groups is striking. As shown in the table that follows, when ranking the mean chain lengths, the ten EPE students rank as the ten lowest mean lengths; the ten IIEP students rank as the ten largest mean lengths. When looking at the individual scores of the native writers, they tend to have shorter chain lengths than the learners. Again it seems that more proficient writers may use shorter co-reference chains in general, or vary their noun phrase constructions enough that their chains cannot be detected by the MUC-7 co-reference coding rules.

Natives vs. Students, Number of Chains per Total Number of Words³⁸

The results that have been examined so far have looked at the absolute number of chains and length of chains. However, the essays varied in length and the number of noun phrases, so to refine the examination, the ratio of the number of co-reference chains to the total number of words in the essays and the ratio of the number of co-reference chains to the total number of noun phrases was also examined. When comparing the

³⁸ The total number of words per essay and the total number of noun phrases per essay is in direct proportion in the student and native essays. Therefore, the number of chains could have been evaluated against the number of NPs or the number of words. Words was chosen in this case simply because most of the statistics are controlled against the number of words.

native writers' essays to the student essays, there is a significant difference in the absolute total number of co-reference chains ($z = -2.03$ [$p = .0424$]).

Table 29: Mean Length of Co-reference Chains: A Comparison of IIEP and EPE Students

<i>Participant ID</i>	<i>Course of Study</i>	<i>Country</i>	<i>Mean</i>	<i>Rank</i>
8	E	T	2.67	1
3	E	K	2.78	2
5	E	K	2.8	3
4	E	K	2.83	4
9	E	C	2.94	5
6	E	C	3	6
7	E	C	3.09	7
2	E	K	3.2	8
1	E	K	3.27	9
10	E	C	3.5	10
18	I	T	3.57	11
16	I	T	4.18	12
19	I	T	4.38	13
20	I	K	4.6	14
15	I	K	4.92	15
14	I	K	5.27	16
11	I	K	5.75	17
17	I	T	7.2	18
13	I	T	8.17	19
12	I	K	10.25	20

However, when the total number of co-reference chains is controlled for the number of words or the number of NPs, there is no significant difference according to the

Wilcoxon Rank Sums test ($z = 1.41$, NS [$p = .1586$], and $z = -.03$, NS [$p = .9760$], respectively). This implies that the number of co-reference chains varies directly with the length of the discourse or the number of referents.

Table 30: Number of Co-Reference Chains in Student vs. Native Speaker Essays,
Wilcoxon Rank Sums Test Results

	z	p	<i>significant difference?</i>
<i>total # of co-reference chains</i>	-2.03	.0424	yes
<i># of co-reference chains per total # words in essay</i>	1.41	NS [$p = .1586$]	no
<i># of co-reference chains per total # of NPs in essay</i>	-.03	NS [$p = .9760$]	no

Figure 23: Comparison of Student Essay Trends in Length

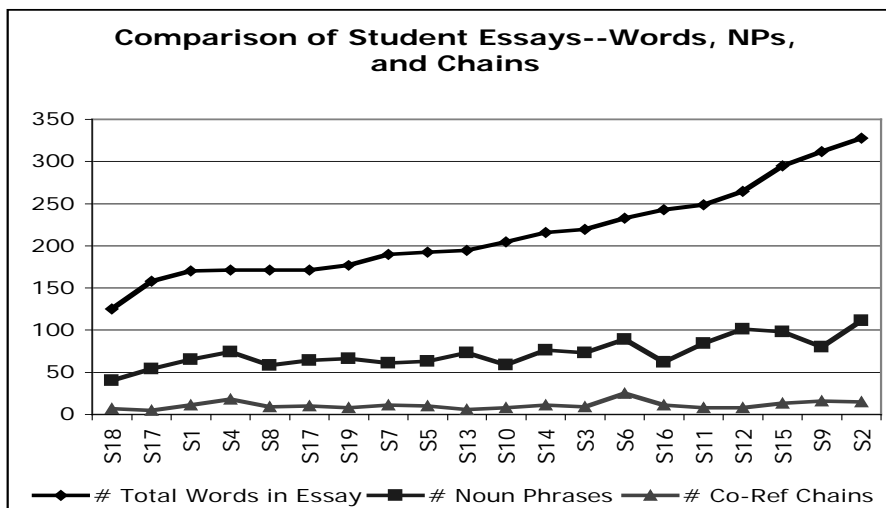
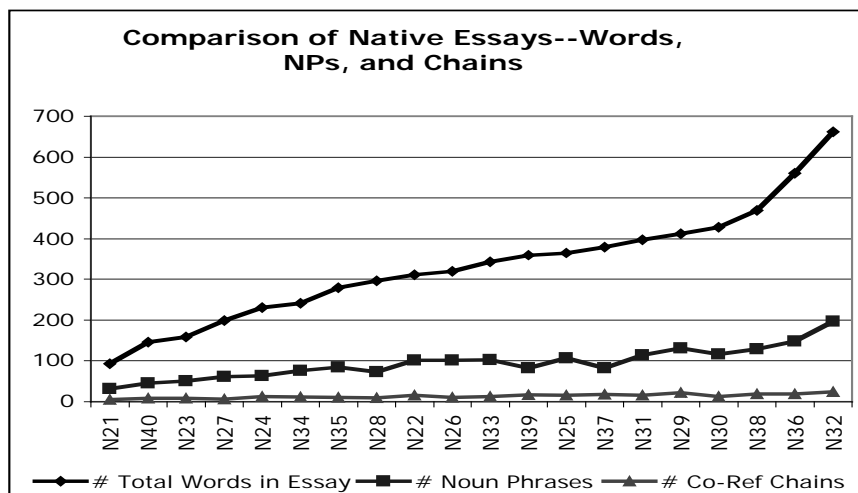


Figure 24: Comparison of Native Essay Trends in Length



The graphs above show this. Although the absolute lengths of the essays are shorter for the students, the relative amounts of words, NPs, and chains are similar. Furthermore, there is no marked difference in the incline shown on the graphs for these three categories.

When comparing the IIEP and EPE students who are at different proficiency levels, the significant difference that was found between the absolute number of co-reference chains is retained when the number of chains is normalized against either the total number of words or number of NPs. See the table that follows. This signifies that there are significant differences in the number of co-reference chains that cannot be attributed to discourse length differences. Looking at the averages and the medians of these three values for the two groups, it seems that the EPE students vary more widely in terms of the total number of words and the number of co-reference chains.

Table 31: Number of Co-Reference Chains in IIEP vs. EPE Students' Essays, Wilcoxon Rank Sums Test Results

	<i>z</i>	<i>p</i>	<i>significant difference?</i>
<i>total # of co-reference chains</i>	-1.89	.0588	yes
<i># of co-reference chains per total # words in essay</i>	-2.12	.0340	yes
<i># of co-reference chains per total # of NPs in essay</i>	-2.27	.0238	yes

The EPE students also had more words and more co-reference chains than the IIEP students, but the average and median number of NPs for the two groups is almost the same. This demonstrates that there is a difference in the number of NPs that are involved in constructing coherent discourse for the two proficiency levels. So, proficiency level is associated with using fewer NPs in co-reference chains, but perhaps having a higher concentration of NPs overall.

Table 32: Average and Median Total Number of Words, NPs, and Co-reference Chains per Essay, IIEP vs. EPE

	<i>IIEP Students</i>	<i>EPE Students</i>
<i>Average # words</i>	209.4	219.3
<i>Median # words</i>	205.5	199
<i>Average # NPs</i>	71.8	73.3
<i>Median # NPs</i>	69.5	69
<i>Average # chains</i>	8.7	13.2
<i>Median # chains</i>	8	11

Although proficiency level is associated with differences in co-reference chains, no significant difference between the two L1 groups for the absolute number of co-reference chains or when the co-reference chains were normalized for the number of words or noun phrases.

Table 33: Number of Co-Reference Chains in Korean vs. Chinese/Taiwanese Students' Essays, Wilcoxon Rank Sums Test Results

	<i>z</i>	<i>p</i>	<i>significant difference?</i>
<i>total # of co-reference chains</i>	.98	NS [p = .3270]	no
<i># of co-reference chains per total # of words in essay</i>	.08	NS [p = .9362]	no
<i># of co-reference chains per total # of NPs in essay</i>	-.83	NS [p = .4066]	no

In conclusion, the most significant differences in the use of co-reference chains were found when comparing the IIEP, or lower proficiency, group and the EPE, or higher proficiency, group. Little difference was detected between the Chinese/Taiwanese and the Korean language learning groups. Furthermore, although there was some significant difference seen in the comparison of the student and native writers in terms of the absolute number of co-reference chains (not adjusted for length of essay or number of NPs), the similarities between the two groups were more striking. Both student and native writers showed patterns in which some NPs were referred to at the beginning and the end of the essay, with different references chained in the middle. Furthermore, for both the L1 and L2 writers, a pattern of having many two member chains with two or three much longer co-reference chains was very common.

The next section will look at the differences in the total number of different kinds of transitions in the essays.

Transitions

In the Simple Transition theory, NPs are categorized as new when they are first mentioned, maintained when they appeared in the previous sentence, and re-mentioned when they have been referred to before, but earlier in the discourse. Looking at transitions is helpful for understanding the use of lexical forms in co-reference chains; however, it is also useful for examining overall how connected the discourse is. This section will address this second issue first by comparing student and native writers, and then by comparing the transitions of the student proficiency and L1 sub-groups. Please note that the terms “transition” and “referent” will be used interchangeably throughout this section and the rest of the paper. Therefore, the phrases “new transition” and “new referent” will be considered equivalent.

New Referents or Transitions

A new referent is one that appears in the discourse for the first time. In this case, “new” refers to the real-world referent and not the specific lexical form. Since co-reference was marked before the transitions were classified, new referents end up being the first appearance of a referent or lexical form in the discourse that is referred to later, or else the only appearance of a particular referent or lexical form. Recall, however, that the co-reference tagging scheme is not sensitive to mark all implied connections between referents and NPs. Therefore, fewer maintained or re-mentioned referents will be detected by the transition cataloguing, and more new referents will be labeled. This over-estimation of the number of new referents is consistent across all essays, though.

Native vs. Student Writers, Number of New Referents

The table that follows shows the number of new referents for each essay.

Table 34: Total Number of New Referents per Essay, Natives vs. Students

<i>ESL Students</i>		<i>Native English Speakers</i>	
<i>participant #</i>	<i>#</i>	<i>participant #</i>	<i>#</i>
1	40	21	21
2	72	22	58
3	57	23	27
4	41	24	38
5	45	25	72
6	40	26	58
7	37	27	43
8	43	28	55
9	49	29	90
10	39	30	64
11	46	31	77
12	27	32	130
13	30	33	64
14	30	34	50
15	47	35	62
16	27	36	92
17	22	37	47
18	22	38	79
19	39	39	49
20	28	40	29

Native writers produced both the smallest and the largest number of new referents. Native writer 21 had the least of all the essays, with only 21 new referents. This is only one less than the lowest number of new referents that a student writer produced. Students 17 and 18 both had only 22 new referents. The largest number of new referents was 130, which was produced by Native 32. This was much more than the

student with the most new referents, though. Student 2 had the most new NPs for a student writer at 72, or 58 fewer new referents than the native writer with the most.

This implies that the native writers varied more in the number of new referents they included in their essays, but also that they tended to have more new referents. When looking at the average and the median number of new referents for these two groups, this is borne out. The table immediately following summarizes these details.

Table 35: Average and Median Total Number of New Transitions per Essay, Native vs. Student Writers

	<i>Students</i>	<i>Natives</i>
<i>Shortest</i>	22, Student #17, #18	21, Native #21
<i>Longest</i>	72, Student #2	130, Native #32
<i>Average #</i>	39.05	60.25
<i>Median #</i>	39.5	58
<i>Wilcoxon Rank Sums Test</i> ($\alpha = .10$)	$z = -3.16$ [$p = .0016$]	

As can be seen, the students' mean and median are very close. The students averaged 39 new referents per essay, which is about 20 fewer new referents than the average or mean number in the native writers' essays. For the natives, the average number of new referents was 60.25, and the median was 58. When the Wilcoxon Rank Sums Test was used to compare the groups, they were found to be significantly different. When the absolute numbers of new referents produced by the native and student writers are compared, the natives are shown to introduce significantly more new referents.

As has been mentioned many times, though, the length of the discourse (number of words and number of NPs) influences this conclusion. The natives tended to write

longer essays, meaning that it is likely that their essays contain significantly more new referents. To account for this, the ratio of the number of new referents to the total number of words was examined. When this is done, there is no significant difference between the student and native writers production of new referents ($z = -.08$, NS [$p = .9362$]). The following table gives the ratios and their ranks. In the next section, the sub-groups of students will be compared.

IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese
Students, Number of New Referents

IIEP vs. EPE Students

As was true for other discourse construction features, a significant difference was found between the proficiency levels but not between the native language families when considering the number of new referents. Students 17 and 18 were the IIEP students with the least number of new referents, having only 22, the least of any student writers. The EPE student with the fewest new referents had 37 and was Student 7. This is 15 more than the IIEP students with the fewest. The IIEP student with the highest number of new referents still had fewer than that EPE student with the fewest new referents.

Student 15 was the IIEP student with the most new referents at 47, but this is 25 fewer than the EPE student with the most new referents, which was Student 2 with 72 new referents. As can be seen in Table 37, the average and median number of new referents was also higher for the EPE than the IIEP students by about 12 referents. In fact, this difference is statistically significant ($z = -2.57$, $p = .0102$, $\alpha = .10$).

Table 36: Ranks of Total Number of New Referents per Total Number of Words, Native vs. Student Writers

<i>participant #</i>	<i>participant category</i>	<i>total # new referents</i>	<i>total # words</i>	<i>total # referents/ total # words</i>	<i>rank</i>
12	IK	27	265	.1019	1
16	IT	27	243	.1111	2
37	N	47	380	.1237	3
39	N	49	359	.1365	4
14	IK	30	216	.1389	5
17	IT	22	158	.1392	6
30	N	64	428	.1495	7
13	IT	30	195	.1538	8
9	EC	49	312	.1571	9
15	IK	47	295	.1593	10
20	IK	28	171	.1637	11
36	N	92	560	.1643	12
24	N	38	231	.1645	13
38	N	79	469	.1684	14
23	N	27	159	.1698	15
6	EC	40	233	.1717	16
18	IT	22	125	.1760	17
26	N	58	319	.1818	18
11	IK	46	249	.1847	19
28	N	55	296	.1858	20
22	N	58	311	.1865	21
33	N	64	343	.1866	22
10	EC	39	205	.1902	23
31	N	77	398	.1935	24
7	EC	37	190	.1947	25
32	N	130	662	.1964	26
25	N	72	365	.1973	27

Table 36—Continued

<i>participant #</i>	<i>participant category</i>	<i>total # new referents</i>	<i>total # words</i>	<i>total # referents/ total # words</i>	<i>rank</i>
40	N	29	146	.1986	28
34	N	50	241	.2075	29
27	N	43	199	.2161	30
29	N	90	412	.2184	31
2	EK	72	328	.2195	32
19	IT	39	177	.2203	33
35	N	62	279	.2222	34
21	N	21	93	.2258	35
5	EK	45	193	.2332	36
1	EK	40	170	.2353	37
4	EK	41	171	.2398	38
8	ET	43	171	.2515	39
3	EK	57	220	.2591	40
$z = -.08, NS (\alpha = .10) [p = .9362]$					

When the number of new referents is divided by the total number of words, the difference between the IIEP and EPE students is still significant ($z = -2.26, p = .0238$), despite the fact that there is no significant difference between the numbers of words produced by these groups ($z = -.08, p = .9362$). The IIEP students tend to have fewer new referents per the number of words in their essays than the EPE students.

Since no significant difference was found in the number of noun phrases or number of words when these groups were compared, the fact that the IIEP students had significantly fewer new referents indicates that they made more connections between their referents than the EPE students did. In fact, a significant difference was also found

between the number of co-reference chains in the IIEP and the EPE essays (when controlled for the number of words, $z = -2.27$, $p = .0238$).

However, the average and the median number of chains per essay is *lower* for the IIEP essays (average 8.7, median 8) than for the EPE essays (average 13.2, median 11). This means that the IIEP students tended to have more referents per chain, while the EPE students had more chains with fewer referents in each chain.

Table 37: Average and Median Total Number of New Referents per Essay, IIEP vs. EPE and Korea vs. China/Taiwan

	<i>IIEP students</i>	<i>EPE students</i>	<i>Korean Students</i>	<i>Chinese/Taiwanese</i>
<i>Most new referents</i>	47, Student #15	72, Student #2	72, Student #2	49, Student #9
<i>Least new referents</i>	22, Student #17, #18	37, Student #7	27, Student #12	22, Student #17
<i>Average # new</i>	31.8	46.3	43.3	34.8
<i>Median # new</i>	29	42	43	38
<i>Wilcoxon Rank Sum ($\alpha = .10$)</i>	$z = -2.57$ [$p = .0102$]		$z = 1.59$, NS [$p = .1142$]	

Table 38: IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test

	z	p	<i>significant difference? ($\alpha = .10$)</i>
<i>total # words</i>	-.08	NS [.9362]	no
<i>total # new referents</i>	-2.57	.0102	yes
<i>ratio of # new referents per total # words</i>	-2.26	.0238	yes

This difference could be an artifact of the essay topics the groups had to choose from. The essay topics that the IIEP students chose are in the table that follows. These topics all encourage a connected story or a narrative with a limited number of main referents or with one main referent.

Table 39: IIEP Essay Topics

<i>Topic title</i>	<i>Description of topic</i>	<i>Number of learners who selected it</i>
being someone else	explanation of how the student's life would be different if they were someone else	3
emotional event	description of an emotional event in the student's life	3
overcoming challenge	description of how the student faced a challenge or obstacle in his/her life	1
most important possession	description of the student's most important possession and explanation of why	1
ceremony or tradition	description of a special ceremony or important tradition in the student's country	2

For example, describing an emotional event in one's life would include the individual him or herself as the main referent and then a limited number of others involved in the activities. Describing a ceremony or tradition would also have one referent that would be linked across a coherent essay, which would be the name of the day or ceremony, or the kind of tradition.

The main difference is that the IIEP topics can be fully explored with a smaller number of referents than required for the EPE essay topics. For the topic of

globalization, for example, an EPE student might have to discuss several ceremonies or traditions that had been affected in order to give sufficient evidence of culture borrowing.

When describing the population graph for the other EPE topic, the number of distinct referents needed would also be potentially quite large. If 30 countries are listed in the graph, then introducing close to 30 separate new referents would be necessary for an intelligible explanation of the point of the graph. The EPE essay topics are presented in the table that follows.

Table 40: EPE Essay Topics

<i>Topic title</i>	<i>Description of topic</i>	<i>Number of learners who selected it</i>
globalization	advantages and disadvantages of culture borrowing and globalization	3
population	description of a graph of past and projected population growth in the developed and developing world	7

The question regarding the essay topics then becomes: Did the test designers choose more contained topics for the IIEP students because it was clear that they would not be able to handle the vocabulary needed for the more complex topics? Were the EPE test topics constructed to require more examples and more referents in order to see if the higher level students could coherently construct a more complicated discourse? This is probably the case. It is not clear what differences would be found if the proficiency levels had written on more similar topics. The native writers' essays, though, contain a mix of personal anecdotes or narratives like those required for the IIEP essays, and a variety of examples, like what was needed for the EPE essays. Therefore, it is interesting to see the similarities and differences between the discourse construction of the student and native writers.

Korean vs. Chinese/Taiwanese Essays

Student 2 was the Korean students with the most new referents, having 72. This is 23 more than the Chinese or Taiwanese student with the most new referents, who was Student 9 with 49 new NPs. The Korean student with the fewest new referents also has more new referents than the Chinese or Taiwanese student with the fewest referents, but the difference is not as great. Student 12 is the Korean student with the least, having only 27 new NPs. Two Taiwanese students tied for the fewest number of new referents for the Chinese/Taiwanese group. These were Students 17 and 18 who had only 22 new referents, five fewer than Korean Student 12. However, when a statistical analysis is done to contrast the two L1 groups, there is no statistically significant difference between the number of new referents either considered by themselves ($z = 1.59, p = .1142$) or in ratio with the number of words per essay ($z = -.83, p = .4066$). It is unclear whether the difference in essay topics had any effect on this outcome, but for most of the discourse features examined, there has been little significant difference between the English L2 writing of the two L1 groups.

Table 41: Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Wilcoxon Rank Sums Test

	z	p	<i>significant difference? ($\alpha = .10$)</i>
<i>total # words</i>	-.08	NS [.9362]	no
<i>total # new referents</i>	1.59	NS [.1142]	no
<i>ratio of # new referents per total # words</i>	-.83	NS [.4066]	no

In the next section, the student and native groups and the student sub-groups will be compared for differences in the use of maintained referents. Since a significant difference was found between the number of new referents in the native and student essays, and between the IIEP and EPE student sub-groups, the prediction is that there should also be a difference in the number of maintained referents in the essays of these groups.

Maintained Referents or Transitions

For the purpose of this study, a maintained referent is one that was referred to in the immediately preceding sentence. Therefore, unlike a new referent, these referents are guaranteed to be activated in short term memory. This means that the set of possible lexical forms for maintained referents is different from that of new referents. Re-mentioned referents are similar to maintained referents in that they have appeared before in the discourse. The distinction between these lies in the fact that re-mentioned referents do not appear in the immediately preceding sentence, so they do not have the same level of activation in memory.

Native vs. Student Writers, Number of Maintained Referents

The table that follows on the next page compares the number of maintained referents in the student and native essays. As can be seen in the table, the smallest number of maintained referents overall was produced by native writers. Natives 21 and 28 had only nine NPs that continued reference from the previous sentence. The smallest number of maintained referents made by a student, though, is only one more than what Natives 21 and 28 did. Student 3 had only ten maintained referents.

Possibly against expectation, the greatest number of maintained referents overall were made by a student. Student 12 had 57 maintained referents, which is nine more than the number made by the native speaker with the most maintained referents. Native 36

had only 48 maintained referents. So, it can be seen that two native writers tied for having the fewest maintained NPs, while a student writer had the most. Does this indicate that students tended to produce more maintained referents on the whole? When looking at the average and median number of maintained NPs for the two groups, this tendency cannot be confirmed.

Table 42: Total Number of Maintained Referents per Essay

<i>ESL Students</i>		<i>Native English Speakers</i>	
<i>participant #</i>	<i># maintained referents</i>	<i>participant #</i>	<i># maintained referents</i>
1	20	21	9
2	22	22	37
3	10	23	22
4	16	24	14
5	12	25	25
6	29	26	35
7	20	27	10
8	12	28	9
9	22	29	25
10	13	30	43
11	28	31	20
12	57	32	42
13	36	33	25
14	40	34	16
15	39	35	17
16	29	36	48
17	28	37	19
18	12	38	31
19	22	39	23
20	30	40	12

The average and the median number of maintained referents differ by at most .75, indicating that the groups as a whole do not greatly differ. See the table that follows for the averages and medians. When the Wilcoxon Rank Sums Test is used to test the significance of the difference between the student and native groups, none is found ($z = .08$, NS [$p = .9362$]). There is no indication that student and native writers are using more or fewer maintained referents than the other group.

Table 43: Average and Median Total Number of Maintained Referents per Essay,
Student vs. Native Writers

	<i>Students</i>	<i>Natives</i>
<i>Shortest # maintained referents</i>	10, Student #3	9, Natives #21, #28
<i>Longest # maintained referents</i>	57, Student #12	48, Native #36
<i>Average # maintained referents</i>	24.85	24.1
<i>Median # maintained referents</i>	22	22.5
<i>Wilcoxon Rank Sums Test ($\alpha = .10$)</i>	$z = .08$, NS [$p = .9362$]	

Because of the ubiquitous problem of differences in number of words affecting numbers of NPs or kinds of NPs, the ratio of the number of maintained referents to the total number of words per essay was also examined. When these ratios are ranked and compared, there is in fact shown to be a significant difference ($z = 2.94$, $p = .0032$) between the use of maintained referents by students and by native writers. From looking at the ranks in the table that follows, it can be seen that the students have more maintained referents per total number of words in their essays.

This difference in maintained referents is interesting because as discussed previously, there was no significant difference between the number of co-reference chains per the total number of words of the student and native writers ($z = 1.41$, $p =$

.1586). All maintained referents must be part of co-reference chains. Therefore, the fact that the number of the *kind* of transition in the co-reference chain is significantly different, while the number of chains was not significantly different means that while both groups may be similar in how many chains they have, how they are constructed is likely to be different. It was also shown that the number of new referents per the total number of words was not significantly different for the student and native groups ($z = -.08, p = .9362$). Thus, there seems to be a difference so far only in how often new referents introduced in discourse are repeated in the next sentence given variations in the length of the essays. This difference could result from many sources. One possibility is that the students are less able to or do not vary their linked references semantically. As discussed, not all of the connections between native writer NPs could be classified by the co-reference tagging rules.

If a native writer had a part-whole relationship, or category-example relationship between two NPs, this would not be marked with this system even though it is the maintenance of reference in some way and may be seen as more desirable or elegant. For example, if a writer is discussing colleges, and names of colleges are used, there is some kind of connection thematically in the discourse, but this relationship would not be marked. In contrast, repeating the exact same vocabulary term would be detected as a maintained reference, while being considered possibly as odd or a simplistic writing style. In this manner, students with less variety of expression and word choice could be found to have more maintained referents.

Differences in essay topic could also again be causing this difference in the use of maintained referents. If a particular topic requires one to talk about a larger pool of referents, there may be fewer maintained referents because moving on to another example may be more important than explaining ideas related to a particular referent.

Table 44: Ranks of Total Number of Maintained Referents per Total Number of Words,
Native vs. Student Writers

<i>participant #</i>	<i>participant category</i>	<i>total # maintained referents</i>	<i>total # words</i>	<i>total # maintained referents/ total # words</i>	<i>rank</i>
28	Native	9	296	.0304	1
3	Student	10	220	.0454	2
37	Native	19	380	.0500	3
27	Native	10	199	.0503	4
31	Native	20	398	.0503	5
24	Native	14	231	.0606	6
29	Native	25	412	.0607	7
35	Native	17	279	.0609	8
5	Student	12	193	.0622	9
10	Student	13	205	.0634	10
32	Native	42	662	.0634	11
39	Native	23	359	.0641	12
38	Native	31	469	.0661	13
34	Native	16	241	.0664	14
2	Student	22	328	.0671	15
25	Native	25	365	.0685	16
8	Student	12	171	.0702	17
9	Student	22	312	.0705	18
33	Native	25	343	.0729	19
40	Native	12	146	.0822	20
36	Native	48	560	.0857	21
4	Student	16	171	.0936	22
18	Student	12	125	.0960	23
21	Native	9	93	.0968	24
30	Native	43	428	.1005	25
7	Student	20	190	.1053	26
26	Native	35	319	.1097	27
11	Student	28	249	.1124	28

Table 44—Continued

<i>participant #</i>	<i>participant category</i>	<i>total # maintained referents</i>	<i>total # words</i>	<i>total # maintained referents/ total # words</i>	<i>rank</i>
1	Student	20	170	.1176	29
22	Native	37	311	.1190	30
16	Student	29	243	.1193	31
19	Student	22	177	.1243	32
6	Student	29	233	.1245	33
15	Student	39	295	.1322	34
23	Native	22	159	.1384	35
20	Student	30	171	.1754	36
17	Student	28	158	.1772	37
13	Student	36	195	.1846	38
14	Student	40	216	.1852	39
12	Student	57	265	.2151	40
$z = 2.94$, significant ($\alpha = .10$) [$p = .0032$]					

Since the number of new NPs per total number of words in the student and native essays was not statistically different, the issue seems to be that the students (when considered all together) are only maintaining referents different and not introducing fewer referents.

Recall that there may be differences between the sub-groups of students, as there was a significant difference between the proficiency levels, but not the L1 groups, in the use of new referents. The next section will look again at these sub-groups.

IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese
Students, Number of Maintained Referents

IIEP vs. EPE Students, Maintained Referents

An EPE student had the fewest maintained referents (Student 3, 10 maintained referents) while an IIEP student had the most maintained referents (Student 12, 57 maintained referents) when looking at the entire group of students. For just the IIEP students, the one student with the fewest maintained NPs did not differ much from the student with the lowest number. IIEP Student 18 had only 12 maintained referents, only two more than EPE Student 3. There is a much greater difference between the IIEP student with the most referents for that group, though, and the EPE student with the most maintained NPs for their group. EPE Student 6 had the most maintained NPs among the EPE students at 29, but this is 28 fewer than the IIEP student with the most. Looking at the averages and the medians for the two groups, they indicate that the IIEP students had more maintained referents than the EPE students. The IIEP average was 32.1 maintained NPs and the IIEP median was 29.5 maintained referents. These are both over 10 more than the average and median number of maintained referents in the EPE essays, which are 17.6 and 18, respectively. Furthermore, this difference was found to be statistically significant ($z = 3.02$, $p = .0026$), as is the difference when the number of maintained referents is divided by the number of words in each essay ($z = 3.25$, $p = .0012$). These details are summarized in the table that follows and in one further on below.

So, IIEP students have significantly more maintained referents than EPE students when their averages are compared. The lower proficiency level students have on average *more* maintained referents.

Table 45: Average and Median Total Number of Maintained Referents per Essay, IIEP vs. EPE and Korea vs. China/Taiwan

	<i>IIEP students</i>	<i>EPE students</i>	<i>Korean Students</i>	<i>Chinese/Taiwanese</i>
<i>Longest # maintained referents</i>	57, Student 12	29, Student 6	57, Student 12	36, Student 13
<i>Shortest # maintained referents</i>	12, Student 18	10, Student 3	10, Student 3	12, Student 8
<i>Average #</i>	32.1	17.6	27.4	22.3
<i>Median #</i>	29.5	18	25	22
<i>Wilcoxon Rank Sum ($\alpha = .10$)</i>	$z = 3.02$ [$p = .0026$]		$z = .38$, NS [$p = .7040$]	

Table 46: Selected Averages for Statistically Significant Differences between IIEP and EPE Students

<i>Discourse feature</i>	<i>IIEP vs. EPE students</i>	
	<i>IIEP average</i>	<i>EPE average</i>
<i>average co-ref chain length</i>	5.83*	3.01
<i>total # co-ref chains</i>	8.7	13.2*
<i>ratio # co-ref chains/ # words</i>	.0425	.0617*
<i>ratio # co-ref chains/ # NPs</i>	.1257	.1782*
<i>total # new transitions</i>	31.8	46.3*
<i>ratio # new transitions/ # words</i>	.1549	.2152*
<i># maintained transitions</i>	32.1*	17.6
<i>ratio # maint. transitions/ # words</i>	.1522*	.0820

*Indicates the higher average.

However, because they also have *fewer* new referents—but no statistically significant difference in the number of NPs overall, it can be said that the IIEP students are creating

more connections between a smaller set of real-world elements. What effect this may have on article and determiner use in chained NPs will be examined later. The table immediately preceding summarizes the averages of the proficiency levels for the statistically significant differences. The next issue is whether there are differences in maintained transitions between the Korean and Chinese/ Taiwanese students.

Table 47: IIEP Essays as Compared to EPE Students' Essays, Maintained Referents, Wilcoxon Rank Sums Test

	<i>z</i>	<i>p</i>	<i>significant difference?</i> ($\alpha = .10$)
<i>total # words</i>	-.08	NS [.9362]	no
<i>total # maintained referents</i>	3.02	.0026	yes
<i>ratio of # maintained referents per total # words</i>	3.25	.0012	yes

Korean vs. Chinese/Taiwanese Students

The Korean and Chinese/ Taiwanese students are similar in the lower boundary of the number of maintained referents, but appear to differ greatly in the upper boundary. Korean Student 3 had the fewest maintained referents of all the Koreans, with only ten. This is only two fewer than the Taiwanese student who had the fewest maintained referents of the other L1 group, which was Student 8, who had 12 maintained referents. The difference between the Korean and Chinese/ Taiwanese students with the most maintained referents is ten times this difference at the lower boundary. Korean Student 12's essay contains 57 maintained referents, which is 21 more than Taiwanese Student 13 who had 36 maintained referents. Even though this upper extreme is quite different, though, overall there was no significant difference ($z = .38$, NS [$p = .7040$]) found

between the number of maintained referents in the essays of the two groups. These details are summarized in a previous table. When the number of maintained referents is normalized for the number of words per essay, there is still no significant difference between the language groups ($z = 0, p = 1$), as can be seen in the table below. As with many other of the discourse relevant features of these essays, there is a difference between the student proficiency levels, but not between the two language groups.

Table 48: Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Maintained Referents, Wilcoxon Rank Sums Test

	z	p	<i>significant difference?</i> ($\alpha = .10$)
<i>total # words</i>	.98	NS [.3720]	no
<i>total # maintained referents</i>	.38	NS [.7040]	no
<i>ratio of # maintained referents per total # words</i>	0	NS [1]	no

So far, in looking at the transitions, there has been found a difference between the native and student writers in:

- The number of new transitions, but only the absolute number and not when normalized versus the number of words in an essay, and in
- The number of maintained transitions, but only when considered in ration with the number of words per essay.

When examining the students in more detail by looking at the IIEP/lower proficiency and EPE/higher proficiency sub-groups, significant differences were found in:

- The absolute number of new transitions,
- The ratio of the number of new transitions to the number of words,

- The absolute number of maintained transitions, and
- The ratio of the number of transitions to the number of words.

Table 49: Total Number of Re-mentioned Referents per Essay, Native vs. Student Writers

<i>ESL Students</i>		<i>Native English Speakers</i>	
<i>participant #</i>	<i># re-mentioned referents</i>	<i>participant #</i>	<i># re-mentioned referents</i>
1	5	21	2
2	16	22	7
3	6	23	2
4	17	24	11
5	6	25	10
6	20	26	8
7	3	27	8
8	3	28	9
9	9	29	16
10	7	30	9
11	10	31	17
12	17	32	25
13	7	33	14
14	6	34	10
15	12	35	6
16	6	36	8
17	4	37	16
18	6	38	19
19	5	39	10
20	6	40	4
$z = -1.68, p = .0950, \text{significant}$			

No significant differences have been found between the Korean and the Chinese/Taiwanese sub-groups in the use of transitions. There is one more kind of simple transition to be examined, the re-mentioned transition, which is investigated next.

Re-mentioned Referents or Transitions

As described earlier, new transitions are the first appearance of a nominal referent, and maintained transitions are NPs for which another co-referring NP appears in the previous sentence. Re-mentioned referents are then NPs which are subsequent mentions of a real-world object already talked about in the discourse, but for whom there is *no* co-referring NP in the previous sentence. More than one sentence intervenes between the two co-referring NPs.

Native vs. Student Writers, Number of Re-mentioned

Referents

The table that is immediately following shows the number of re-mentioned referents produced by each student and native writer. As the table shows, the range of the number of re-mentioned transitions for the student and native writer groups does not appear to be that different. Natives 21 and 23 have only two re-mentions, which is the fewest among the native writers. This is only one less than the three re-mentions produced by both Students 7 and 8, who at three re-mentions have the fewest for the students. The range for both groups is about a difference of about 20 referents. The native writer with the most re-mentioned referents is Native 32, whose essay has 25 re-mentioned transitions. Student 6 has the most re-mentions for the student group, with 20. The difference between the two groups is the distribution over this range. Ten native speakers have a number of re-mentions in the double digits, while only six students do. Moreover, the average and the median number of re-mentions is higher for the native writers. This difference is in fact statistically significant ($z = -1.68$, $p = .0950$, $\alpha = .10$). This is summarized in the previous table.

Table 50: Average and Median Total Number of Re-mentioned Referents per Essay,
Native vs. Student Writers

	<i>Students</i>	<i>Natives</i>
<i>Shortest # re-mentioned referents</i>	3, Students #7, #8	2, Natives #21, #23
<i>Longest # re-mentioned referents</i>	20, Student #6	25, Native #32
<i>Average # re-mentioned referents</i>	8.55	10.55
<i>Median # re-mentioned referents</i>	6	9.5
<i>Wilcoxon Rank Sums Test ($\alpha = .10$)</i>	$z = -1.68, p = .0950$	

As was seen when looking at the number of new transitions and the total number of co-reference chains, though, there is a statistically significant difference between the groups only when the absolute number of re-mentions is tested. The native writers have significantly more total words per essay and total NPs per essay than the student writers. If they wrote more, then it is to be expected that the native writers would have more re-mentioned transitions. When the number of re-mentions is controlled by dividing them by the number of words per essay, there is no significant difference between the number of re-mentions used by the student versus the native writers ($z = .95, p = .3422$). The ratios for each student and native writer are given in the next table.

Next the difference between the student proficiency and L1 sub-groups will be investigated.

IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese

Students, Number of Re-mentioned Referents

This study has already looked at the number of new and maintained referents for the two student sub-groups.

Table 51: Ranks of Total Number of Re-mentioned Referents per Total Number of Words

<i>participant #</i>	<i>participant category</i>	<i>total # re-ment. referents</i>	<i>total # words</i>	<i>total # re-ment. referents/ total # words</i>	<i>rank</i>
23	Native	2	159	.0126	1
36	Native	8	560	.0143	2
7	Student	3	190	.0158	3
8	Student	3	171	.0175	4
30	Native	9	428	.0210	5
21	Native	2	93	.0215	6
35	Native	6	279	.0215	7
22	Native	7	311	.0225	8
16	Student	6	243	.0247	9
26	Native	8	319	.0251	10
17	Student	4	158	.0253	11
3	Student	6	220	.0273	12
25	Native	10	365	.0274	13
40	Native	4	146	.0274	14
14	Student	6	216	.0278	15
39	Native	10	359	.0279	16
19	Student	5	177	.0282	17
9	Student	9	312	.0288	18
1	Student	5	170	.0294	19
28	Native	9	296	.0304	20
5	Student	6	193	.0311	21
10	Student	7	205	.0341	22
20	Student	6	171	.0351	23
13	Student	7	195	.0359	24
32	Native	25	662	.0378	25
29	Native	16	412	.0388	26
11	Student	10	249	.0402	27
27	Native	8	199	.0402	28

Table 51—Continued

<i>participant #</i>	<i>participant category</i>	<i>total # re-ment. referents</i>	<i>total # words</i>	<i>total # re-ment. referents/ total # words</i>	<i>rank</i>
38	Native	19	469	.0405	29
15	Student	12	295	.0407	30
33	Native	14	343	.0408	31
34	Native	10	241	.0415	32
37	Native	16	380	.0421	33
31	Native	17	398	.0427	34
24	Native	11	231	.0476	35
18	Student	6	125	.0480	36
2	Student	16	328	.0488	37
12	Student	17	265	.0642	38
6	Student	20	233	.0858	39
4	Student	17	171	.0994	40
$z = .95$, NS [$p = .3422$] ($\alpha = .10$)					

In fact, there was a significant difference found between the IIEP and EPE students for the absolute number of new and maintained transitions, as well as for the number when controlled per the number of words. No significant differences were found between the Korean and Chinese/Taiwanese sub-groups. For re-mentioned transitions then, one might predict a significant difference also to be found when comparing the IIEP and EPE students, but not when comparing the Korean and Chinese/ Taiwanese students. This prediction, however, is not fully correct. There are no significant differences between either sub-group in terms of re-mentioned transitions.

IIEP vs. EPE Students

The EPE/ higher proficiency students have a slightly larger range in the number of re-mentions, spanning from a low of three (Students 7 and 8) to a high of 20 (Student

6). The IIEP/ lower proficiency students have a smaller range by four, with a low of four re-mentions (Student 17) and a high of 17 (Student 12). Despite this small difference in the range, and the fact that the EPE students have a slightly higher mean and median number of re-mentions, there is no statistically significant difference between the IIEP and EPE students' production of re-mentioned referents ($z = .30, p = .7642$). This is summarized in the subsequent table.

Table 52: Average and Median Total Number of Re-mentioned Referents per Essay, IIEP vs. EPE and Korea vs. China/Taiwan

	<i>IIEP students</i>	<i>EPE students</i>	<i>Korean Students</i>	<i>Chinese/Taiwanese</i>
<i>Longest</i>	17, Student #12	20, Student #6	17, Student #12	20, Student #6
<i>Shortest</i>	4, Student #17	3, Students #7, #8	5, Student #1	3, Students #7, #8
<i>Average #</i>	7.9	9.2	10.1	7
<i>Median #</i>	6	6.5	8	6
<i>Wilcoxon Rank Sum</i>	$z = .30, NS [p = .7642] (\alpha = .10)$		$z = 1.20, NS [p = .2302]$	

In fact, there is no significant difference between the number of re-mentions when normalized to the number of words either. Again see the preceding table. This lack of difference is not surprising since there is no significant difference in the number of words in the essays of the IIEP and EPE students. However, it is surprising that there is a difference between the IIEP and EPE students' use of new and maintained transitions, but not re-mentioned transitions. It is also surprising that neither group uses that many re-mentioned transitions (IIEP average is 7.9, EPE average is 9.2) compared to the number of new (IIEP average is 31.8, EPE average is 46.3) and retained (IIEP average is 32.1, EPE average is 17.6) transitions they have.

Table 53: IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test, Re-mentioned Referents

	<i>z</i>	<i>p</i>	<i>significant difference?</i> ($\alpha = .10$)
<i>total # words</i>	-.08	NS [.9362]	no
<i>total # re-mentioned referents</i>	.30	NS [.7642]	no
<i>ratio of # re-mentioned referents per total # words</i>	.23	NS [.8180]	no

The factors limiting the number of re-mentioned referents for both groups are related to the fact that in order to have maintained or re-mentioned transitions, NPs must be linked in a chain of reference:

- Given the number of NPs each individual has, there are really not that many co-reference chains linking NPs. The average number of ratio of chains to NPs is only .1257 for the IIEP students, and .1782 for the EPE students.
- The majority of chains link only two NPs, and these two NPs are not usually very far apart in the discourse. The mode length for all but three of the students was two linked referents. If the NPs are close, then they are not likely to be separated by more than one sentence, meaning the most linked NPs would be new-maintained referent pairs with no re-mention of the object or concept.
- All the essays had one to three very long chains. Because these chains were very long, though, the NPs were spread throughout the discourse and linked the entire pieces together. This means that any two NPs in a long chain are not likely to be very far apart. Again, if the NPs are close, they are not likely to be separated by more than one sentence, and the most common relationship between two NPs is probably maintained-maintained.

To summarize, the patterns of co-reference in these essays imply that re-mentioned referents are not that common nor that important to the discourse. These conclusions make it less surprising that there is no difference between the IIEP and EPE students only in this kind of transition.

Korean vs. Chinese/Taiwanese Students

The range in the number of re-mentioned referents is smaller for the Korean than for the Chinese/ Taiwanese students by five. The range for the Korean students is 12, with Student 1 having the fewest re-mentions (only five), and Student 17 having the most (12 re-mentions). The range for the Chinese/ Taiwanese students is 17, with Students 7 and 8 having the fewest re-mentions (only three), and Student 6 having the most (20 re-mentions). However, more Korean students have more re-mentions than do the Chinese/ Taiwanese students. Both the average and median number of re-mentions is higher for the Korean students. The Korean students' average is 10.1, and their median number of re-mentions is eight. In contrast, the Chinese/ Taiwanese students' average is only seven, with a median value of six. These minor differences, though, are not significant ($z = 1.2$, $p = .2302$). Nor is there any significant difference in the ratio of the number of re-mentioned referents to the number of words ($z = 1.59$, $p = .1118$). See the table that follows. The fact that there is no significant difference in the use of re-mentioned transitions is not surprising since no significant differences have been found between the L1 groups except for the number of NPs.

Conclusions about Transitions

There are many interesting conclusions that can be drawn from the comparisons made between the groups regarding their use of transitions. Some of these are:

- There is no significant difference between the Chinese/ Taiwanese and the Korean groups regarding their production of new, maintained, or re-mentioned referents.

Table 54: Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Wilcoxon Rank Sums Test, Re-mentioned Referents

	<i>z</i>	<i>p</i>	<i>significant difference?</i> ($\alpha = .10$)
<i>total # words</i>	.98	NS [.3720]	no
<i>total # re-mentioned referents</i>	1.21	NS [.2302]	no
<i>ratio of # re-mentioned referents per total # words</i>	1.59	NS [.1118]	no

- Native and student writers differ in the absolute number of their new and re-mentioned transitions, with native writers averaging more of each kind. However, native writers had significantly more NPs and words than the students, and when the number of new and re-mentioned referents are divided by the number of words, there is no longer any significant difference. This indicates that the number of new and re-mentioned referents vary directly with the number of words. Therefore, perhaps the significant difference between the absolute values is not pointing out a major difference between the two groups.
- There was a significant difference found between the student and native writers in the ratio of the number of maintained referents to the total number of words. In fact, the students tended to have more maintained referents per total number of words than the native writers. Since the native writers have more co-reference chains than the students, this indicates that the students are linking their referents more often with maintained transitions than the native writers.
- The EPE and IIEP students varied significantly in their use of new and maintained transitions. The EPE / higher proficiency students tended to have more new referents, which may indicate that they have a wider range of vocabulary to use to talk about a topic or refer to a real-world object or concept. This could also be the

impetus behind the pattern that the IIEP/ lower level students tended to have more maintained transitions than the EPE students. These differences might also have been caused by the difference in the kind of writing topics presented to the two student groups. The IIEP topics required the students to talk about a smaller pool of real-world referents than the EPE students were asked to discuss, so the IIEP students would have had to introduce fewer referents and had the opportunity to link those they introduced more often in an adjacent clause.

The next table summarizes the averages for significant differences between the groups.

Table 55: Summary of Averages for Statistically Significant Differences between Groups

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/ Taiwanese students</i>	
	<i>student average</i>	<i>native average</i>	<i>IIEP average</i>	<i>EPE average</i>	<i>Korean average</i>	<i>Chinese/ Taiwan</i>
<i>total # new transitions</i>	39.05	60.25*	31.8	46.3*		
<i>ratio # new transitions/ # words</i>			.1549	.2152*		
<i># maintained transitions</i>			32.1*	17.6		
<i>ratio # maint. transitions/ # words</i>	.1171*	.0748	.1522*	.0820		
<i># re-mentioned transitions</i>	8.55	10.55*				
<i>ratio # re-ment. transitions/ # words</i>						

The next section will look at the discourse feature of syntactic position. Syntactic position in a clause is important to co-reference and discourse in many ways. For example, syntactic position in subsequent clauses can affect anaphora resolution, and there is a pattern in English of old information being in the subject position and new information being in the object position due to pronunciation and focus stress.

Syntactic Position

The relative syntactic positions of two co-referring NPs can be important in creating discourse cohesion because it can effect anaphora resolution. Consider this example:

Example 57: Syntactic Position Affects Anaphor Meaning Resolution

Ana called Maria on the phone.

She was upset.

The pronoun in the second sentence has two potential antecedents: Ana or Maria. However, without any extra words to force the reading of Maria as the antecedent, the fact that *she* and *Ana* are both in subject position causes their linkage and the second sentence is read as Ana is the woman who is upset. Another reason why syntactic position is important to discourse in English is that the neutral sentence stress position is towards the end of the sentence. This means that objects often are new information and subjects are often old information. Still one more reason why the syntactic position of NPs is important is that it can give indications of how complex the syntactic structures of the writers are. If a group tends to have more NPs as objects of prepositions, then it is possible that they have more longer NPs that have PP complements or adjuncts.

In this section, the number of NPs in a variety of syntactic positions will be examined, and later the positions and forms of NPs in co-reference chains will be investigated.

Classification System for Syntactic Position

The NPs in the essays were classified as being subjects, verbal objects, objects of prepositions, genitive specifiers, complements of copula verbs, part of apposition phrases, in comparison constructions, or in the titles of the essays. Separately categorizing appositions and comparisons was desirable because these are not always accurately produced by non-native speakers, and if the NPs involved were marked distinctly, these constructions could be found and compared more easily.

Subjects

The examples below have the whole subject NPs italicized. Note that two of the subjects consist of an NP + PP. With such a situation, the NP in the PP is classified as the object of a preposition, which is the case with *fat people*. The entire NP *the number of fat people* is the subject, in which the head noun *number* is in bold. In the native example, *core values* is the head noun, but the entire italicized NP is the subject. The embedded NP examples would be listed as appositions. The native example also shows how varied the length of the subject NPs can be. Some are single words, such as *we*, and some are quite long and contain several PPs or relative clauses.

Example 58: NP Subjects in Student and Native Essays

Student 2: In these days in Korea, *the **number** of fat people* is increasing, and *the **reason** of that* must be fatty western food.

Native 32: ***Core values** such as responsibility, acceptance, tolerance and many others* come in part from having the knowledge and understanding which *we* obtained from general education courses.

Verbal Objects

The issues of length and marking the entire NP, not just the head or the complements are also true in the case of verbal objects and the other categories. In the examples below, it can be seen that the object NPs can be in the main clause of a sentence, or a dependent clause. *Multiple disciplines* and *perseverance* are in dependent clauses, while *myself* is in the main clause. Objects are also of course found in tensed and tenseless clauses. *Anything* is in a tenseless clause, but the other object NPs below are in tensed clauses.

Example 59: NP Verbal Objects in Student and Native Essays

Student 17: When I have *perseverance*, I will encourage *myself* to meet *anything*.

Student 24: It is rare that one problem does not span *multiple disciplines*.

Objects of Prepositions

NP objects of prepositions can be embedded in larger NPs or else be part of the verb phrase. In Student 14's example, the PP *for one day* is an adjunct of the verb phrase and *one day* is classified as an NP object of a preposition. In the example from Native 25, the long NP *the general education requirement for proficiency in composition* has two internal PP constituents: *for proficiency in composition* and *in composition*. Therefore, *proficiency in composition* is an NP marked as a PP object, and *composition* is classified the same way. *The general education requirement for proficiency in composition* is a subject.

Example 60: NP Objects of Prepositions in Student and Native Essays

Student 14: So if I am a very rich person *for a one **day***, I will make a association that help them.

Native 25: *The general education requirement for **proficiency in composition** demands that all students, regardless of the **discipline that interests them**, are proficient in constructing a letter, writing a proposal, or pitching a product.*

Genitive Specifiers

There are fewer genitive specifiers than the number of subjects, verbal objects, or objects of prepositions, and the majority of these specifiers are possessive pronouns and not the Saxon genitive as in the examples below. The examples also show two complications in sorting NPs as genitive specifiers. First, what should be done with *own*? The MontyTagger automatic part-of-speech tagging program tags *own* as an adjective, so it was treated this way.

Example 61: NP Genitive Specifiers in Student and Native Essays

Student 1: So many people is meeting by on-line chatting and sharing ***their own** information*, such as ***their** hobby*, foods, and ***their** life style*.

Native 36: For example, ***my master's** in library science program* includes management and technology courses in addition to the traditional research and cataloging courses.

Second, what to do with *master's*? Is it a Saxon genitive with an understood or empty head noun? In this case, *my* was classified as a genitive specifier and *master's* was used as a placeholder for the understood head NP “degree.” Native 36’s example also illustrates the fact that the native speakers’ writing is not error-free. *My master's in*

library science program sounds awkward at best because if *library science program* is a compound noun, then the head singular count noun *program* has no determiner.

Copula Complements

Copula complements are another syntactic position in which students commonly make errors—from omitting the copula, to omitting an article, to selecting an awkward article. For example, in Student 11’s excerpt below, *a dictator* is slightly awkward because typically there is only one dictator in a country at a time, and in this case, *a* implies one out of a set. With the meaning of one dictator out of all of those in the history of Korea, *a* works a little better, but is still questionable in regards to evaluating the truth value of Korea having a dictator since they are currently ostensibly a democracy with a president. Native 30’s excerpt also gives another example of how the natives’ writing is not error-free. There is an incorrect verb form in that *take* is a base form and not a gerund in *was able to completely avoid take*.

Example 62: NP Complements of Copula Verbs in Student and Native Essays

Student 11: I would *be a dictator of my country Korea*, if I could *be anybody* for one day because of the following three main reasons.

Native 30: I was also able to completely avoid take a basic science, which *was a bonus*.

Appositions

Up to this point, the syntactic positions used as classifications have been relatively uncontroversial, but appositions, comparisons, and titles are not usually considered as their own categories. A separate category for appositions was chosen because of examples like that of Student 3 below.

Example 63: NP Appositions in Student and Native Essays

Student 3: Second, the developing countries, *India, Indonesia, Nigeria, Mexico, Saudi Arabia*.

Native 27: If “bar” is narrowly defined as a place that only serves alcoholic beverages *i.e.—no food, no music, no other entertainment and no amusement devices (e.g. pool tables, pinball machines, chess boards)*, it will be circumvented by the addition of minimal food service (as simple as microwave pizza), or a pool table.

In this excerpt, the country names are listed without any overt syntactic architecture such as complementizers, verbs, or phrases like *such as*. For ease of coding, then, a separate apposition category was considered. This course was ultimately chosen because in the MUC-7 co-reference coding rules, how to mark the co-reference of NPs in appositions was considered separately. As there are special reference issues surrounding appositions, they were classed separately. Native 27’s excerpt shows again how in some circumstances only punctuation in writing, such as a dash, a comma, or parentheses, indicates an apposition and there is no complementizer, adverb, or morphological structure used. In speech, the punctuation would ostensibly be pausing or pitch changes.

Comparisons

Comparisons were given their own category mainly because they are multi-word, complicated constructions with word order and function word requirements. One difference between EPE and IIEP students is likely to be skill in writing comparisons accurately, as they are difficult for lower proficiency students. Comparisons are taught in IIEP grammar classes at the UI, but are not seen as a necessary part of the curriculum in the grammar classes in which EPE students are placed. Student 10’s excerpt shows how in one of the essay topics presented to the EPE students, comparisons between many

different countries in terms of their population were necessary. Therefore, there may be more examples of comparisons in the EPE essays.

Example 64: NP Comparisons in Student and Native Essays

Student 10: The population structure of US is suitable to its developing speed, not *as old as Japan and China* who are getting more and more burden, and not *too young, as India*.

Native 39: I think that this would be a major mistake because it would, for many students, turn higher education into *nothing more than a vocational school*.

Titles

Titles were classified separately also for convenience's sake. Very few of either the students or the natives used titles. However, as shown in the two example titles below, the entire title usually consists of just an NP. One difference in the student and native titles was that the student titles were formal and clearly viewed as part of the overall essay, such as the title of a newspaper article. The students used capital letters and centered the titles on their pages.

Example 65: NPs in Titles in Student and Native Essays

Student 7: *Culture Difference*

Native 23: *letter to the editor for retaining general education requirements*

The native speakers, on the other hand, such as Native 23, used informal titles that seemed to function only to indicate the essay topic that was chosen. They did not use

capital letters, and placed titles on their writing mainly when they were letters to the editor. The fact that letters are not a form of writing that require titles shows that they did not see the titles as part of the piece of writing.

Now that examples have been given illustrating what the syntactic categories used mean in this research, a comparison of the frequency of these in the sub-groups' essay will be completed in the following sections.

Native vs. Student Writers, NPs in Various Syntactic Positions

Comparing the Student and Native writers for the absolute numbers of all of these categories, the only statistically significant differences between the two groups were in the number of verbal objects ($z = -2.29, p = .0220$), and the number of objects of prepositions ($z = -2.79, p = .0054$). See the first landscaped table for a summary.

In the cases of both the verbal and prepositional objects, the native writers averaged more NPs in that position than the students. However, again it is important to remember that the native writers had significantly more words and more NPs per essay, which could partially cause this result. In fact, when the number of verbal and prepositional objects are normalized per the number of words per essay, there is no remaining significant difference between the two groups. See the table that follows on the next page.

It is interesting, though, that when controlled for the number of words, a significant difference between the student and native writers pops up in regards to the use of NP subjects and copula complements. In both cases, the *student* writers have a higher average ratio of NPs in that position to the number of words.

Table 56: Comparison of Statistical Differences between Native and Student Writers’
Syntactic Positions of NPs

<i>comparison</i>	<i>subject</i>	<i>verbal object</i>	<i>object of preposition</i>	<i>genitive specifier</i>	<i>copula complement</i>	<i>apposition</i>
<i>total # of NPs in position</i>	$z = -1.54$ $p = .1236$ NS	$z = -2.30$ $p = .0220$ S*	$z = -2.79$ $p = .0054$ S*	$z = -.87$ $p = .3898$ NS	$z = -.41$ $p = .6818$ NS	$z = -.62$ $p = .5352$ NS
<i>total # NPs in position/ total # words</i>	$z = 2.89$ $p = .0038$ S*	$z = .65$ $p = .5156$ NS	$z = -.70$ $p = .4840$ NS	$z = 1.08$ $p = .2802$ NS	$z = 2.03$ $p = .0424$ S*	$z = .70$ $p = .4840$ NS

The student ratio for subjects is .1057, compared to the natives’ .0795. For copular complements, the student ratio is .0250, compared to the natives’ .0115. This means in effect that the student writers have more subject NPs per word than the native writers, and also that the students have more copula complement NPs per word than the native writers. It is not clear why this would happen. One possible reason for the difference in the copula complements, though, could be that the students are using more of them than the native writers as an effect of having studied them extensively in their English classes. *Be* is an irregular verb in English, so it would be covered more in classes, but it would also receive a lot of airtime in English classes because both Korean and Chinese students tend to omit the copula. This possibility could be an area of further research.

IIEP vs. EPE Students, Syntactic Positions of NPs

The IIEP and EPE students have significantly different numbers of NPs only in two positions: subject and verbal object. This difference is significant not only when the number of subjects and verbal objects is considered alone, but also when they are considered against the total number of words per essay. These details are summarized in the second landscaped table, and also in the table below.

Table 57: Comparison of Statistical Differences between IIEP and EPE Student Writers'
Syntactic Positions of NPs

<i>comparison</i>	<i>subject</i>	<i>verbal object</i>	<i>object of preposition</i>	<i>genitive specifier</i>	<i>copula complement</i>	<i>apposition</i>
<i>total # of NPs in position</i>	$z = 2.34$ $p = .0192$ S*	$z = 2.27$ $p = .0232$ S*	$z = -1.21$ $p = .2262$ NS	$z = .91$ $p = .3628$ NS	$z = 1.29$ $p = .1970$ NS	$z = -1.51$ $p = .1310$ NS
<i>total # NPs in position/ total # words</i>	$z = 3.02$ $p = .0026$ S*	$z = 2.04$ $p = .0414$ S*	$z = -1.29$ $p = .1970$ NS	$z = .60$ $p = .5486$ NS	$z = .91$ $p = .3628$ NS	$z = -1.59$ $p = .1118$ NS

It is not only interesting that there is a difference only in the subjects and objects, and that the difference holds for the absolute and normalized numbers of NPs, but also in that the IIEP students average more subject and object NPs than the EPE students. Averages and medians are given in the second landscaped table. This is unexpected because the IIEP students are the lower proficiency students, but they have higher averages. In the significant difference between the native and student subject and verbal object NPs, the natives had more verbal object NPs and a higher ratio between the number of subjects and the number of words. It is possible that this is another artifact of the essay topic differences, but overall it is unclear why the IIEP students would be more native-like in these two cases.

Korean vs. Chinese/ Taiwanese Students, NPs in Various Syntactic Positions

When considering other discourse features, differences have often been found between the IIEP and EPE student sub-groups, but few differences have been found between the Korean and Chinese/ Taiwanese language groups. For syntactic positions, there is in fact only one difference between the Korean and Chinese/ Taiwanese students:

There is a statistically significant difference between these groups only in the ratio of the number of copular complements³⁹ to the number of words per essay ($z = 1.81, p = .0702$). The statistics are summarized in the third landscaped table, and in the table that follows.

In this case, the Korean students have a higher ratio of NPs as copula complements to words (.0374) than the Chinese/ Taiwanese students (.0126). This means that the Korean students had more copula complements in their essays than the Chinese students. It is not clear what would account for this difference, but this may or may not affect the number of errors in determiners made by Korean students as questions about co-reference and old/ new information often lead to L2 mistakes in determiners in this position. The example repeated below in fact shows a Korean students' mistake in determiner selection after the copula. This issue will be considered in more depth when the student errors are examined.

Example 66: Error in an NP Complement of a Copula Verb in a Korean Student's Essay

Student 11: *I would be a **dictator** of my country Korea, if I could be anybody for one day because of the following three main reasons.*

³⁹ In all the essays, only one mistake was made in which the copula was omitted. Student 18 wrote: "She embraced [*embraced*] me tidly [*tightly*] and said that it [*was*] just the same thing she wanted to do." In this case, the noun phrase *the same thing she wanted to do* was counted as the complement of the copula. This could be simply a writing error on the part of Student 18, or Student 18 could be using a zero morpheme as a copula, or something else. As in other situations, the intention was guessed at and used to analyze the structure.

Table 58: Syntactic Positions of NPs in Native and Student Essays

	<i>subject</i>	<i>verbal object</i>	<i>object of preposition</i>	<i>genitive specifier</i>	<i>copula complement</i>	<i>apposition</i>	<i>comparison</i>	<i>title</i>
<i>Native total</i>	522*	442*	646*	115*	73	78	15	3*
<i>Student total</i>	447	293	455	113	105*	95*	19*	2
<i>Native average</i>	26.1*	22.1*	32.3*	5.75*	3.65	3.9	--	--
<i>Student average</i>	22.35	14.65	22.75	5.65	5.25*	4.75*	--	--
<i>Native median</i>	23.5*	19*	33.5*	6*	3	3	--	--
<i>Student median</i>	20	15	18	4	3	3.5*	--	--
<i>Wilcoxon Rank Sums Test ($\alpha = .10$)</i>	z = -1.54, [p = .1236] NS	z = -2.29, [p = .0220] S*	z = -2.79, [p = .0054] S*	z = -.86, [p = .3898] NS	z = -.41, [p = .6818] NS	z = -.62, [p = .5352] NS	--	--

Table 59: Syntactic Positions of NPs in IIEP and EPE Student Essays

	<i>subject</i>	<i>verbal object</i>	<i>object of preposition</i>	<i>genitive specifier</i>	<i>copula complement</i>	<i>apposition</i>	<i>comparison</i>	<i>title</i>
<i>IIEP total</i>	262*	171*	230*	63*	40	23	6	1
<i>EPE total</i>	185	122	225	50	65*	72*	13*	1
<i>IIEP average</i>	26.2*	17.1*	23*	6.3*	4	2.3	--	--
<i>EPE average</i>	18.5	12.2	22.5	5	6.5*	7.2*	--	--
<i>IIEP median</i>	25.5*	18*	14	4.5*	4*	1.5	--	--
<i>EPE median</i>	17.5	12	20.5*	3.5	2	5*	--	--
<i>Wilcoxon Rank Sums Test ($\alpha = .10$)</i>	z = 2.34, [p = .0192] S*	z = 2.27, [p = .0232] S*	z = -1.21, [p = .2262] NS	z = .91, [p = .3628] NS	z = 1.29, [p = .1970] NS	z = -1.51, [p = .1310] NS	--	--

Table 60: Syntactic Positions of NPs in Korean and Chinese/Taiwanese Student Essays

	<i>subject</i>	<i>verbal object</i>	<i>object of preposition</i>	<i>genitive specifier</i>	<i>copula complement</i>	<i>apposition</i>	<i>comparison</i>	<i>title</i>
<i>Korea total</i>	253*	164*	198	64*	80*	39	9	1
<i>Chinese/Taiwanese total</i>	194	129	257*	49	25	56*	10*	1
<i>Korea average</i>	25.3*	16.4*	19.8	6.4*	8*	3.9	--	--
<i>Chinese/Taiwanese average</i>	19.4	12.9	25.7*	4.9	2.5	5.6*	--	--
<i>Korea median</i>	24.5*	15.5*	21*	4.5*	5*	2.5	--	--
<i>Chinese/Taiwanese median</i>	18	13	18	3.5	2	4*	--	--
<i>Wilcoxon Rank Sums Test</i> ($\alpha = .10$)	z = 1.36, [p = .1738] NS	z = .98, [p = .3270] NS	z = -.45, [p = .6528] NS	z = .45, [p = .6528] NS	z = 1.59, [p = .1118] NS	z = -.15, [p = .8808] NS	--	--

Table 61: Comparison of Statistical Differences between Korean and Chinese/Taiwanese Student Writers' Syntactic Positions of NPs

<i>comparison</i>	<i>subject</i>	<i>verbal object</i>	<i>object of preposition</i>	<i>genitive specifier</i>	<i>copula complement</i>	<i>apposition</i>
<i>total # NPs in position</i>	$z = 1.36$ $p = .1738$ NS	$z = .98$ $p = .3270$ NS	$z = -.45$ $p = .6528$ NS	$z = .45$ $p = .6528$ NS	$z = 1.59$ $p = .1118$ NS	$z = -.15$ $p = .8808$ NS
<i>total # NPs in position/ total # words</i>	$z = .60$ $p = .5486$ NS	$z = .53$ $p = .5962$ NS	$z = -.68$ $p = .4966$ NS	$z = .08$ $p = .9362$ NS	$z = 1.81$ $p = .0702$ S*	$z = -.60$ $p = .5486$ NS

Conclusion

Because syntactic positions are important for co-reference and discourse cohesion, they will be reconsidered when the NPs in chains are looked at in more depth. In just examining the number of NPs in various positions, few differences were actually found, and those that were discovered were most often differences between the native and student writers, or differences between the IIEP students and the EPE students. The differences are summarized in the table on the next page.

When the natives and the students were compared, there was a difference in the number of verbal objects and objects of prepositions. However, as this could be partly caused by the native writers having significantly more words and NPs in their essays than the students, the ratio of number of objects to number of words was also examined. When the length of the essays is accounted for, there is in fact no significant difference between these two groups in the number of verbal objects and the number of prepositional objects. It is surprising, though, that when the ratios are considered, a

significant difference in the ratio of subjects to words and the ratio of copula complements to words comes out. The students had more subjects and more copula complements per word than the native writers. The reason for this needs to be investigated further.

Table 62: Group Averages for Significant Findings Relating to Syntactic Positions of NPs

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/ Taiwanese students</i>	
	<i>student average</i>	<i>native average</i>	<i>IIEP average</i>	<i>EPE average</i>	<i>Korean average</i>	<i>Chinese/ Taiwan</i>
<i># subjects</i>			26.2*	18.5		
<i># subjects/ # words</i>	.1057*	.0795	.1260*	.0855		
<i># verbal objects</i>	14.65	22.1*	17.1*	12.2		
<i># verbal objects/ # words</i>			.0821*	.0580		
<i># objects of prepositions</i>	22.75	32.3*				
<i># prep. objects/ # words</i>						
<i># genitive specifiers</i>						
<i># genitive specifiers/ # words</i>						
<i># copula complements</i>						
<i># copula comp.s/ # words</i>	.0250*	.0115			.0374*	.0126
<i># appositions</i>						
<i># appositions/ # words</i>						

Darkly shaded boxes indicate no significant difference was found. * indicates the higher average.

When the IIEP and EPE students were compared, it was found that the lower-proficiency IIEP students had significantly more subjects and verbal objects in their essays than the EPE students. This is another unexpected finding. When the Korean and Chinese/ Taiwanese sub-groups were compared, the only significant difference found was in the number of copula complements to the total number of words. The Korean students in fact had more copulas per word than the Chinese/ Taiwanese students. The difference between the Korean and Chinese/ Taiwanese students in this ratio, and the difference in the native and student writers in this ratio, could be affected by explicit grammar instruction in the classroom of the copula and determiners and NPs after the copula. In many of these cases, more research needs to be pursued.

The next topic to be investigated is the number of word types used by the student and native writers. At times so far in this study, the fact that students may not be varying their vocabulary as much as native speakers has been considered as a cause of some of the differences in discourse use by the two groups. This next section will look at the variety of words used by the writers.

Total Number of Distinct Words (Word Types)

word types = number of distinct kinds of words; like dictionary entries, each word is different; for example, *bird, the, fish, and...*

word tokens = number of times that any particular distinct word is used; for example, in the sentence *The big dog barked at the small dog* there are two tokens of *the*, two tokens of *dog*, one token of *bark...*

The number of distinct word types is an indication of how variable the vocabulary used by a writer is. Although a piece of writing may have 500 total word tokens, there may be only 250 distinct types of words, with several of the types used multiple times. A writer with a less developed vocabulary may have, for example, only 200 word types in a 500 word essay. They are re-using a more limited number of words. It is therefore

possible that the fewer the words, the more connections there are likely to be between some of them. In this section, this will be examined.

Other conclusions can be made by looking at the number of distinct word types. This may indicate areas of difference in co-reference, for example if one group is using more pronouns than another. This is just another way to look at how discourse is constructed differently by different groups.

Native vs. Student Writers, Number of Types

As may be expected, the native writers do in fact tend to use more word types than the students. Student 4 has the fewest types of any student, with only 66. The student with the most word types is Student 2, who uses 162 distinct word types in his/her essay. The range from the smallest number of word types to the largest in the essays by the student writers is 96, which is about a third of the range in the number of word types used by native writers. The native writer with the fewest different words is Native 21, who uses 72 different types. This is 222 fewer different words than the native writer with the most types, which is Native 32 with 294 different word types. For the native writers, the most and least types corresponds with the essays with the most and the least total number of words or word tokens. Native 32 wrote the longest essay (662 words), and Native 21 wrote the shortest (93 words). For the students, Student 2 did write the longest essay (328 words) and have the most distinct word types, but Student 4, who had the smallest number of word types, did not write the shortest essay. This means that Student 4 (whose essay had 170 words) re-used more vocabulary words than Student 18, who did write the shortest student essay (with 125 words). The number of word types in each essay is presented in the table below.

Table 63: Total Number of Distinct Word Types per Essay

<i>ESL Students</i>		<i>Native English Speakers</i>	
<i>participant #</i>	<i>#</i>	<i>participant #</i>	<i>#</i>
1	96	21	72
2	162	22	160
3	135	23	96
4	66	24	129
5	106	25	185
6	89	26	153
7	104	27	124
8	88	28	168
9	132	29	205
10	106	30	186
11	152	31	220
12	121	32	294
13	114	33	181
14	95	34	144
15	144	35	160
16	108	36	269
17	80	37	156
18	88	38	229
19	106	39	153
20	89	40	107

The difference in the number of distinct word types in the native and student essays is significant ($z = -3.98, p = .0001$). In fact, the native writers have a much wider range from the lowest number to the highest number of distinct word types than the student writers. Also, the average and the median number of distinct word types in the

native writers' essays is at least 50 more than the average and median for the students'.

See the following table.

Table 64: Average and Median Total Number of Word Types per Essay, Native vs. Student Writers

	<i>Students</i>	<i>Natives</i>
<i>Fewest types</i>	66, Student 4	72, Native 21
<i>Most types</i>	162, Student 2	294, Native 32
<i>Average # types</i>	109.05	169.55
<i>Median # types</i>	106	160
<i>Wilcoxon Rank Sums Test</i>	$z = -3.98, p = .0001$	

However, as with other aspects of the essays that have been examined, this difference could be a function of essay length. Therefore, the ratio of the number of word types to the total number of words or word tokens was also investigated in regards to differences between the native and the student writers. Again, as with so many other features in these essays, when the length of the essays is controlled for, there is no longer any significant difference between the groups ($z = -.51, p = .6100$). See the ratios in the table that follows on the next page.

Since any particular word type from an open class is not likely to be used more than once or twice, it is not surprising that the more words or word tokens there are in an essay, the more distinct word kinds there are.

Table 65: Ranks of Word Types Based on the Number of Tokens of Each Word Type,
Native vs. Student Writers

<i>participant #</i>	<i>participant category</i>	<i>total # word types</i>	<i>total # words</i>	<i>total # types / total # words</i>	<i>rank</i>
6	Student	89	233	.3820	1
4	Student	66	171	.3860	2
37	Native	156	380	.4105	3
9	Student	132	312	.4231	4
39	Native	153	359	.4262	5
30	Native	186	428	.4346	6
14	Student	95	216	.4398	7
32	Native	294	662	.4441	8
16	Student	108	243	.4444	9
12	Student	121	265	.4566	10
26	Native	153	319	.4796	11
36	Native	269	560	.4804	12
15	Student	144	295	.4881	13
38	Native	229	469	.4883	14
2	Student	162	328	.4939	15
29	Native	205	412	.4976	16
17	Student	80	158	.5063	17
25	Native	185	365	.5068	18
22	Native	160	311	.5145	19
8	Student	88	171	.5146	20
10	Student	106	205	.5171	21
20	Student	89	171	.5205	22
33	Native	181	343	.5277	23
7	Student	104	190	.5474	24
5	Student	106	193	.5492	25
31	Native	220	398	.5528	26
24	Native	129	231	.5584	27
1	Student	96	170	.5647	28

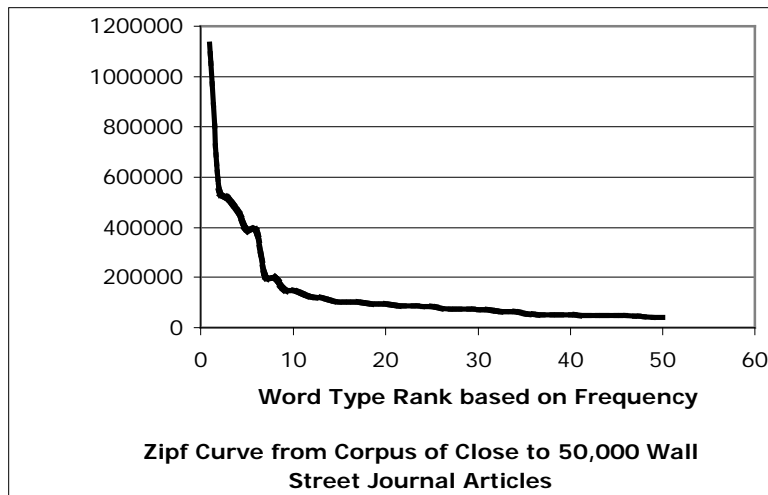
Table 65—Continued

<i>participant #</i>	<i>participant category</i>	<i>total # word types</i>	<i>total # words</i>	<i>total # types / total # words</i>	<i>rank</i>
28	Native	168	296	.5676	29
35	Native	160	279	.5735	30
13	Student	114	195	.5846	31
34	Native	144	241	.5975	32
19	Student	106	177	.5989	33
23	Native	96	159	.6038	34
11	Student	152	249	.6104	35
3	Student	135	220	.6136	36
27	Native	124	199	.6231	37
18	Student	88	125	.7040	38
40	Native	107	146	.7329	39
21	Native	72	93	.7742	40
$z = -.51, NS [p = .6100]$					

In fact, if the number of tokens of each different word type is plotted against the word type's rank based on the number of its tokens, a curve that decreases exponentially and levels off quickly results. This shows that most words are rarely used, and some words are very, very common. This pattern was studied by linguist George Kingsley Zipf (see Li 2004). The figure that follows shows what a Zipf curve, or the plot of the word occurrences to their rank typically looks like.

This information leads to the question: Are there differences between the native and the student writers in regards to the word types ranked by their number of tokens?

Figure 25: A Zipf Curve Based on Word Distributions in Wall Street Journal Articles (Li 2004)



In the Wall Street Journal corpus used to make the graph above, and in most other corpora of English, the most frequent word kinds are closed class words, such as determiners, conjunctions, pronouns, and other grammatical function words. For example, see the table on the next page.

When looking at the word types with the most tokens in the native writing half of the corpus used in this study, the top three most frequently used word kinds for every writer are words such as those in the table: *a, as, and, I, if, in, is, of, that, the, this, to, and we*. However, for the learner half of the corpus, it can only be said that the word types with the most tokens for every student is function word for a closed class word (*and, he, I, in, is, of, the, and to*). Six of the twenty student writers have a common noun as their second most commonly used word. See the landscaped table.

Table 66: Three English Corpora and their High Frequency Function Words (Li 2004)

<i>corpus name/ source</i>	<i># total words in corpus</i>	<i># of top ranking positions that are closed class/function words</i>
423 short Time Magazine articles	over 90,000 words	more than the top 20: <i>the, of, to, a, and, in, that, for, was, with, his, is, he, as, on, by, at, it, from, but</i>
IGB TREC Vol. 3 Corpus	336,310 documents 125,720,891 total words 508,209 word types	the top nine: <i>the, of, to, and, in, is, for, The⁴⁰, that</i>
46,449 Wall Street Journal Articles	19,000,000 words	the top nine: <i>the, of, to, a, in, and, that, for, is</i>

⁴⁰ The concordance program that was used to count the number of tokens of each distinct word types count capitalized and lower case words as different tokens. This was necessary in order to keep proper nouns counted separately.

Table 67: Ten Most Common Word Types in Each Essay

Participant	<i>most common words, from first most common on the left, to tenth most common on the right</i>									
	1 st	2 nd	3 rd	4 th	5 th	6 th	7 th	8 th	9 th	10 th
1	is	of	the	to	their	in	culture	our	a	and
2	of	is	and	the	food	in	western	Korean	people	Korea
3	the	countries	of	in	and	population	to	because	developed	is
4	the	population	is	and	countries	in	country	that	China	India
5	the	of	and	population	growth	to	a	economic	is	graph
6	the	population	of	and	in	China	countries	increase	as	has
7	the	culture	of	and	is	more	our	country	a	are
8	the	is	population	and	growth	in	of	selected	that	a
9	in	to	the	population	a	growth	India	people	will	be
10	the	will	population	to	and	countries	in	a	as	developing
11	to	and	the	I	of	my	in	all	be	for
12	he	I	and	my	to	the	God	father	very	a
13	I	my	a	and	to	for	job	no	this	was
14	I	and	a	will	be	for	them	they	can	help
15	the	I	to	and	of	that	was	in	life	but

Table 67—Continued

Participant	<i>most common words, from first most common on the left, to tenth most common on the right</i>									
	1 st	2 nd	3 rd	4 th	5 th	6 th	7 th	8 th	9 th	10 th
16	I	the	afraid	and	confused	is	to	a	would	in
17	I	perseverance	will	to	and	have	important	in	the	do
18	and	I	to	friend	we	afraid	that	together	was	at
19	the	wedding	to	and	they	their	a	couple	of	party
20	the	it	are	I	is	so	and	you	beautiful	country
21	is	to	the	a	and	classes	important	in	knowledge	of
22	of	the	and	education	a	to	general	in	can	requirements
23	a	I	as	education	of	general	my	particular	subject	be
24	of	the	to	a	education	would	general	in	students	an
25	the	to	of	in	students	a	and	for	this	education
26	to	the	we	of	must	and	have	in	person	understand
27	to	is	that	will	a	and	as	it	no	bar
28	the	to	if	is	requirements	in	and	be	for	of
29	and	the	a	to	of	history	in	general	I	is
30	to	that	the	I	a	of	and	in	education	not
31	a	the	in	to	of	and	are	is	education	students
32	of	and	the	to	a	in	we	education	use	are

Table 67—Continued

Participant	<i>most common words, from first most common on the left, to tenth most common on the right</i>									
	1 st	2 nd	3 rd	4 th	5 th	6 th	7 th	8 th	9 th	10 th
33	the	to	that	of	are	age	I	can	in	it
34	the	is	to	a	education	of	and	I	courses	general
35	the	to	of	and	a	college	in	students	study	for
36	the	to	and	I	in	of	a	for	is	that
37	the	this	in	to	underage	would	be	for	from	a
38	the	in	to	and	students	of	education	a	I	for
39	of	the	to	a	in	and	his	knowledge	that	have
40	of	a	the	education	students	and	are	general	I	beyond

Although essay length may again play a part in what word types have the most tokens, the fact that all the students' most tokens are for function words indicates that there may be a difference between the student and native writers concerning the frequency of use of function words such as determiners.

Looking at how many word types have only one or two tokens may also point out differences in the vocabulary and skill at creating discourse. For the students, Student 14 had the lowest percentage (52%) of word types with only one token. For the native writers, Native 37 had the lowest percentage of single tokens, with 58% of the word types appearing only one time. So, Native 37 and Student 14 had fewer unique words than others in their groups. At the other end of the spectrum, Student 3 and Native 21 had the highest percentage of their word types used only once. It is interesting to note that Native 21 had the shortest essay of all the native writers. When the student and native writers' percentages were statistically compared, the native writers were found to have a significantly higher percentage of word types used only once ($z = -2.08, p = .0376$). This means that the native writers used a more varied vocabulary. See these details summarized in the table below.

Table 68: Percentage of Word Types with Only One Token, Native vs. Student Writers

	<i>Students</i>	<i>Natives</i>
<i>Lowest % used once</i>	52%, Student 14	58%, Native 37
<i>Highest % used once</i>	77%, Student 3	85%, Native 21
<i>Average % used once</i>	64%	70%
<i>Wilcoxon Rank Sums Test</i>	$z = -2.08, p = .0376$	

This conclusion is supported by the fact that the learners had a higher percentage of words used twice than the native speakers. This difference was also statistically significant ($z = 3.08, p = .0020$). See the table below.

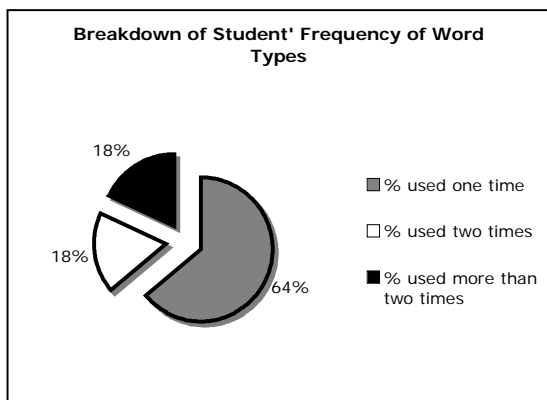
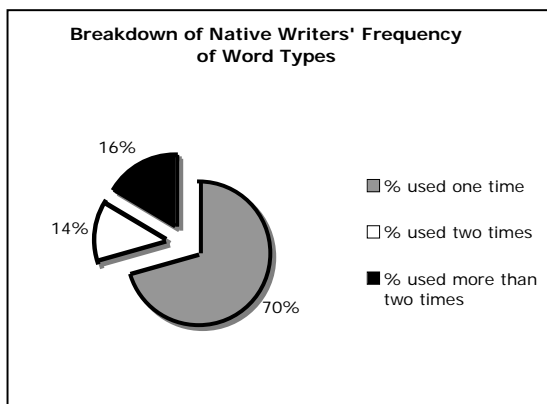
Table 69: Percentage of Word Types with Exactly Two Tokens, Native vs. Student Writers

	<i>Students</i>	<i>Natives</i>
<i>Lowest % used twice</i>	9%, Student 2	7%, Native 21
<i>Highest % used twice</i>	28%, Student 16	22%, Native 24
<i>Average % used twice</i>	18%	14%
<i>Wilcoxon Rank Sums Test</i>	$z = 3.08, p = .0020$	

Looking at the differences graphically in a pie chart (see the next page), it is noticeable how most of the word types are single tokens, and how few words have two or more tokens. Furthermore, it can be seen that the native writers have more distinct word types in their essays.

The next sections will compare the student sub-groups.

Figure 26: Comparison of the Percentage of Tokens of Different Word Types, Native vs. Student Writers



IIEP vs. EPE Students and Korean vs. Chinese/Taiwanese
Students, Number of Word Types

Because the students are all language learners, there is not expected to be as big a difference between the sub-groups. However, if patterns in the number of word types follow other patterns in the data, then there is more likely to be a difference between the IIEP and EPE students than between the two L1 backgrounds.

IIEP vs. EPE Students, Number of Word Types

The IIEP students have a slightly smaller range from the lowest number of distinct word types to the highest, but overall there is no significant difference between the IIEP and EPE students. The IIEP student with the most word types is Student 11, who has 152 unique words. This is very similar to the EPE student with the most word types, which is Student 2, who has 162 distinct word types. The IIEP student and EPE student with the fewest word types are also not very far apart. For IIEP, Student 17 has the fewest word types, with only 80. This is higher than the number of word types used by the EPE student with the fewest word types. Student 4 had only 66 distinct word types, which is almost 20 fewer than the lower proficiency student with the least. As stated, though, there is no *significant* difference between these sub-groups in regards to number of distinct word types ($z = .38, p = .7040$). See Table 70.

When the number of word types is divided by the number of total words, there is still no significant difference between the IIEP and EPE students in terms of number of word types ($z = .68, p = .4966$). This is unexpected in that it implies that the two groups have about as equally varied or complex vocabularies. See Table 71.

Table 70: Average and Median Total Number of Word Types per Essay, IIEP vs. EPE and Korea vs. China/Taiwan

	<i>IIEP students</i>	<i>EPE students</i>	<i>Korean Stud.s</i>	<i>Chinese/Taiwanese</i>
<i>Most types</i>	152, Student 11	162, Student 2	162, Student 2	132, Student 9
<i>Fewest types</i>	80, Student 17	66, Student 4	66, Student 4	80, Student 17
<i>Average # types</i>	109.7	108.4	116.6	101.5
<i>Median # types</i>	107	105	113.5	105
<i>Wilcoxon Rank Sum</i>	$z = .38, NS [p = .7040]$		$z = 1.21, NS [p = .2262]$	

This conclusion is held up by the fact that there is no significant difference between the IIEP and EPE students in regards to the percentages of the word types with only one or two word tokens, as shown in Table 72.

Table 71: IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test, Number of Word Types

	<i>z</i>	<i>p</i>	<i>significant difference?</i>
<i>total # words</i>	-.08	NS [.9362]	no
<i>total # word types</i>	.38	NS [.7040]	no
<i>ratio of # types per total # words</i>	.68	NS [.4966]	no

Table 72: IIEP Essays as Compared to EPE Students' Essays, Wilcoxon Rank Sums Test, Percentage of Word Types with Only One or Two Tokens

	<i>z</i>	<i>p</i>	<i>significant difference?</i>
<i>% word types used only once</i>	.08	.9362	no
<i>% word types used twice</i>	.45	.6528	no

Korean vs. Chinese/ Taiwanese Students, Number of Word Types

Similar to comparing the IIEP and EPE sub-groups for number of word types, there is very little difference when comparing the Korean and Chinese/ Taiwanese student groups. The Chinese/ Taiwanese students have a smaller range between the fewest and greatest number of distinct word tokens, but the difference is not significant ($z = 1.21, p = .2262$). The range for the Chinese/ Taiwanese students is 52, with Student 9

having the most word types (132 types), and Student 17 having the fewest (80 types). The range for the Korean students is 96, almost double that for the Chinese/ Taiwanese group. Korean Student 2 has the most word types, having 162 unique words. Student 4 has the fewest word types for the Korean writers, with only 66 distinct types. Despite this, the averages of the two groups are not significantly different. The Korean students had on average 116.6 distinct word types in their essays, while the Chinese students on average only used 101.5 distinct types. The table below summarizes these details.

Table 73: Average and Median Total Number of Distinct Word Types per Essay, IIEP vs. EPE and Korea vs. China/Taiwan

	<i>IIEP students</i>	<i>EPE students</i>	<i>Korean Stud.s</i>	<i>Chinese/Taiwanese</i>
<i>Most types</i>	152, Student 11	162, Student 2	162, Student 2	132, Student 9
<i>Fewest types</i>	80, Student 17	66, Student 4	66, Student 4	80, Student 17
<i>Average # types</i>	109.7	108.4	116.6	101.5
<i>Median # types</i>	107	105	113.5	105
<i>Wilcoxon Rank Sum</i>	$z = .38$, NS [$p = .7040$]		$z = 1.21$, NS [$p = .2262$]	

When the number of distinct word types is normalized according to the total number of words per essay, there is still no significant difference between the Korean and Chinese/ Taiwanese students in terms of how many distinct word types they use. This continues the pattern of there not being much difference between the Korean and Chinese/ Taiwanese students in the aspects of the discourse features of their essays. Furthermore, there is no significant difference between the groups in the number of distinct word types with only one or two tokens. All of this information is summarized in the next table.

Table 74: Korean Students' Essays as Compared to Chinese and Taiwanese Students' Essays, Wilcoxon Rank Sums Test, Number of Word Types, Percentage of Types with Only One or Two Tokens

	<i>z</i>	<i>p</i>	<i>significant difference?</i>
<i>total # words</i>	.98	NS [.3720]	no
<i>total # types</i>	1.21	NS [.2262]	no
<i>ratio of # types per total # words</i>	0	NS [1]	no
<i>% of word types used once</i>	.08	.9362	no
<i>% of word types used twice</i>	-.53	.5962	no

The next section will look specifically at the differences between the groups in their use of articles, determiners, and pronouns.

Articles, Demonstratives, and Pronouns

Counting the number of word types permitted examination of the frequency of different kinds of articles, demonstratives, and pronouns in the student and native writers' essays. The next section looks at the differences in frequencies between the students and native writers.

Student L2 English vs. Native English Writers

In the table that follows on the next page, the number of times each writer used *a/an*, *the*, a demonstrative, or a pronoun is given. A breakdown of how the different kinds of pronouns or demonstratives were used by student and native writers will be given later.

Table 75: Number of Articles, Demonstratives, and Pronouns per Essay

<i>participant #</i>	<i># a/an</i>	<i># the</i>	<i># demonstratives</i>	<i># pronouns</i>
1	4	8	5	19
2	4	11	8	7
3	2	14	4	4
4	0	15	6	2
5	4	18	3	6
6	0	20	2	8
7	4	10	4	15
8	4	13	6	5
9	7	15	7	4
10	4	21	3	35
11	3	9	6	30
12	5	8	3	30
13	8	3	4	33
14	10	3	3	31
15	3	26	8	36
16	5	12	9	25
17	0	4	2	18
18	0	1	3	19
19	4	13	2	16
20	1	11	1	26
21	3	4	2	4
22	11	13	10	10
23	8	2	3	18
24	12	13	7	6
25	11	22	16	12
26	5	16	8	31
27	4	3	7	7
28	4	19	4	8
29	13	15	8	20
30	14	14	19	31

Table 75—Continued

<i>participant #</i>	<i># a/an</i>	<i># the</i>	<i># demonstratives</i>	<i># pronouns</i>
31	18	15	12	22
32	23	22	10	34
33	6	17	17	26
34	8	11	4	15
35	8	13	3	5
36	18	30	16	33
37	5	8	13	14
38	9	19	16	23
39	8	12	10	20
40	6	5	4	6

Frequency of Use of *the*

Comparing the ranges of individual values for the student and native groups, they are closest for *the* and for pronouns. Student 18 had the fewest uses of *the* with only one, and Student 15 had the most *the*, using it 26 times. This makes the range in the students' frequency of use of *the* 25. The range for the native writers is similar at 28. The native writer with the fewest instances of *the* was Native 23 who used it twice. Native 36 used *the* the greatest number of times with a total of 30 tokens.

In addition to the ranges, the average number of times *the* was used per essay is very close for each group. The students averaged 11.75 instances of *the* per essay, and the native English writers averaged 13.65. When the Wilcoxon Rank Sums Test was used to test whether the difference between these averages was significant, no significant difference was found when testing the absolute number of times *the* was used ($z = -1.30$, $p = .1936$). Neither was a significant difference found between the student and native writers when testing the ratio of the number of uses of *the* to the total number of words

per essay ($z = 1.60, p = .1096$). This second finding is interesting since the native writers typically wrote longer essays, which might have led to a significant difference in the ratios.

Number of Pronouns

As mentioned above, the student and the native writers also had very similar ranges in how many pronouns were used per essay. The student range was 34—Student 4 had only two pronouns, and Student 15 had the most pronouns with 36. The native range was 30. However, unlike the students, the native writer with the longest essay had the most pronouns (Native 32, 34 pronouns), and the native writer with the shortest essay had the fewest pronouns (Native 21, 4 pronouns). Therefore, it seems that the frequency of pronouns is related to essay length for the natives but not necessarily for the students.

The averages and the medians for the two groups were not that different either. The students averaged 18.45 pronouns per essay, while the native writers averaged 17.25 per essay. See the averages and medians summarized in the table below.

Table 76: Average Number of Articles, Demonstratives, and Pronouns for Native and Student Writers

	<i>ESL Students averages</i>		<i>Native writers averages</i>	
	<i>average</i>	<i>median</i>	<i>average</i>	<i>median</i>
<i>a/an</i>	3.6	4	9.7	8
<i>the</i>	11.75	11.5	13.65	13.5
<i>total # articles</i>	15.35	15	23.35	23
<i>demonstratives</i>	4.45	4	9.45	9
<i>pronouns</i>	18.45	18.5	17.25	16.5

No significant difference was actually found when the counts of the number of pronouns by themselves were tested ($z = -.03, p = .9760$). However, unlike what was seen with the use of *the*, when the number of pronouns was controlled by the number of words, a significant difference was found ($z = 1.79, p = .0734$). The student writers had significantly more pronouns given the length of their essays than the native writers.

Aside from there being a difference in the concentration of pronouns in the student and native essays, the pronouns that the groups are using most often differ. The table that follows shows the five most commonly used pronouns for each group.

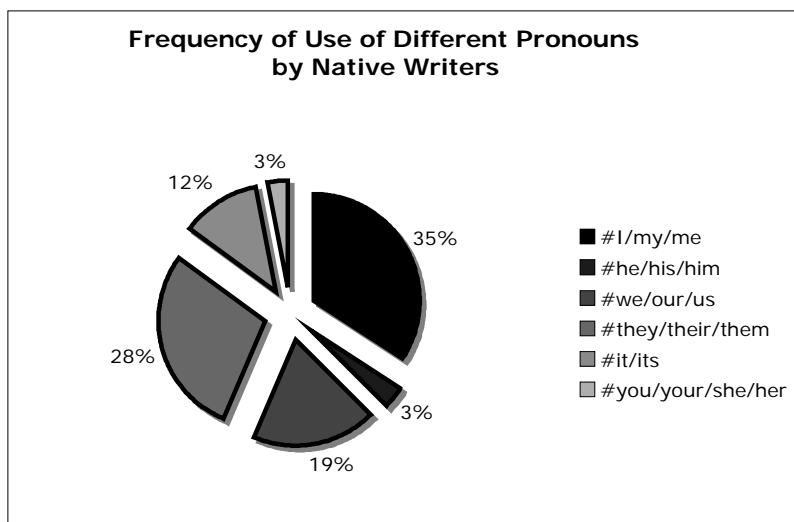
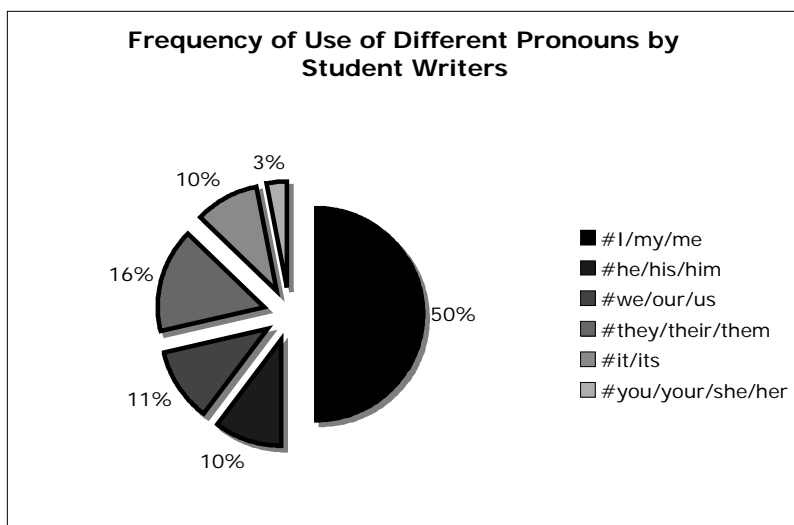
Table 77: Comparison of the Pronouns Used Most Often by Student and Native Writers

<i>Student essays</i>		<i>Native essays</i>	
<i>pronoun, from most to least common</i>	<i>number of occurrences</i>	<i>pronoun, from most to least common</i>	<i>number of occurrences</i>
I	127	I	86
my	44	they	41
it	31	it	40
he	26	their	36
their	26	we	32

I is most common in both the student and native essays. The second most common pronoun for the native writers, though, is *they*, while for the students it is *my*. So, for the students, two co-referential pronouns are the most frequent. For the native writers, the frequency of *they*, *it*, and *their* rather than other pronouns could be due to their having the choice only to write on two relatively more formal topics. This could also be the reason for there being significantly more pronouns in the student essays. Less formal, easier topics allow students to talk about themselves. The IIEP topics clearly have been

selected to allow students to write about their own experiences, and first person pronouns do not require overt NP antecedents in the prior discourse. This might account for part of why the students have statistically more pronouns.

Figure 27: Comparison of Native and Student Writer Use of Pronouns



Frequency of Use of Demonstrative Determiners

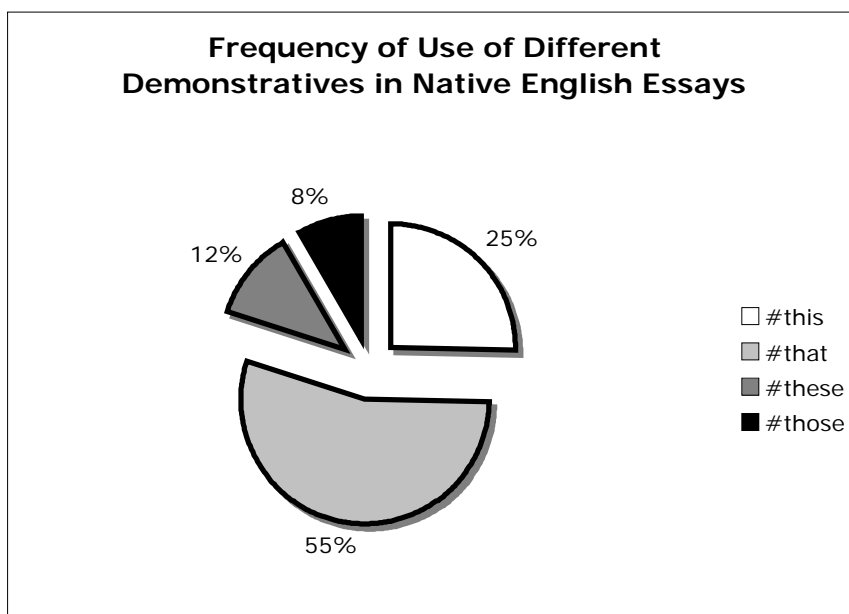
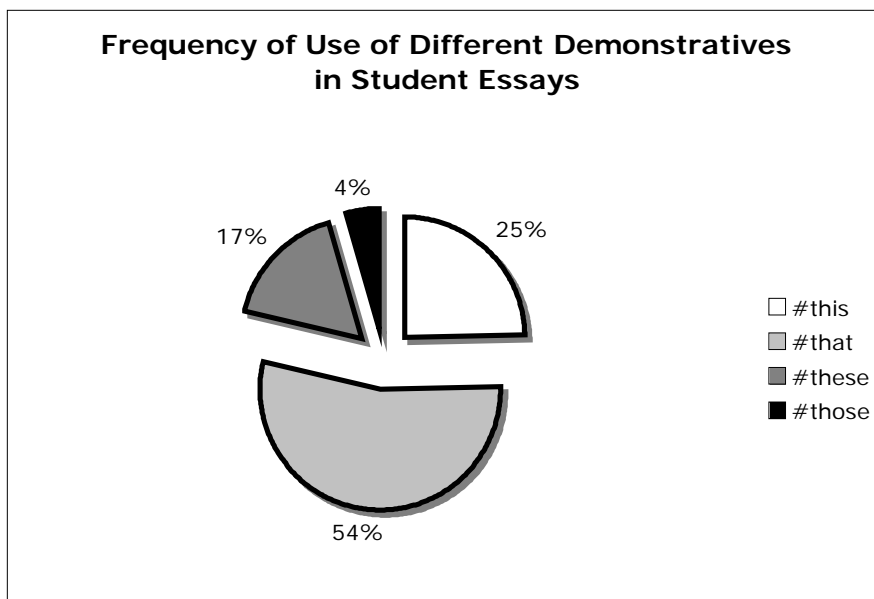
For both just the number of demonstratives ($z = -3.52, p = \leq .0004$) and the number of demonstratives per the total number of words in each essay ($z = -2.11, p = .0348$), the student and native writers' essays were significantly different. The student using the fewest demonstratives was Student 20, who used only one. The student with the most demonstratives was Student 16, who had nine demonstratives in his/her essay. The range of the student demonstrative frequencies is about half of the range of those of the native writers. The native writer with the fewest demonstratives was Native 21, who had only two. Native 30 had the most with 19 demonstratives. The average number of demonstratives written by the students (4.45 per essay) was also about half of the average for the natives (9.45).

Despite the differences in the amount of determiner use, the frequency rank order of the demonstratives is the same for the natives and the students. Both groups used *that* the most, followed in rate of occurrence by *this*, *these*, and then *those*. The figures on the next page show that the main differences in the percentages of the two groups are that the students use *those* less than the native writers.

Use of *a/an*

One interesting result regarding these frequencies is that the student writers use *a/an* a significantly smaller amount of times than the native writers ($z = -4.36, p < .00003$). This is in part surprising because no significant difference was found in the frequency of *the*. In looking at the native writers' essays, the use of *a/an* seems related to total number of words, but this does not seem to be the case for the student writers. The native writer with the highest number of *a/an* was Native 32 with 23 occurrences, and the native writer with the fewest was Native 21 with only three uses of *a/an*. Native 32 has the longest essay in the native group, while Native 21 has the shortest.

Figure 28: Comparison of Native and Student Writer Use of Demonstratives



For the students, the longest and the shortest essay writers do not correspond to the essays with the most and the least number of *a/an*. The student with the most occurrences of *a/an* was Student 14, who had ten instances, about half the number in Native 32's essay. When looking at the students with the smallest number of articles, a big difference between the native and student writers' use of articles, determiners, and pronouns comes out: Four different students (Students 18, 17, 6, and 4) had no instances of *a/an* at all in their essays. In contrast, every single native writer had at least one use of *a/an*. The cause of this difference is unknown.

Robertson (2000)'s discourse rule the "lexical transfer principle" is problematic when this data is considered in more detail. This transfer principal predicts that the Chinese/Taiwanese students should be making such substitutions as replacing *the* with *this* and *a/an* with *one*. The underuse of *a/an* found here would be consistent with this possibility. However, there is limited evidence to support the use of *one* as a marker of indefiniteness. *One* is not used very often in student essays. See the table below.

Table 78: Use of *one* in Student Essays

<i>participant #</i>	<i># tokens of one</i>	<i>participant #</i>	<i>#tokens of one</i>
1	0	11	1
2	7	12	0
3	0	13	0
4	0	14	2
5	0	15	2
6	3	16	2
7	0	17	3
8	1	18	1
9	0	19	2
10	0	20	0

These uses of *one* are also predominantly quantificational, pronominal, and used with the meaning “someone”, and not marking indefiniteness. For instance, Student 2 has the most occurrences of *one*, but none of the seven contexts is *one* an indefinite determiner. See the examples below.

Example 67: Quantificational, Pronominal, and “Someone” Uses of *one* in Student 2’s Essay

quantificational:

- (1) “**one** of the most important word which express the world”
(*one of the most important words expressing the concept of globalization*)
- (2) “**one** of the biggest changes of Korea”

pronominal:

- (3) “Western clothes are much more comfortable to behave than Korean **one**.”
(*it is much easier to move around in traditional Western clothes than traditional Korean clothes*)
- (4) “Korean traditional foods are totally different from western **one**.”
- (5) “...burgersize and cokesize is smaller and less fatty than western **one**”
(*the size of fast food meals is smaller in Korea than in the US, and the meals are less fatty in Korea*)
- (6) “Korean food is much better to health than western **one**”

“someone” meaning:

- (7) “Including housing, wearing and eating is most significant factor of **one**’s life.”
(*shelter, clothing, and food are three basic needs of human beings*)

Moreover, Student 2 is Korean and not Chinese. Even assuming that Korean and Chinese would both transfer this discourse rule to English, the validity of this hypothesis is questionable. In these excerpts, Student 2 is making quite a few grammatical and word choice errors, but the use of *one* is native-like, except for the lack of plural agreement on some of the pronominal uses. Thus, it is not that Student 2 is achieving native-like competence on the whole. Student 2 just does not use *one* as an indefinite.

In all of the student essays, there are only three times when *one* might be being used as an indefinite. See the examples below.

Example 68: Possible Indefinite Uses of *one* in Student 15 and Student 17's Essays

Student 15 (Korean):

- (1) "During the conflicts, **one** resident fired himself and the situation got worse and worse."

(During the conflicts, one?? a?? resident set himself on fire and the situation got worse and worse.)

Student 17 (Taiwanese):

- (2) "On the **one** hand, when I do **one** thing. I may have different troubles."

(For example, when I do a thing?? that thing?? something??, I may have various problems.)

- (3) "Nevertheless, I think perseverance is **one** especially important possession in my life."

(Nevertheless, I think perseverance is one of the most?? an?? especially important quality that I possess.)

In Student 15's example, *one* could be marking a specific, indefinite referent, which would fit with Robertson's theory, or it could be quantificational and mean that only one resident protested through self-immolation. For Student 17, "one thing" could mean *something*, *a thing*, or *that thing*, but given the context, *a thing* is the least likely reading

of *one*. The sentence before (2) is, “In this essay, I will argue in favor of the most important possession is perseverance.” So, (2) is actually an example or supporting point of the main purpose of the essay stated in the sentence before it. The topic focuses on doing and persevering, so introducing a new discourse referent with *a thing* would not make sense here. The student returns to the topic of doing within the same paragraph, but never specifies what is being done: “Troubles will make me feel very sad and I won’t to do anymore.” *Anything* or *something* are the most appropriate interpretations of the semantics of *one thing* here.

As a result of stating the meaning of *one* in (2) is *something*, there is only one example that seems to fit Robertson’s discourse rule, (3) above from Student 17. If only the Chinese/Taiwanese students are considered, this is one out of twelve total uses of *one* with any meaning or function. There are 36 uses of *a/an* in these same essays. So, even if there is one example that might support Chinese students using *one* in place of *a/an* in L2 English, this does not seem to happen enough to truly consider this a pattern rather than an infrequent idiosyncratic use.

Comparison of IIEP and EPE Student Sub-groups and Korean and Chinese/Taiwanese Students

Similar to findings seen often in this study, there are significant differences between IIEP and EPE students in some categories, but no differences between the Korean and Chinese/Taiwanese students regarding the frequencies of their use of articles, demonstratives, and pronouns. See the statistical summary in the table that follows.

The IIEP students and EPE students are significantly different in their use of *the*, articles all together, and pronouns. In fact, the EPE students are using *the* and articles as a whole more than the IIEP students are. The EPE students are averaging 14.5 uses of *the* per essay, and 17.8 total articles per essay. This is significantly higher than the average nine uses of *the* and 12.9 total articles in the IIEP essays. These significant

differences arise even when the absolute number is controlled for the number of words per essay.

This conclusion may only partly be true, though, as the IIEP students have significantly more pronouns than the EPE students, both when just the counts are considered ($z = 2.95, p = .0032$) and when these counts are normalized for the number of words per essay ($z = 2.87, p = .0042$). The IIEP students are averaging 26.4 pronouns per essay, over twice the average number in the EPE students' essays, which is 10.5. (See the summarized averages of the significant differences found in the table at the end of this section.) This could be yet one more effect of the essay topic on the discourse construction. The IIEP students were asked to write about themselves, which would require or allow more pronouns than the formal essays on impersonal topics. The impersonal topics would require more definite description NPs.

This does not explain what might be happening with the number of instances of *a/an*. When the IIEP and EPE students are compared for frequency of *a/an*, there is no significant difference. Neither is there one between the Korean and Chinese/Taiwanese students. However, there was a significant difference between the native and the student use of *a/an*.

Conclusions about Articles, Demonstratives, and Pronouns

The most interesting differences in the use of articles, demonstratives, and pronouns are:

- The native writers use *a/an* significantly more often than the students. For the students, no significant differences were discovered between the sub-groups. Robertson (2000) does predict a lower use of *a/an* by L2 English students, but predicts this is due to substitution of *a/an* with *one* by L1 Chinese students. Examples from the student essays, though, show only one instance when this may

in fact be occurring. Other uses of *one* are for quantification, pronominalization, or to mean *individual*.

Table 79: Summary of Statistical Findings for Differences between Groups in Articles, Demonstratives, and Pronouns, Wilcoxon Rank Sums Test

Discourse feature	Student vs. Native writers		IIEP vs. EPE students		Korean vs. Chinese/ Taiwanese students	
	<i>z</i>	<i>p</i> ($\alpha = .10$)	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
# <i>a/an</i>	-4.36	< .00003*	.68	.4966	-.91	.3628
# <i>a/an</i> / # words	-3.00	.0026*	-.08	.9362	-.60	.5486
# <i>the</i>	-1.30	.1936	-2.12	.0348*	.08	.9362
# <i>the</i> / # words	1.60	.1096	-2.12	.0348*	.15	.8808
total # articles	-2.43	.0150*	-1.81	.0702*	-.38	.7040
# articles/ # words	.16	.8728	-2.27	.0232*	-.23	.8180
# demonstratives	-3.52	≤ .0004*	-.45	.6528	.45	.6528
# demonstratives/ # words	-2.11	.0348*	-.98	.3270	.23	.8180
# pronouns	-.03	.9760	2.95	.0032*	.15	.8808
# pronouns/ # words	1.79	.0734*	2.87	.0042*	-.38	.7040

Shaded boxes indicate no significant difference was found. * indicates significance above $\alpha = .10$.

- The native and student writers do not significantly differ in the frequency of use of *the*. However, the IIEP and EPE students do significantly differ in this area. This could be a result of the demands on the NP structure by the topic. The IIEP students wrote on more personal topics which would be more likely to support pronoun use, and they did in fact use significantly more pronouns than the EPE students.

- Student writers had significantly fewer demonstratives than native writers.
- There were more significant differences in usage when comparing the IIEP and EPE students that when comparing the Chinese/Taiwanese and Korean students. Also, an effect of the essay topics again seemed to influence these sub-group differences.

The next section will look at the frequencies of common nouns, proper nouns, and other kinds of nominals.

Table 80: Group Averages for Significant Differences Found in Articles, Demonstratives, and Pronouns

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/Taiwanese students</i>	
	<i>student average</i>	<i>native average</i>	<i>IIEP average</i>	<i>EPE average</i>	<i>Korean average</i>	<i>Chinese/Taiwan average</i>
<i># a/an</i>	3.6	9.7*				
<i># a/an/ # words</i>	.0165	.0300*				
<i># the</i>			9	14.5*		
<i># the/ # words</i>			.0404	.0690*		
<i>total # articles</i>	15.35	23.35*	12.9	17.8*		
<i># articles/ # words</i>			.0582	.0842*		
<i># demonstratives</i>	4.45	9.45*				
<i># demonstratives/ # words</i>	.0206	.0279*				
<i># pronouns</i>			26.4*	10.5		
<i># pronouns/ # words</i>	.0893*	.0516	.1280*	.0520		

Darkly shaded boxes indicate no significant difference was found. * indicates the higher average.

Proper Nouns, Common Nouns, and Other Kinds of Noun

Phrases

In this section, “other NPs” is a category for lexical items such as pronouns and bare demonstratives. The number of common nouns, proper nouns, and other kinds of NP were counted for these reasons:

- One of the EPE essay topics required citing country names from a graph. If the EPE students were found to have more proper nouns than the IIEP students, then that would be another indicator that essay topic played a large part in the use of different kinds of nominals in the L1 and L2 writing examined here.
- Anaphoric hierarchies, one of the alternative models of discourse, rank proper nouns, common nouns, and pronouns/bare demonstratives at different levels of activation in memory. These hierarchies are presumed to be universal. Pronouns and bare demonstratives are ranked as being highly activated in working memory, or closer to the center of attention. Common and proper nouns are generally low on anaphoric hierarchies because a lot of semantic information about the referent is included in their forms. First occurrences are usually lower on the scale, with later mentions ranking higher. If students are using overly informative reference, which is a pattern than has been seen in L2 production, then the students may have more common and proper nouns per total number of words than the native writers.

The table that follows on the next page shows the number of common nouns, proper nouns, and other NPs in each essay. For the common nouns, the number appears to be linked to essay length. The shortest student essay has the fewest common nouns (Student 18, 19 common nouns), and the longest student essay has the most (Student 2, 90 common nouns). The same is true for the native writers. Native 21, with the shortest essay, has the fewest common nouns (26 nouns), and Native 32, with the longest essay, has the most common nouns for the native writers (141 common nouns).

Table 81: Number of Common Nouns, Proper Nouns, and Other NPs per Essay

<i>participant #</i>	<i># common nouns</i>	<i># proper nouns</i>	<i># other NPs</i>
1	43	1	21
2	90	13	8
3	46	21	7
4	38	32	4
5	48	8	7
6	44	37	8
7	41	5	15
8	40	12	5
9	43	31	6
10	37	15	7
11	50	1	33
12	32	7	62
13	37	1	36
14	43	0	33
15	56	2	38
16	32	2	28
17	30	0	24
18	19	1	20
19	46	0	20
20	36	3	25
21	26	1	4
22	85	1	15
23	32	0	19
24	55	0	8
25	89	6	12
26	55	4	42
27	52	0	9
28	58	2	13
29	87	11	29
30	74	1	42

Table 81—Continued

<i>participant #</i>	<i># common nouns</i>	<i># proper nouns</i>	<i># other NPs</i>
31	78	6	30
32	141	7	49
33	66	2	34
34	55	2	19
35	76	1	8
36	106	3	38
37	64	0	18
38	100	6	24
39	61	0	23
40	38	2	5

The length of the essay seems related to the number of “other” NPs for the native writers, but not for the students. Native 32 has the longest essay, and the most other kinds of NPs (49 of them). Native 21 has the fewest other kinds of NP, with only 4, and also has the shortest native essay. For the students, Student 4 has the fewest other kinds of NP with only 4, and Student 12 has the most, with 62 instances of pronouns, bare demonstratives, and the like. However, neither of these students had the shortest or longest essays.

Another difference between the student and native groups is that for both the common nouns and the other kinds of NP, the native writers have a wider range of values, and higher averages. See the averages summarized in the table on the next page. These have not been normalized by the number of words per essay.

In contrast, the native writers have fewer proper nouns on average than the student writers. Five native writers have no proper noun phrases in their essays at all (Natives 23, 24, 27, 37, and 39).

Table 82: Average Number of Common Nouns, Proper Nouns, and Other NPs for Native and Student Writers

	<i>ESL Students averages</i>		<i>Native writers averages</i>	
	<i>averages</i>	<i>medians</i>	<i>averages</i>	<i>medians</i>
<i>common nouns</i>	42.55	42	69.9	65
<i>proper nouns</i>	9.6	4	2.75	2
<i>other NPs</i>	20.35	20	22.05	19

The native writer with the most proper NPs is Native 29, who has only eleven proper nouns. However, the student writer with the highest number of proper nouns has three times this amount (Student 6, 37 proper nouns). Student 6 took the EPE test, and selected the topic of describing the population graph, so this is not unexpected. The proper nouns are all country names. The IIEP students (Students 11-20), who did not have the option of this topic, have much fewer proper nouns per essay and are more similar to the native writers in this case. Three of the IIEP students have no proper nouns at all (Students 14, 17, and 19). So, again, topic is influencing the use of noun phrases and other discourse relevant lexical items.

When the student and native writers are compared using the Wilcoxon Rank Sums Test, significant differences are found in the number of common nouns and in the number of proper nouns. For the common nouns, there is a significant difference in both the total number of common nouns ($z = -3.65, p = .0004$), and for the number of common nouns per words in the essays ($z = -4.06, p \leq .00006$). It is not clear what would cause this difference, unless it is the fact that perhaps the use of more proper nouns by the students left fewer opportunities to use common nouns. For the proper nouns, the only significant difference was found when comparing the ratio of the number of proper nouns

to the total number of words per essay ($z = 2.27, p = .0232$). Given the large amount of proper nouns in the EPE essays, this is not unexpected.

When the EPE and IIEP essays are compared, significant differences were found in the ratio of the number of common nouns to the number of words, the total number of proper nouns, the ratio of proper nouns to words, the total number of other kinds of NPs, and the ratio of other NPs to the number of words per essay. These results are summarized in the next table.

Table 83: Summary of Statistical Findings for Differences between Groups in Common Nouns, Proper Nouns, and Other NPs—Wilcoxon Rank Sums Test

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/ Taiwanese students</i>	
	<i>z</i>	<i>p</i> ($\alpha = .10$)	<i>z</i>	<i>p</i> ($\alpha = .10$)	<i>z</i>	<i>p</i> ($\alpha = .10$)
<i># common NPs</i>	-3.65	.0004*	-1.29	.1970	1.81	.0702*
<i># common NPs/ total # NPs</i>	-4.06	$\leq .00006^*$	-1.89	.0588*	.76	.4472
<i># proper NPs</i>	1.16	.2460	-3.17	.0016*	-.08	.9362
<i># proper/ total # NPs</i>	2.27	.0232*	-3.40	.0006*	-.08	.9362
<i># other NPs</i>	-.73	.4654	3.63	.0004*	.76	.4472
<i># other/ total # NPs</i>	.68	.4966	3.70	.0002*	-.30	.7642

Shaded boxes indicate no significant difference was found. * indicates significance above $\alpha = .10$.

The EPE students average more common nouns and more proper nouns, but the IIEP students have more other kinds of NPs. As has been discussed, these differences may relate to differences in writing topics more than being caused by proficiency differences.

Examining essays on a controlled topic designed to elicit specifically pronouns, common nouns, or proper nouns would be an area of potential further research.

When comparing the Korean and the Chinese/Taiwanese students, as with other features, very little significant difference was discovered. In fact, the only significant difference was in the number of common nouns ($z = 1.81, p = .0702$). There was no significant difference in the ratio of common nouns to the total number of words. It is not clear why the Korean students have more total common nouns than the Chinese/Taiwanese students. The table on the next page summarizes the averages of the groups for the significant differences that were found.

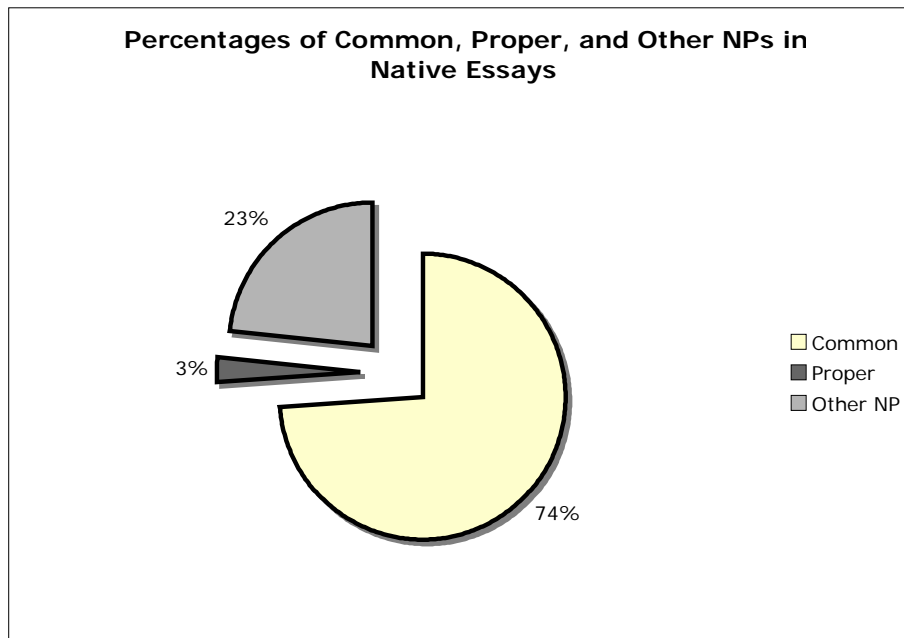
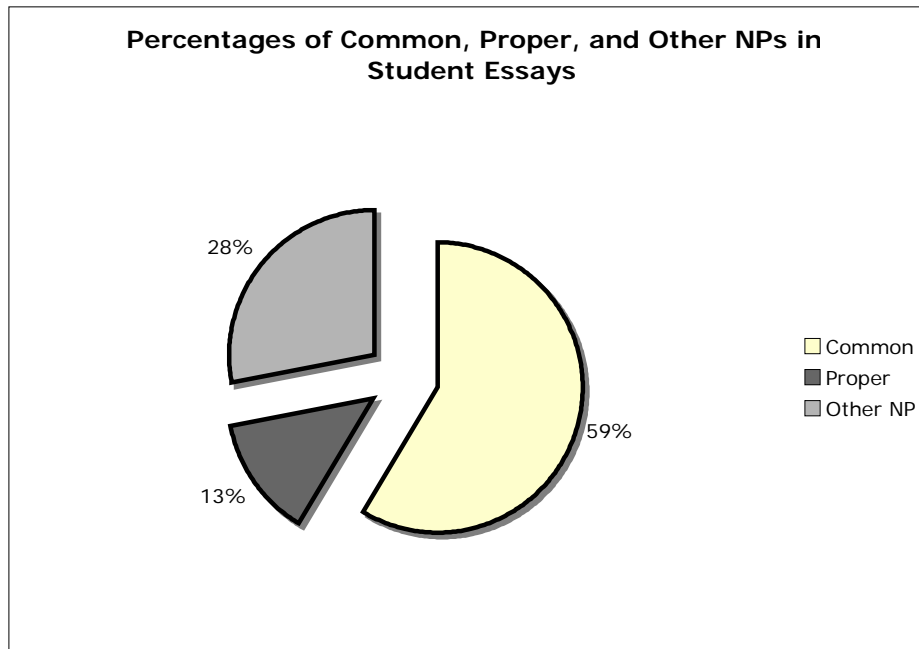
The next section summarizes what was found regarding research goals one and two.

Table 84: Group Averages for Significant Differences Found in Common Nouns, Proper Nouns, and Other NPs

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/Taiwanese students</i>	
	<i>student average</i>	<i>native average</i>	<i>IIEP average</i>	<i>EPE average</i>	<i>Korean average</i>	<i>Chinese/Taiwan average</i>
<i># common NPs</i>	42.55	69.9*			48.2*	36.9
<i># common NPs/ total # NPs</i>	.5885	.7514*	.5367	.6402*		
<i># proper NPs</i>			1.7	17.5*		
<i># proper/ total # NPs</i>	.1273*	.0259	.0220	.2326*		
<i># other NPs</i>			31.9*	8.8		
<i># other/ total # NPs</i>			.4413*	.1272		

Darkly shaded boxes indicate no significant difference was found. * indicates the higher average.

Figure 29: Percentages of Proper Nouns, Common Nouns, and Other Kinds of NPs in Student and Native Essays



Summary of Significant Findings, Research Goals One and

Two

Goal 1: To describe some characteristics of co-reference and discourse construction in native and L2 essays in English.

Goal 2: To determine where there are significant differences in co-reference and discourse construction between (a) native and L2 writers, (b) L2 writers from China and Korea, and (c) L2 writers of different proficiency levels.

Several statistical comparisons of the native and student writers, lower proficiency IIEP students and higher proficiency EPE students, and the Korean and Chinese/Taiwanese students. These groups were compared by looking at the number of words in their essays, the number of noun phrases, co-reference chains, noun phrase transitions and more features of their essays that might give indications of their overall skill at and manner of creating coherent discourse in their essays. The two tables at the end of this section summarize the significant findings.

Native English vs. L2 English/Student Writers

When comparing the native and student writers, some of the significant findings were expected. For example, the native English writers had longer essays, and a wider variety of vocabulary words. In addition, because of the nature of one of the EPE essay topics requiring the use of several proper nouns, the native writers had more common nouns while the students averaged more proper nouns. Furthermore, because the IIEP essay topics required more discussion of personal experiences, the students were found to use more pronouns. The most commonly used pronoun for either group was *I*.

Some of the differences between the native and student writers were unexpected. For instance, the student writers had more NP subjects and more complements of copulas than the native writers. Regarding the copulas, it is possible that students have more copula verbs as an effect of explicit grammatical instruction in the classroom. Both Korean and Chinese students often omit the copula, or use articles incorrectly for copula

complements. There are therefore routinely covered in ESL grammar books of various levels. This could make these constructions more familiar. Alternatively, students could acquire them earlier or find them easier to use than other constructions. This is an area for further research.

Two surprising differences were found between the native and student writers that specifically dealt with determiners and articles. The native writers used *a/an* significantly more than the students, and used more articles. The native writers also used more demonstrative determiners. Other research on articles in L2 English has found or predicts an overuse of *a/an*, in contrast to what was encountered in these essays. It is also interesting that there were no significant differences found in the use of *the* by the two groups. Other research has suggested that Chinese students should omit *the*, or overuse it in certain contexts. This topic will be returned to in the error analysis.

IIEP (Lower Proficiency) vs. EPE (Higher Proficiency)

Students

When looking at the significant differences between the IIEP and EPE students, some seem to indicate differences in discourse structure based on essay topic, and some seem due to proficiency differences. The IIEP students had significantly more pronouns, and maintained transitions. They also had longer co-reference chains. These differences could result from the fact that their essays required personal storytelling. The students used pronouns to talk about themselves, and had fewer key referents to include based on the topic. The EPE students had more new transitions because they had to introduce more referents in order to give examples that supported their essay arguments.

The differences in the number of co-reference chains, the syntactic positions of NPs, and articles could indicate differences in proficiency or in essay topic. The EPE students had more co-reference chains, and more articles. The fact that they used more *the* could indicate that the EPE students are better with its use. Recall that there was no

significant difference when comparing all the students to the native writers in their use of *the*. That the IIEP and EPE students do not differ in their use of *a/an* indicates that both groups are using *a/an* similarly, but less often than the native writers. Another possible indicator of proficiency differences is that the IIEP students had more NP subjects and verbal objects. A sentence of the structure NP subject-verb-NP object is one of the simplest syntactic clauses. This result seems to imply that the IIEP students are using more simple sentences, and fewer prepositional phrases and relative clauses.

Korean vs. Chinese/Taiwanese Students

There were very few differences found between the Korean and Chinese/Taiwanese student groups, signifying that L1 (or this pair of L1s) is not as important as proficiency regarding differences in use of L2 English. Significant differences between the two first language groups were only found in the number of common nouns, the number of noun phrases, and the concentration of the number of copula complements per essay. In all three cases, the Korean students had higher averages than the Chinese/Taiwanese students.

Table 85: Summary of Statistically Significant Differences by Group

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/ Taiwanese</i>	
	<i>z</i>	<i>p</i> ($\alpha = .10$)	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
<i>total # words</i>	-3.03	.0024*	-.08	.9362	.98	.3720
<i>total # NPs</i>	-2.22	.0264*	.08	.9362	2.27	.0232*
<i>ratio # NPs/ # words</i>	3.81	$\leq .0001^*$.30	.7642	1.44	.1498
<i>average co-ref chain length</i>	.68	.4966	3.78	.0002*	.23	.8180
<i>total # co-ref chains</i>	-2.03	.0424*	-1.89	.0588*	.98	.3270
<i>ratio # co-ref chains/ # words</i>	1.41	.1586	-2.12	.0340*	.08	.9362
<i>ratio # co-ref chains/ # NPs</i>	-.03	.9760	-2.27	.0238*	-.83	.4066
<i>total # new transitions</i>	-3.16	.0016*	-2.57	.0102*	1.59	.1142
<i>ratio # new transitions/ # words</i>	-.08	.9362	-2.26	.0238*	-.83	.4066
<i># maintained transitions</i>	.08	.9362	3.02	.0026*	.38	.7040
<i>ratio # maint. transitions/ # words</i>	2.94	.0032*	3.25	.0012*	0	1
<i># re-mentioned transitions</i>	-1.68	.0950*	.30	.7642	1.20	.2302
<i>ratio # re-ment. transitions/ # words</i>	.95	.3422	.23	.8180	1.59	.1118
<i># subjects</i>	-1.54	.1236	2.34	.0192*	1.36	.1738
<i># subjects/ # words</i>	2.89	.0038*	3.02	.0026*	.60	.5486
<i># verbal objects</i>	-2.29	.0220*	2.27	.0232*	.98	.3270
<i># verbal objects/ # words</i>	.65	.5156	2.04	.0414*	.53	.5962
<i># objects of prepositions</i>	-2.79	.0054*	-1.21	.2262	-.45	.6528
<i># prep. objects/ # words</i>	-.70	.4840	-1.29	.1970	-.68	.4966
<i># genitive specifiers</i>	-.86	.3898	.91	.3628	.45	.6528
<i># genitive specifiers/ # words</i>	1.08	.2802	.60	.5486	.08	.9362

Table 85—Continued

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/ Taiwanese</i>	
	<i>z</i>	<i>p</i> ($\alpha = .10$)	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
# copula complements	-4.1	.6818	1.29	.1970	1.59	.1118
# copula comp.s/ # words	2.03	.0424*	.91	.3628	1.81	.0702*
# appositions	-.62	.5352	-1.51	.1310	-.15	.8808
# appositions/ # words	.70	.4840	-1.59	.1118	-.60	.5486
# word types	-3.90	.0001*	.38	.7040	1.21	.2262
# types/ # words	-.51	.6100	.68	.4966	0	1
# single use types/ total # types	-2.08	.0376*	.08	.9362	.08	.9362
# double use types/ total # types	3.08	.0020*	.45	.6528	-.53	.5962
# a/an	-4.36	< .00003*	.68	.4966	-.91	.3628
# a/an/ # words	-3.00	.0026*	-.08	.9362	-.60	.5486
# the	-1.30	.1936	-2.12	.0348*	.08	.9362
# the/ # words	1.60	.1096	-2.12	.0348*	.15	.8808
total # articles	-2.43	.0150*	-1.81	.0702*	-.38	.7040
# articles/ # words	.16	.8728	-2.27	.0232*	-.23	.8180
# demonstratives	-3.52	≤ .0004*	-.45	.6528	.45	.6528
# demonstratives/ # words	-2.11	.0348*	-.98	.3270	.23	.8180
# pronouns	-.03	.9760	2.95	.0032*	.15	.8808
# pronouns/ # words	1.79	.0734*	2.87	.0042*	-.38	.7040
# common NPs	-3.65	.0004*	-1.29	.1970	1.81	.0702*
# common NPs/ total NPs	-4.06	≤ .00006*	-1.89	.0588*	.76	.4472
# proper NPs	1.16	.2460	-3.17	.0016*	-.08	.9362
# proper/ total # NPs	2.27	.0232*	-3.40	.0006*	-.08	.9362
# other NPs	-.73	.4654	3.63	.0004*	.76	.4472
# other/ total # NPs	.68	.4966	3.70	.0002*	-.30	.7642

Table 86: Summary of Group Averages for Significantly Different Findings

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/ Taiwanese</i>	
	<i>student average</i>	<i>native average</i>	<i>IIEP average</i>	<i>EPE average</i>	<i>Korean average</i>	<i>Chinese/ Taiwan</i>
<i>total # words</i>	214.35	332.5*				
<i>total # NPs</i>	72.55	94.9*			80.9*	64.2
<i>ratio # NPs/ # words</i>	.3420*	.2907				
<i>average co-ref chain length</i>			5.83*	3.01		
<i>total # co-ref chains</i>	10.95	13.8*	8.7	13.2*		
<i>ratio # co-ref chains/ # words</i>			.0425	.0617*		
<i>ratio # co-ref chains/ # NPs</i>			.1257	.1782*		
<i>total # new transitions</i>	39.05	60.25*	31.8	46.3*		
<i>ratio # new transitions/ # words</i>			.1549	.2152*		
<i># maintained transitions</i>			32.1*	17.6		
<i>ratio # maint. transitions/ # words</i>	.1171*	.0748	.1522*	.0820		
<i># re-mentioned transitions</i>	8.55	10.55*				
<i>ratio # re-ment. transitions/ # words</i>						
<i># subjects</i>			26.2*	18.5		
<i># subjects/ # words</i>	.1057*	.0795	.1260*	.0855		
<i># verbal objects</i>	14.65	22.1*	17.1*	12.2		
<i># verbal objects/ # words</i>			.0821*	.0580		
<i># objects of prepositions</i>	22.75	32.3*				
<i># prep. objects/ # words</i>						
<i># genitive specifiers</i>						
<i># genitive specifiers/ # words</i>						

Table 86—Continued

<i>Discourse feature</i>	<i>Student vs. Native writers</i>		<i>IIEP vs. EPE students</i>		<i>Korean vs. Chinese/ Taiwan</i>	
	<i>student average</i>	<i>native average</i>	<i>IIEP average</i>	<i>EPE average</i>	<i>Korean average</i>	<i>Chinese/ Taiwan</i>
<i># copula complements</i>						
<i># copula comp.s/ # words</i>	.0250*	.0115			.0374*	.0126
<i># appositions</i>						
<i># appositions/ # words</i>						
<i># word types</i>	109.05	169.55*				
<i># types/ # words</i>						
<i># single use types/ total # types</i>	.6449	.7002*				
<i># double use types/ total # types</i>	.1792*	.1392				
<i># a/an</i>	3.6	9.7*				
<i># a/an/ # words</i>	.0165	.0300*				
<i># the</i>			9	14.5*		
<i># the/ # words</i>			.0404	.0690*		
<i>total # articles</i>	15.35	23.35*	12.9	17.8*		
<i># articles/ # words</i>			.0582	.0842*		
<i># demonstratives</i>	4.45	9.45*				
<i># demonstratives/ # words</i>	.0206	.0279*				
<i># pronouns</i>			26.4*	10.5		
<i># pronouns/ # words</i>	.0893*	.0516	.1280*	.0520		
<i># common NPs</i>	42.55	69.9*			48.2*	36.9
<i># common NPs/ total # NPs</i>	.5885	.7514*	.5367	.6402*		
<i># proper NPs</i>			1.7	17.5*		
<i># proper/ total # NPs</i>	.1273*	.0259	.0220	.2326*		
<i># other NPs</i>			31.9*	8.8		
<i># other/ total # NPs</i>			.4413*	.1272		

CHAPTER IV. RESULTS, RESEARCH GOAL THREE

In this section, data and conclusions analyzed to satisfy the third research goal will be examined. This research goal is:

Goal 3: To discover patterns of error in the selection of articles, determiners, pronouns, and other lexical items contributing to co-reference and discourse cohesion.

In the previous section, one interesting finding was that there was no significant difference between the student and native writers in the frequency of use of *the*. In this section, errors made by students will be examined to determine if there are, for example, substantive differences in how *the* is used despite their being no real frequency difference.

Overall Number of Errors in NPs

The number of errors in NPs and the percentage of the total number of NPs with errors is summarized in the next table given. When looking at the IIEP students, Student 18 (from Taiwan) had the fewest errors in NPs with only five, while Student 15 (from Korea) had the most NP errors in that group, with a total of 26. When the EPE students are considered in comparison, it can be seen that they have more total NP errors than the IIEP students. However, the EPE students tended to write longer essays, so there were more opportunities to make a mistake—meaning more NPs written. The EPE student with the most NP errors was Student 2 (from Korea), who had a total of 43 NP errors, close to twice the number of errors made by the IIEP student with the most total errors in NPs. The EPE students with the fewest NP errors also had about twice the number of total errors than the IIEP student with the fewest. Students 7 and 9 (both from China) tied for the fewest number of NP errors for the EPE students, with only twelve of them each, but this is still over twice the number of mistakes made by most accurate IIEP student, Student 18.

Looking at the percentages of errors, the EPE students still have more mistakes than the IIEP students, but there is not as wide a margin between the two groups. The EPE students on average had errors in 31% of their noun phrases, with 69% on average being acceptable (with standard deviation of .09 and median of 32%). In contrast, the IIEP students only averaged having errors in 22% of their total noun phrases, having on average 78% of their NPs be acceptable (with standard deviation of .07 and median of 21%). This is surprising because the EPE students have all scored high enough on the TOEFL to be admitted to the University of Iowa¹, while the IIEP students have not, and in general are at a lower English proficiency level than the EPE students.

It is also surprising that there is some noticeable difference in the number of errors between the two language backgrounds. Up to this point, very few differences have been seen between the Korean and the Chinese student sub-groups. However, the Korean students had a higher average percentage of NP errors than the Chinese and Taiwanese students. The Korean students averaged errors in 29% of their NPs, while the Chinese and Taiwanese students averaged errors in only 23% of their noun phrases. See the next table for individual values.

A Note on Optionality

When looking at the range of the percentages of error presented in the table above, the concept of optionality in the production of language learners comes to mind. In the past in child language acquisition studies, some authors have chosen a particular percentage of correct forms as an indication that a certain grammatical feature has been acquired.

¹ Please refer back to the beginning of the section on data collection for the exact requirements for admittance.

Table 87: Summary of Number of Noun Phrases with Problems in Student Essays

<i>participant</i>	<i>total # NPs</i>	<i>total # NPs with problems</i>	<i>% acceptable NPs</i>	<i>% problematic NPs</i>
5, EPE, Korea	63	25	60%	40%
1, EPE, Korea	65	22	66%	34%
2, EPE, Korea	111	38	66%	34%
19, IIEP, Taiwan	66	22	67%	33%
20, IIEP, Korea	64	19	70%	30%
4, EPE, Korea	74	21	72%	28%
10, EPE, China	59	16	73%	27%
15, IIEP, Korea	98	26	73%	27%
6, EPE, China	89	23	74%	26%
16, IIEP, Taiwan	62	16	74%	26%
3, EPE, Korea	73	17	77%	23%
8, EPE, Taiwan	58	12	79%	21%
12, IIEP, Korea	101	21	79%	21%
17, IIEP, Taiwan	54	11	80%	20%
13, IIEP, Taiwan	73	13	82%	18%
7, EPE, China	61	10	84%	16%
11, IIEP, Korea	84	12	86%	14%
14, IIEP, Korea	76	11	86%	14%
9, EPE, China	80	10	87%	13%
18, IIEP, Taiwan	40	5	87%	13%

For example, maybe it can be said that a learner has mastered the use of determiners and NPs in discourse in the new language when 80% of such forms are correct.

The problem is that there is no intrinsic reason why 80% should be chosen over some other level of correctness. For instance, one might say that a learner has acquired a structure and its meanings after they no longer perform at chance or randomly. This might indicate a threshold of maybe 51% of the forms correct. On the other hand, maybe

this threshold is too low. Perhaps it would be better to say that getting 80% of the NPs correct indicates mastery. However, that would be saying that 20% of the forms are incorrect due to something other than understanding the structure of NPs and the meanings of the functional morphemes used in them. This is clearly a much higher percentage than what would be seen in native writer production.

This leads to a consideration of other complications inherent in using such a system to evaluate mastery of a form. First, it is not clear what ultimate state of attainment might be common or possible for the majority of adults. What is the best that someone could possibly get? If due to cognitive considerations or the structure of the adult brain after a first language is acquired it is only possible to ever be at most 90% correct in the use of determiners and the structure of noun phrases, then 80% would indicate much higher proficiency than if 100% accuracy were possible. However, it has long been discussed and observed that there is a wide range in the abilities of adults to learn a new language. Some students learn languages to the point that they are proficient enough to be almost indistinguishable from monolingual native speakers of the target language. In contrast, others stop improving and seem to fossilize at a much lower skill level in manipulating the new grammar and forms. Despite knowing that what is ultimately possible for any particular individual may be very different from that for others, no answer has been found yet regarding exactly how to determine the ultimately possible level of attainment of an individual based on their intelligence, cognitive skills, language acquisition device, or whatever determines this in adults.

At this point, then, we are led back to the long-considered issue of the difference between first or child language acquisition and adult acquisition. The same individuals that vary so widely in adult language learning end up being fairly indistinguishable in their first languages. There is clearly a fundamental difference between comparing the language capability of native and non-native speakers, though. Native speakers vary in their ability to manage standard written forms formally taught in school and they differ in

the size of their vocabularies and ability to spell or express their ideas coherently. Unless they suffer from cognitive deficits, native English speakers all know what is English (syntax) for their dialect, what is not, and what is questionable. Adult learners of English are obviously learning this piece that native speakers already know. However, differences in the language skills of native speakers are recognized and exploited on tests such as the SAT, LSAT, and GRE. In perhaps not the same sense as when used to describe non-native writers, native writers do not in fact share the same level of ultimately attained proficiency in manipulating and learning language.

This highlights that it is also of course a fallacy to state that monolingual English speakers do not make mistakes in their writing, or that they agree on what is “correct.” First, the fact that native speakers do make mistakes can be seen in the performance of some of the writers in this study. See the table that follows.

Table 88: Noun Phrase Errors in the Essays of Native English Writers

<i>participant</i>	<i>NPs w/errors</i>	<i>NPs correct</i>	<i>participant</i>	<i>NPs w/errors</i>	<i>NPs correct</i>
21, native	-- --	100% (32)	31, native	-- --	100% (114)
22, native	-- --	100% (101)	32, native	1% (2)	99% (195)
23, native	-- --	100% (51)	33, native	-- --	100% (103)
24, native	-- --	100% (63)	34, native	3% (2)	97% (74)
25, native	1% (1)	99% (106)	35, native	4% (3)	96% (82)
26, native	2% (2)	98% (99)	36, native	1% (2)	99% (146)
27, native	-- --	100% (61)	37, native	1% (1)	99% (81)
28, native	-- --	100% (73)	38, native	1% (1)	99% (128)
29, native	-- --	100% (131)	39, native	-- --	100% (82)
30, native	-- --	100% (116)	40, native	2% (1)	98% (44)

Almost half of the native writers made at least one mistake in the structure of noun phrases. The number of mistakes is not high, either considered separately or in comparison to the performance of the students in the study, but they are there. Additionally, perhaps the native writers would have been able to fix these errors if they had been given time to revise. However, they were tested under similar conditions to what the students experienced, and it is not known how many of the errors made by the students could have been corrected by the students if they had been given the opportunity. Moreover, the native speakers typing on computers may have a somewhat lower error rate than they would have had when writing if they had assistance from spelling or grammar checkers (they were not instructed not to use these tools). The students did not have access to such aids.

Second, it is well-known, but often glossed over, that native writers at times disagree on what is acceptable or clear for determiners and noun phrase structure in the context of a full piece of writing. This is one of the complications in error analysis. There is no “gold standard” for reference and discourse construction. Often more than one possibility would be acceptable—for example, *this thing* and *the thing* are often interchangeable as are *this thing* and *it*. Minor meaning differences in the focus of the discourse may result, but these substitutions do not always lead to miscommunication. A further complication is that while there are definitely structures that native writers agree do not sound like English, there are typically many ways to correct such errors. Therefore, classifying the mistakes and determining what a writer should have been targeting cannot be stated with complete certainty. The best examples illustrating these two points happen when native writers edit each others’ work. It is often seen that one writer selects a full definite NP with *the* for a maintained reference, which is then changed by a friendly reviewer to a pronoun, which is then changed by an editor to a demonstrative determiner plus full NP, and later the original writer changes it back to *the* plus NP before publication.

So, some errors are worse than others, and the definition of what is and is not *well constructed* English is different from the definition of what structures can be considered English at all, and this complicates the issue of error analysis and the evaluation of optionality in learners' language. Therefore, with all the complications described, the best way to try to make sense of the possible optionality in the student writing and what it says about their proficiency is to examine the patterns and actual forms used by the students in more detail. This will be done in the sections and sub-sections that follow.

An Examination of Possible Patterns in the Overall Error

Percentages

When the error percentages and other information given in the student table from the beginning of this section are taken together, at first glance there do not seem to be any salient characteristics that pattern with having either more or less frequent errors.

When reviewing the error percentages given in the table about students given previously in this section, there do not seem to be any patterns of characteristics that stand out: Both IIEP and EPE students have high, medium, and low numbers of errors. There are long essays with a high percentage of errors, and there are short one. Chinese, Taiwanese, and Korean students are mixed over the range of error percentages. When the Wilcoxon Rank Sums Test was used to examine this data, there was found to be no significant difference between the IIEP and EPE students' error rates ($z = -1.29, p = .1970$). However, there was a significant difference between the Korean students and the Chinese and Taiwanese students ($z = 1.66, p = .0970$). The Korean students made significantly more errors in their NPs than the Chinese and Taiwanese students did. Given that in the rest of the study differences between the two language backgrounds were rarely found, but differences between the two class groups was, this is very interesting. If there are so few differences in the discourse structure of the Korean and the Chinese/Taiwanese students, then it could be that the Korean students do not

understand determiner meaning and noun phrase structure as well. However, the significant differences between the two language groups were that the Korean students wrote significantly more NPs, had significantly more NPs as the complements of copulas, and had significantly more common NPs. The simple fact that the Korean students wrote more NPs means that they had more of an opportunity to make mistakes, but in addition, the position of copula complement is one that is a tricky situation for co-reference because the referent is new, but not new, and linked, but not linked. More detail about whether these possibilities are true will be examined later.

Aside from a statistical analysis, another way to search for patterns in the error percentages is to compare them to the other evaluations of the students that are available. For example, the EPE students all have a TOEFL score and a list of classes that they were required to take after the EPE test that might be linked in some way to their rate of NP errors. The IIEP students have teacher ratings of their essays that can be examined, as well as the teachers' decision regarding what level in which to place the students. For both groups, there is the added information of what other kinds of syntactic and semantic errors were made. See the table that follows.

The information in this table shows that there is no strong connection between the number of errors in NPs and TOEFL scores, course placement, or writing rating. Many comparisons support this lack of connection. Looking at the EPE student with the fewest number of errors, Student 9, it seems there might be some connection between class placement and percentage of NP errors since Student 9 was not required to take any further ESL classes. However, if you look at the IIEP student with the lowest error percentage, this link appears to be broken: Student 18 only has errors in 13% of their essay's NPs, but was only placed in level two out of four. In contrast, IIEP Student 19 had an overall error rate of 33%, but was placed in the highest level, level four. There seems to be even less connection between class placement and error rates when Student 19 and Student 20 are compared.

Table 89: Error Percentages Compared with Other Judgments of Students' Proficiency

<i>participant</i>	<i>% problem NPs</i>	<i>writing score/ TOEFL score</i>	<i>IIEP level, ESL classes held for</i>	<i>other kinds of errors</i>
5, EPE, Korea	40%	247 [◊]	ESL Conversation, Grammar, Pronunciation, Reading, Writing	informal vocabulary, verb tenses, word choice
1, EPE, Korea	34%	217 [◊]	IIEP Writing [§] ESL Conversation, Grammar, Writing	prepositions, sentence structure, subject-verb agreement, word choice, word order
2, EPE, Korea	34%	220 [◊]	ESL Grammar, Reading	connecting words, prepositions, resultative verb constructions, sentence structure, subject-verb agreement, verb forms, verb tenses, word choice
19, IIEP, Taiwan	33%	4 [†]	level 4 [¶]	conditional sentences, sentence structure ^Δ , subject-verb agreement, verb tenses, word choice
20, IIEP, Korea	30%	3 [†]	level 2 [¶]	causative verb constructions, passive verbs
4, EPE, Korea	28%	243 [◊]	ESL Conversation, Reading, Writing	subject-verb agreement, verb tenses
10, EPE, China	27%	597 [€]	ESL Grammar, Writing	connecting words [‡] , subject-verb agreement
15, IIEP, Korea	27%	4 [†]	level 4 [¶]	causative verb constructions, relative clauses, sentence structure ^Δ , verb forms, verb tenses, word choice
6, EPE, China	26%	600 [€]	ESL Pronunciation	subject-verb agreement, verb tenses
16, IIEP, Taiwan	26%	unknown	level 3 [¶]	modal verbs, prepositions, verb tenses, word choice
3, EPE, Korea	23%	247 [◊]	ESL Conversation, Grammar, Pronunciation, Reading, Writing	missing conjunctions, sentence fragments, subject-verb agreement

Table 89—Continued

<i>participant</i>	<i>% problem NPs</i>	<i>writing score/ TOEFL score</i>	<i>IIEP level, ESL classes held for</i>	<i>other kinds of errors</i>
8, <i>EPE, Taiwan</i>	21%	220 [◇]	IIEP [§] Communication, Grammar, Reading, Writing	modal verbs, passive verbs, relative clauses, sentence structure, subject-verb agreement
12, <i>IIEP, Korea</i>	21%	4 [†]	level 4 [¶]	conditional sentences, passive verbs, relative clauses, sentence structure, verb forms, verb tenses
17, <i>IIEP, Taiwan</i>	20%	2 [†]	level 3 [¶]	connecting words, modal verbs, sentence structure, word choice
13, <i>IIEP, Taiwan</i>	18%	5 [†]	level 4 [¶]	sentence structure, verb tenses, word choice
7, <i>EPE, China</i>	16%	227 [◇]	ESL Conversation, Reading, Writing	verb tenses
11, <i>IIEP, Korea</i>	14%	6 [†]	level 4 [¶]	complex sentence structure* only, relative clauses, word forms
14, <i>IIEP, Korea</i>	14%	3 [†]	level 3 [¶]	conditional sentences, modal verbs, sentence structure, verb tenses, word form
9, <i>EPE, China</i>	13%	267 [◇]	none ^Ω	verb tenses, word order
18, <i>IIEP, Taiwan</i>	13%	3 [†]	level 2 [¶]	connecting words, prepositions, sentence structure, word choice, word form

◇ Computer-based TOEFL score.

∅ Paper-based TOEFL score.

† IIEP essays are scored holistically, meaning that a score is given for the overall quality or impression of the essay. Essays receive scores on a scale from one to six, with six being the highest score and one being the lowest. These essays are read and scored by two independent raters. A third rater's score is used when the first two raters do not agree.

Table 89—Continued

-
- § IIEP classes are generally lower in level than ESL classes. ESL classes usually are for advanced students. IIEP classes are usually for high-intermediate students and below.
- ¶ Students in the Iowa Intensive English program are usually divided into four levels, with level four being the most advanced, and level one being the least advanced. These levels correspond roughly to: Level 1—beginner, level 2—high beginner/ low intermediate, level 3—intermediate, level 4—high intermediate/ low advanced.
- Ω Not being held for any classes means that a student was not required to take any more English classes for non-native speakers before taking Rhetoric, beginning normal coursework, or graduating.
- Δ Sentence structure means, for example, missing subjects or objects, incorrect verb subcategorization.
- ‡ Connecting words consist of, for example, subordinating and coordinating conjunctions.
- * Complex sentence structure means joining two or more clauses.
-

Student 20 was also placed into level two, two levels below Student 19, but Student 20 had fewer NP errors and fewer other kinds of grammatical errors as well. Similar concerns about the link between class placement and error rates can be found when looking at the EPE students. For example, Student 2 and Student 1 both had NP error rates of 34%, but Student 2 was only required to take ESL grammar after testing, while Student 1 had to take IIEP writing, ESL writing, and ESL grammar.

These highlighted points are not meant to imply that the teachers did not rate the students well. The teachers were reacting to many different aspects of the students' productions when assigning classes and scores. NP errors were only a small part of what they had to consider. Regarding the last example, with Student 1 and Student 2, it could be that Student 2's NP errors did not interfere with the connection of meaning or discourse over the course of the essay. This could happen if, although a mistake or a

structure that was not English was recognized, the co-reference chains were not affected. Student 1's errors may have more seriously affected the linking of referents in the essay, requiring more instruction in writing and building discourse in a US English style. The TOEFL scores in fact do not match the error rates of the students very well either. While it is true that the EPE student with the lowest NP error rate had the highest computer-based TOEFL score of any of the EPE students at 267, Student 5, the student with the highest error rate (reaching 40%), had the second highest computer-based TOEFL score with a score of 247, and both were high enough for admission to the University of Iowa.

The next section will continue to explore the details of the errors and the patterns of NP use by looking more closely at co-reference chains.

Number of Errors in NPs in Co-reference Chains

In co-reference chains, the same real-world entity is referred to by at least two different NPs. The fact that these NPs in co-reference chains are connected makes them interesting to study in regards to NP form and determiners because the second or subsequent mentions of an NP with the same referent often are realized in a different form from the first mention (such as full NP, pronoun). In second language learners' writing, looking at co-reference chains may indicate whether the writer is choosing to use the same lexical form for every reference to a particular object/idea, or whether there are changes in the lexical form based on the position of the referent in the discourse relative to the other members of the same chain. The table that follows shows just the number and percentage of errors in NPs that are chained.

On average, there are fewer mistakes in the NPs in co-reference chains than there are when all the NPs are considered together. The students average mistakes in only 18% of the NPs that are in co-reference chains, versus 24% when all the NPs are considered together.

Table 90: Summary of Number of Noun Phrases in Student Essays in Co-reference

Chains that have Problems

<i>participant</i>	<i>total # NPs in chains</i>	<i>total # chained NPs with problems</i>	<i>% acceptable NPs in chains</i>	<i>% problematic NPs in chains</i>
20, IIEP, Korea	46	15	67%	33%
4, EPE, Korea	51	16	69%	31%
5, EPE, Korea	28	8	71%	29%
2, EPE, Korea	53	13	75%	25%
3, EPE, Korea	25	6	76%	24%
19, IIEP, Taiwan	35	8	77%	23%
1, EPE, Korea	36	8	78%	22%
6, EPE, China	73	16	78%	22%
16, IIEP, Taiwan	46	9	80%	20%
12, IIEP, Korea	82	15	82%	18%
14, IIEP, Korea	57	10	82%	18%
15, IIEP, Korea	64	11	83%	17%
10, EPE, China	28	4	86%	14%
17, IIEP, Taiwan	37	5	86%	14%
8, EPE, Taiwan	24	3	87%	13%
7, EPE, China	34	4	88%	12%
13, IIEP, Taiwan	49	5	90%	10%
9, EPE, China	47	4	91%	9%
18, IIEP, Taiwan	24	2	92%	8%
11, IIEP, Korea	45	2	96%	4%

When the Wilcoxon Rank Sums Test is used to compare the student sub-groups, just as with the situation when all the NPs are considered together, there is a significant difference between the two language backgrounds ($z = 2.27, p = .0232$), but not between the IIEP and EPE students ($z = -1.06, p = .2892$). The Korean students tended to have more errors in their chained NPs than the Chinese and Taiwanese students. However, the

table seems to indicate that all the students have at least some proficiency at selecting the appropriate morphemes, lexical forms, and syntax for NPs in chains. None of the students make mistakes in more than one-third of the NPs that they have linked by co-reference.

Number of Errors in NPs that are Not in Co-reference

Chains

When considering the unlinked NPs, a different story emerges. First, there is a wider range between the lowest error percentage and the highest for the mistakes in unchained NPs. Second, seven of the students have error rates of over 40% in unchained NPs. This means that about one-third to one-half of the students are getting close to 50% error rates, and it also seems to imply that students have more trouble selecting forms for referents that are not linked. It could be that the students do better with co-reference chains because the prior reference makes the selection of the next form easier somehow. Still, it is just as possible that the students are not altering their lexical forms and syntactic constructions for the subsequent references, and yet the form they stick with is correct for one of the contexts, lowering their error rates for chained NPs. More about this will be discussed shortly. The table that follows lists the number and percentages of errors and correct uses of unchained NPs.

When the Wilcoxon Rank Sums Test was used to compare the student sub-groups, no significant differences were found between either the IIEP and the EPE students ($z = -.38, p = .7040$) or the Korean and Chinese/ Taiwanese students ($z = .08, p = .9362$). So, while Korean students make significantly more mistakes in the number of errors in chained NPs, there is no corresponding difference when the unchained NPs are considered. However, it has often been mentioned thus far that there is a limited amount of information that can be gleaned from frequency data.

Table 91: Summary of Number of Noun Phrases in Student Essays NOT in Co-reference Chains that have Problems

<i>participant</i>	<i>total # unchained NPs</i>	<i>total # unchained NPs with problems</i>	<i>% acceptable unchained NPs</i>	<i>% problematic unchained NPs</i>
5, EPE, Korea	35	17	51%	49%
1, EPE, Korea	29	14	52%	48%
19, IIEP, Taiwan	31	14	55%	45%
2, EPE, Korea	57	25	56%	44%
6, EPE, China	16	7	56%	44%
15, IIEP, Korea	34	15	56%	44%
16, IIEP, Taiwan	16	7	56%	44%
10, EPE, China	31	12	61%	39%
12, IIEP, Korea	19	6	62%	38%
17, IIEP, Taiwan	17	6	65%	35%
13, IIEP, Taiwan	24	8	67%	33%
8, EPE, Taiwan	34	9	74%	26%
11, IIEP, Korea	39	10	74%	26%
3, EPE, Korea	48	11	77%	23%
7, EPE, China	26	6	77%	23%
4, EPE, Korea	23	5	78%	22%
20, IIEP, Korea	18	4	78%	22%
18, IIEP, Taiwan	16	3	81%	19%
9, EPE, China	33	6	82%	18%
14, IIEP, Korea	19	1	95%	5%

The next section will begin to look at specific examples and patterns of error in the essays of individuals in order to try to paint a fuller picture of what the students in this study did.

Common Error Patterns

In the students' essays, four different types of mistakes in NPs were most common. Two of these are determiner mistakes: the over-use of the determiner *the*, and the omission of the determiner *the*. One kind of mistake relates to understanding the semantics of the head noun *or* the ability to use plural functional morphemes accurately. These mistakes would be those that seem to be mistakes in the number of the head noun. Finally, students commonly made mistakes in the syntax of the noun phrases, such as in the word order.

Over-use of *the*

Despite the fact that in frequency counts students were not found to use *the* in an amount significantly from native writers, there were still a number of cases in which *the* was used when it was not necessary. The majority of the NPs with such errors are in generic statements, which are actually not commonly included in the discussion in previous studies or in discourse theories. The students with the highest number of errors due to over-using *the* were Student 5 (EPE, Korea), Student 12 (IIEP, Korea), Student 15 (IIEP, Korea), and Student 20 (IIEP, Korea). Although it seems from these four students that the over-use of *the* may be related to L1, there are not enough students included in this study to determine if this pattern is significant and true or visible only because of a small sample window limiting the perception of other possibilities. The next few subsections will discuss specific examples from the student essays.

Double Determiners

The most infrequent kind of over-use of *the* seen in the student essays was the use of *the* plus another determiner. In fact, there were only two instances of this in all twenty student essays. This infrequency suggests that these could be simply surface, production errors and not errors that indicate problems in the underlying structure of the noun

phrases. However, there is no way to determine this from these essays. In the examples that follow, it is interesting that *the* appears before the other determiner.

Examples 69: Double Determiners in Student Essays

- (a) I describe some of the our country's traditional things. (#20, Korea, IIEP)

I will describe some of our country's traditional things.

- (b) While the another² big country, India, will replace China to be the biggest country in the world. (#10, China, EPE)

While another big country, India, will...

Names

More commonly seen than the use of double determiners is the use of *the* with names that do not require it. There are names in English that are used with *the*, but these names follow certain patterns. In these patterns, *the* is used when the names are, for example:

- plural, as in *the United States*,
- describing areas or regions, as with *the Ukraine*, or
- the names of rivers or valleys or other items in which the last word is a singular count noun, as in *the Nile Valley*, *the Mississippi River*, or *the Industrial Revolution*.

As can be seen in the examples that follow, the names with which these students used *the* did not fit any of these categories.

² So, it is debatable whether *another* is really a determiner or not. Regardless, as a quantifier or whatever, it is still not a form that can be used with *a/an* or *the* in English.

Examples 70: Names Incorrectly Used with *the*

- (a) Moreover, he was very strict person and he wanted be a perfect person as the God. (#12, Korea, IIEP)

...he wanted to be as perfect a person as God.

- (b) For example, when I made my decision to come Iowa, I would be a little confused about some questions, such as where the Iowa is, how to go, and where to live. (#16, Taiwan, IIEP)

For example, when I made my decision to come to Iowa, I was a little confused about some questions, such as where Iowa was...

- (c) The Korea: I'm from Korea. My country is very beautiful place. ... I want you to go to the Korea. (#20, Korea, IIEP)

Korea: I'm from Korea. ... I want you to go to Korea.

- (d) First, Most famous thing is the food. Their are kimchi, bulgogi, and chongkujang---. For example, The Kimchi is most famous Korean food. (#20, Korea, IIEP)

First, the most famous thing is the food. There are foods like kimchi, bulgogi, and... For example, kimchi is the most famous Korean food.

In example (a) from Student 12, *the* is used with the singular name *God*. For Student 12, though, this formation is clearly not a production error, but a grammatical or semantic rule. *God* appears in Student 12's essay eight times, and every time it appears with *the*. There is no variation in the article choice with this word. In other students' essays, though, it is not clear which form is preferred, or whether it is the case that the students accept the word both with the determiner and without. The other three examples given above demonstrate this optionality. The words *Iowa* in (b), *Korea* in (c), and *kimchi* in (d) are all used both with and without *the*. In (b), *Iowa* and *the Iowa* are used in the same sentence. In these samples, there are no obvious patterns true for more than one student that would indicate why the article is used on one instance and not in another. Furthermore, there are not enough examples from one student to determine whether any

student's interlanguage grammar has a rule stating something like "use *the* before the name when it is a subject."

Generic Statements

Most of the instances of the over-use of *the* in fact occur in general statements, or when there is generic reference. See the examples that are given next.

Examples 71: Over-use of *the* in General Statements/ with Generic Reference

- (a) It is true for many people there are many things to make them confused or afraid. I am the same with these people because I consider that as a human being, we always are confused and afraid when we arrived in the new environment. (#16, Taiwan, IIEP)

...we are always confused and afraid when we arrive in a new environment.

- (b) Real activity is really essential to the change of society. In conclusion, the challenging spirit and real activity are pivotal to the life. (#15, Korea, IIEP)

...are pivotal to life.

- (c) So I would like to explain about the culture borrowing in the several aspects. (#1, Korea, EPE)

So, I would like to explain several aspects of culture borrowing.

- (d) A long time ago, our country didn't accept the culture from other countries. (#7, China, EPE)

A long time ago, our country didn't accept cultural influence/??culture from other countries.

In (a) through (c) above, there is no problem with the number of the referent or noun phrase. When the nouns are singular, they in fact refer to single entities, and when plural, they refer to plural entities. However, *the* does not fit the generic context. Native writers may prefer *a* in (a) because it is a generic singular count noun. In (b) and for the two NPs in (c), no article or determiner would be preferred. This is because of the

semantics of the lexical items and the referents: *Life* and *culture borrowing* are both generic referents used in an abstract or non-count manner, and *several aspects* is a plural count noun with generic reference. Example (d) is slightly different because there is a word choice error. It is not completely clear what idea the writer was trying to convey. However, the position of the sentence in the discourse, as the start of a new sub-topic in the discourse, and the lack of previous reference to cultures from other countries supports this being a general description. The meaning of generic statements generally does not mesh with the meaning of *the* unless the existence of only one instance of the entity referred to adds the definiteness to the NP. For example, in our galaxy, *the sun* is unique. Even in generic statements about the sun, *the* will be used because the real-world object builds the definite nature of the referent in spite of the generic meaning of the statement. The singularity over-rides the generic description.

The same optionality that was seen with the names above is also seen with these general statements, though. Students use *the* for one NP referring to an object, and in the same writing do not use it. See the next set of examples.

In the examples with the names above, at least one of the students attached the use of *the* to every instance of a particular lexical item (*God*), but with these general statements this is either not happening, or there are no chains long enough to indicate that any student is doing this. Using the alternative theories of discourse presented in the literature review of this study, the table below was made to look for commonalities in regards to which general referents appear with *the* unnecessarily.

Examples 72: Optionality Shown in the Use of *the* with the Lexical Item *baby*

- (a) Second, some advanced nations people tend to be reluctant to get the baby. (#5, Korea)

...some advanced nations' people tend to be reluctant to have babies/a baby.

- (b) Also, the other reason of the decreasing population is the weak economic growth and the people are afraid of giving birth to baby.

Another reason for the decreasing population is weak economic growth causing people to be reluctant to have children.

In looking at the table, these referents do seem to share common discourse features. Most are in chains of co-reference that span sections of the essays. Most are in the syntactic position of the object of a preposition, and most are what Smith (2003) calls a Primary Referent. A Primary Referent is an NP important to the semantics of a sentence and the movement of the discourse, but it is not the topic. They are typically in a patient theta-role, and they help to progress ideas in a discourse that is atemporal (meaning time cannot be used to propel the text construction). (See Smith 2003:17.) The problem is that these commonalities that can be seen by using the alternative theories of discourse do not really add to our understanding of the errors of the over-use of *the*. These features are shared because of the semantics of general statements, and do not explain why *the* would be selected over *a/an*, a demonstrative determiner, or no determiner.

Specific Indefinites

In this section, the view of specificity of Ionin, Ko, and Wexler (2003), which is based on Fodor and Sag (1982), will be used³.

³ Please refer to I.C.1.b.ii on page 20 and I.C.1.c.ii on page 27 for more information.

Table 92: Comparison of Features of NPs in Generic Statements with the Over-use of *the*

<i>NP/ participant</i>	<i>co-reference chain member</i>	<i>transition status</i>	<i>syntactic position</i>	<i>primary referent (Smith 2003)*</i>	<i>predicate type (Smith 2003)</i>	<i>discourse mode (Smith 2003)</i>
<i>the new environment (16, Taiwan, IIEP)</i>	yes	new/ first mention	object of preposition	yes (change of state)	General Stative	Argument
<i>the life (15, Korea, IIEP)</i>	yes	re-mention	object of preposition	yes (dependent on situation for existence)	General Stative	Argument
<i>the culture borrowing (1, Korea, EPE)</i>	yes	new/ first mention	object of preposition	yes (does not exist independent of the event)	Abstract Entity/ proposit'n	Information
<i>the several aspects (1, Korea, EPE)</i>	no	new/ first mention	object of preposition	no	Abstract Entity/ proposit'n	Information
<i>the culture (7, China, EPE)</i>	yes	new/ first mention	verbal object	yes (undergoes change of state)	General Stative	Information
<i>the baby (5, Korea, EPE)</i>	yes	new/ first mention	verbal object	no	General Stative	Argument

* Please see Section I.C.2.iv.α on page 54 for a complete description of Smith's (2003) *Modes of Discourse* theory.

This perspective is that for an NP to be specific, a writer or speaker must intend to refer to the real-world referent, and not just announce its existence. Thus, unless the writer *knows* the referent—has seen, touched, or otherwise experience the real-world item—it is not specific. The examples given below are therefore specific, or known to

the writer, but not definite, in that the reader cannot pick out the exact referent from a group of possibilities. Therefore, *the*, which implies the reader has knowledge of or can distinguish one individual referent, is in error.

The only possible exception to this is example (a) below. In this case, if there is shared background knowledge that leads the reader to understand that there can be only one social activism group for poor laborers in Korea, then *the* is appropriate and not an over-use because the referent would be definite. However, there is nothing in the section of the essay prior to this example that would indicate anything to the reader about whether one or more social activism groups for poor laborers. It also does not seem reasonable for a writer to expect a US English teacher (the audience for the essay) to locate the referent of this NP based on shared knowledge. Although possible, it is also unlikely that any reader of this essay would have knowledge of one or more social activism groups in Korea. Furthermore, it is unlikely that the real-world entity referred to in (a) is definite because there is only one social activism group in Korea. For this combination of reasons, *the* in example (a) below has been classified as an over-use.

In examples (b) and (c), the referents are also not definite despite the use of a definite determiner. For instance, by the time (b) appears, Student 6 has already stated, “Several countries shows the trend of slowly decrease of population, such as Italy, Japan, and Russia.” So, there are four countries mentioned as having trends of decreasing population.

Unless all four countries’ have the exact same rate of decrease, which is highly unlikely, there is more than one trend. Therefore, the reader cannot definitively locate the referent in memory as *the* implies.

Example (c) above is obviously slightly different because of the demonstrative determiner. This is not a case of over-use of *the* because *the* does not appear in the NP.

Examples 73: Over-use of *the* in Specific, Indefinite NPs

- (a) After that event, I quitted newspaper. I thought that it was not helpful to the other people to report the event indirectly. So I joined the social activity group for the poor labors. (#15, Korea, IIEP)

...so I joined a social activism group for poor laborers.

- (b) I think dur to the policy of one child one family, China successfully control the increase of population. From 2050, China will show the tendency of slowly decrease in population. (#6, China, EPE)

From 2050 on, China will show a trend of slowly decreasing population.

- (c) While the another big country, India, will replace China to be the biggest country in the world. Some other developing countries like Nigeria will also has a significant growth in their population. On the other hand, the population of those developed countries, such as Italy and Japan, will undergo a reduction in the following 50 years. (#10, China, EPE)

..the population of developed countries, such as...

However, the deixis of *those* causes the reader to search the previous discourse for clues to define the referent. As with *the*, the implication is that the reader can already do so based on prior knowledge or previous discourse. In this case, the problem is that the narrowing information that is being pointed to has actually been placed after the NP in question. So, it can be seen that (c) is not exactly the same as (a) and (b), but the issues regarding the appropriate selection of functional morphology are the same—an indicator that the reader should know the exact referent of the NP was used when this was not possible.

Ionin and Ionin, Ko, and Wexler (2003) surmise that language learners will over-use *the* for specific, indefinite referents because of the co-effects of the Fluctuation Hypothesis and the Article Choice Parameter. First, the Article Choice Parameter states that languages either preferentially mark definiteness or specificity. Second, the

Fluctuation Hypothesis proposes that language learners can switch their parameter settings back and forth. These switches cause what appears as optional use of different forms in their writing.⁴ See the table of their predictions on the next page.

Table 93: Predictions of L2 Article Use in English Using the Fluctuation Hypothesis and the Article Choice Parameter (Ionin, Ko, & Wexler 2003)

<i>Context</i>	<i>Definite referent: target the</i>	<i>Indefinite referent: target a</i>
<i>Specific referent</i>	correct use of <i>the</i>	overuse of <i>the</i>
<i>Non-specific referent</i>	overuse of <i>a</i>	correct use of <i>a</i>

Using this table, it can be seen that for specific indefinite referents it is predicted that learners will vary between using *a/an* as the correct target form, and making an error by over-using *the* (refer to the darkly outlined box). The examples of over-use of *the* given above in fact follow this pattern. This does not mean that all of the noun phrase errors in article use in the learner corpus seem to fit Ionin's predictions, though. For instance, the example presented next does not.

The Article Choice Parameter/Fluctuation Hypothesis predicts that non-specific, indefinite referents appear with the correct use of *a/an* as the article. However, here Student 10 over-uses *the*. The intended referent of *the way* is indefinite because the reader does not know the manner in which developing countries might resolve concerns developing due to a rapidly increasing population. Also, if *any way* to resolve these problems exists, there is unlikely to be only one of them, meaning that the referent is not definite by being unique.

⁴ Please refer to I.C.1.b.ii on page 20 and I.C.1.c.ii on page 27 for more information.

Example 74: Over-use of *the* with a Non-Specific, Indefinite Referent

It can be concluded, from the graph, that the developing countries will still have to face the big pressure from drastically growing population. Though in some of these countries, economy are growing rapidly, however, they still need to find the way the solve the problems caused by the expanding need from their people. (#10, China, EPE)

...they still need to find a way to solve the problems caused by the expanding needs of their people.

Furthermore, the referent of *the way* is non-specific because the writer does not know the solution to these problems either. If the writer did know such a solution, it most likely would have been included in the essay to strengthen the author's argument. Therefore, it is clear that at least one of the errors by one of the student writers does not support Ionin's hypothesis.

What is not clear is exactly what this allows one to conclude about Ionin's philosophy. First, the example above could be a production error, and not an error that resulted from an actual misunderstanding of the meaning of articles or the reference of the NP. Second, there are simply not enough NP error examples in the learner corpus to be able to confirm or deny the validity of the Article Choice Parameter/ Fluctuation Hypothesis theory. Given the larger number of examples in which *the* is over-used with specific indefinites, this theory may be plausible, but further research is necessary.

Summary of the Over-use of *the*

One of the most common mistakes the students in this study made in their NPs was the use of *the* when another determiner or no determiner would have been more appropriate. In most cases, these errors occurred in generalizations. The table below shows, though, that in contexts where *a/an* might be preferred by a native writer, there were not that many NPs in which the students used *the* instead.

Table 94: Determiners in *a/an* Contexts

<i>participant</i>	<i>correct use of a/an</i>	<i>incorrect choice of a/an</i>	<i>incorrect use of the</i>	<i>incorrect zero determiner</i>
1, EPE, K (4)	75% (3)	25% (1)	-- --	-- --
2, EPE, K (5)	40% (2)	-- --	-- --	60% (3)
3, EPE, K (2)	100% (2)	-- --	-- --	-- --
4, EPE, K (1)	-- --	-- --	100% (1)	-- --
5, EPE, K (3)	100% (3)	-- --	-- --	-- --
6, EPE, C (5)	-- --	-- --	40% (2)	60% (3)
7, EPE, C (4)	100% (4)	-- --	-- --	-- --
8, EPE, T (3)	100% (3)	-- --	-- --	-- --
9, EPE, C (8)	88% (7)	-- --	-- --	13% (1)
10, EPE, C (6)	50% (3)	-- --	17% (1)	33% (2)
11, K, IIEP (2)	100% (2)	-- --	-- --	-- --
12, K, IIEP (6)	83% (5)	-- --	-- --	17% (1)
13, T, IIEP (10)	70% (7)	10% (1)	-- --	20% (2)
14, K, IIEP (12)	75% (9)	8% (1)	-- --	17% (2)
15, K, IIEP (7)	43% (3)	57% (4)	-- --	-- --
16, T, IIEP (5)	80% (4)	-- --	20% (1)	-- --
17, T, IIEP (0)	-- --	-- --	-- --	-- --
18, T, IIEP (0)	-- --	-- --	-- --	-- --
19, T, IIEP (5)	80% (4)	-- --	-- --	20% (1)
20, K, IIEP (2)	50% (1)	-- --	-- --	50% (1)

Additionally, the next table shows that there are no overwhelmingly glaring patterns of error in what the students used in places in which no determiner might have been preferred by a native writer.

These two tables highlight the fact that individual student's error patterns are some of the interesting findings in this study. Student 5, Student 12, Student 15, and Student 20 all over-used *the* more than the other students.

Table 95: Mistakes in Contexts Appropriate for Zero Determiners

<i>participant</i>	<i>substitute a/an</i>	<i>substitute the</i>	<i>substitute a demonstrative</i>	<i>total # zero D context mistakes</i>
1, EPE, K	-- --	100% (3)	-- --	3
2, EPE, K	100% (1)	-- --	-- --	1
3, EPE, K	-- --	100% (3)	-- --	3
4, EPE, K	-- --	-- --	-- --	0
5, EPE, K	14% (1)	86% (6)	-- --	7
6, EPE, C	-- --	-- --	-- --	0
7, EPE, C	-- --	100% (2)	-- --	2
8, EPE, T	50% (1)	50% (1)	-- --	2
9, EPE, C	-- --	-- --	-- --	0
10, EPE, C	25% (1)	50% (2)	25% (1)	4
11, K, IIEP	-- --	-- --	-- --	0
12, K, IIEP	-- --	100% (8)	-- --	8
13, T, IIEP	-- --	-- --	-- --	0
14, K, IIEP	-- --	-- --	-- --	0
15, K, IIEP	-- --	100% (5)	-- --	5
16, T, IIEP	-- --	100% (1)	-- --	1
17, T, IIEP	-- --	-- --	-- --	0
18, T, IIEP	-- --	-- --	-- --	0
19, T, IIEP	-- --	100% (2)	-- --	2
20, K, IIEP	-- --	100% (4)	-- --	4

The majority of the times when they did so were in general statements, which are situations in which no determiner is often preferred by native writers. The next section looks at the other side of the coin in regards to the students' use of *the*—cases in which it was omitted but a determiner might have been preferred. Again, it is important to remember that different native writers sometimes prefer different determiners under the

same conditions. These are cases in which such changes only result in minor meaning differences, and are different from times when the use or lack of a particular determiner with a full NP creates a structure that is not English.

Omission of *the* and Definite Determiners

In the last chapter, it was discussed that the frequency counts showed that students were not using *the* in any significantly different amount from native writers. The last section then described how the frequency counts did not hint at how often students used *the* in contexts in which native writers might not. This section looks at how the frequency counts were not significantly different partly because this over-use co-occurred with the omission of *the* as well. While the majority of over-use cases were in generic statements, the majority of omissions of definite determiners occur in cases in which there is only one member in the set of what could be the referent. In other words, the referent is unique, and this is known either from the semantics of the NP or from the previous part of the essay. The students with the highest number of errors due to omission of definite determiners are Student 2 (EPE, Korea), Student 3 (EPE, Korea), Student 4 (EPE, Korea), Student 6 (EPE, China), and Student 9 (EPE, China).

Out of these five, only one⁵ had any instances of the over-use of *the*. It cannot be inferred, though, that omitting *the* and over-using it are generally mutually exclusive. Considering the students who had the highest number of over-uses of *the*—Student 5 (EPE, Korea), Student 12 (IIEP, Korea), Student 15 (IIEP, Korea), and Student 20 (IIEP, Korea)—only Student 12 had no errors due to omitting *the*. If there is any trend here, it is merely that students that frequently omit *the* do not tend to over-use *the*. It is interesting, although not a pattern that can be substantiated as being significant, that all of the

⁵ Student 3 had three NPs in which *the* was added when no determiner might have been preferred. Out of these three instances, two in fact referred to the same real-world entity and shared a co-reference chain. In fact, the NPs were the same except for one was plural and one was singular: *the humans* and *the human*.

students who commonly omit *the* are EPE students, while three of the four over-users are IIEP students. Since EPE students are generally at a higher proficiency level than IIEP students, there could be some effect of level regarding who omits and who over-uses. As stated before, there are not enough students in this study to determine if this pattern is significant or only visible because the small sample limits the perception of other possibilities. The next few sub-sections will discuss specific examples from the student essays.

Names

Names are unique and definite because of their inherent qualities: Names denote an individual. Individuals are unique referents. If the referent attached to the name is not known by the reader, the reader at least knows that the writer can only be referring to one object or person. As mentioned in the section on the over-use of *the*, names can appear with or without a determiner depending on what the real-world entity they are attached to is. Most names do not require a determiner. Some cases in which names appear with a definite determiner like *the* are when the names are:

- plural, as in *the United States*,
- describing areas or regions, as with *the Ukraine*, or
- the names of rivers or valleys or other items in which the last word is a singular count noun, as in *the Nile Valley*, *the Mississippi River*, or *the Industrial Revolution*.

In the student essays, *the United States* was the name that appeared without its necessary *the* for the most number of times. The topic of the EPE essay in which students were asked to describe a population graph influenced this. Students had to describe the differences in population trends for over twelve different countries, and *the United States* was the only one listed on a graph that appears with an article. This

difference and the imbalance between the number of countries with and without articles most likely increased the number of omissions of *the*.

Unique or Defined by the Inherent Semantics of the Referent or Shared Background Information

Another type of referent with which many of the EPE students omitted definite determiners are those whose uniqueness/definiteness is inherent to the semantics of the NP. For instance, in the example below, there can be only one *most famous thing*. The superlative adjective means that there is only one possible referent, so such NPs are always definite.

Example 75: Superlatives are Definite

We have many traditional things. First, Most famous thing is the food. (#20, Korea, IIEP)

...*the most famous thing is the food.*

For Ionin, Ko, and Wexler (2003), superlatives could still be non-specific if the writer has not seen or experienced the real-world entity. This does not mean that the NP would not still be definite. In the example above, the writer does in fact know the entity being referred to, so this NP is specific and definite and a correct use of *the* is what Ionin's theory predicts will be used. Note that Ionin's Article Choice Parameter/Fluctuation Hypothesis does not make predictions about when determiners will be omitted by language learners.

In contrast, Gundel, Hedberg, and Zacharski's (1993) Givenness Hierarchy does make predictions about when language learners might omit determiners in noun phrases that are not generic. The Givenness Hierarchy assigns the meaning of "uniquely

identifiable” to *the* in English. This is to distinguish *the* from other definite determiners such as *this* or *that* which are not only uniquely identifiable, but also somehow active or salient in working memory. GHZ’s cross-linguistic comparison of discourse forms predicts that Chinese writers may use *that* or else no determiner for uniquely identifiable referents if they are transferring noun phrase meanings and structures from their L1. Regarding Russian, Japanese, and Chinese, Gundel, Hedberg, and Zacharski state that “a noun with no preceding determiner in these languages can be interpreted as either uniquely identifiable (definite) or merely referential or type identifiable (indefinite) (GHZ 1993:284).” They do not examine Korean so it is unclear if the same predictions might be made for transfer from it to English. More research would be necessary in order to determine whether transfer is taking place in some of these instances, and in order to determine the predictions for Korean. The appearance of both correct usage of article and incorrect usage or omission in the same essay of course confuses the picture more. For Student 20, whose example is above, this is only one of three cases of omission of a definite determiner, and there were also five instances of over-use of *the* in other NPs. So, Student 20’s essay is not filled with one consistent kind of error.

Two other examples of referents that are definite because they are uniquely identifiable are given below. These cases are slightly different than the one described for superlatives. With superlatives, there is truly only one possible referent that could match the meaning of the NP. In the examples below, the shared background knowledge of how the world works makes the referents uniquely identifiable. With both superlatives and the background knowledge cases below, the reader and writer are relying on information not given in the discourse to supply the definiteness of the NP. In both the NPs below, the student writers have incorrectly omitted *the*.

Examples 76: NPs Uniquely Identifiable Due to Shared Background Knowledge and Semantics

- (a) About 100 years ago, every people weared Korean traditional clothes, but in these days there are no people who wear that in a street. People's wearing is almost same to western culture not only on outside. (#2, Korea, EPE)

What people wear is almost the same as in Western culture, not only on the outside.

- (b) It was our first time to come abroad without any relative or friend, and other classmate or friend. Everything was exciting but unknown. In that wild, strange country, we lived together, talked together, and even went toilet together. (#18, Taiwan, IIEP)

...we lived together, talked together, and even went to the bathroom together.

In (a), *outside* is uniquely identifiable because we know that although many places can be described as *outside*, it is still unique because *outside* is not *inside*. For (b), although the reader and the writer do not share knowledge of exactly what bathroom (*toilet*) specifically is being referred to, enough knowledge is shared about what bathrooms are and how people use them to make this referent uniquely identifiable: the closest room of this type in the place where these friends were at any one of a number of particular occasions in the new country.

In contrast to what is seen with the NPs in this section, other referents rely on what has been stated by the writer in the noun phrase or in the prior sections of the essay to make them uniquely identifiable. The next section provides an example of such a case.

Unique or Defined by the Context of the Essay Discourse

A writer can provide enough description in either a noun phrase or in the sentences and paragraphs preceding a noun phrase to make the referent uniquely identifiable, definite, and needing *the* even when it is the first mention of the essay.

Furthermore, native speakers do make mistakes in article use in their writing. The following example demonstrates both of these things.

Example 77: Noun Phrase Description Uniquely Identifies a Referent/ A Native Writer's

Mistake of Omitting *the*

This would make it more difficult for underage students to gain access to alcoholic beverages. It would also cut down on number of underage individuals driving home from bars, potentially reducing drunk driving in this population.” (#37, Native English)

It would also cut down on the number of underage individuals driving home...

Number by itself does not make a uniquely identifiable referent, but there can only be one *number of underage individuals driving home from bars*. In this case, the noun phrase itself has enough description from prepositional phrases and a reduced relative clause to define the real-world entity. The fact that a native writer produced this error shows that errors of omission are one type of mistake natives make. In contrast, there were no native writers who made an error in which they substituted one article or determiner form for another. This might indicate that the mistakes non-native writers make that are substitutions are not just production errors, but rather misunderstandings of the appropriate underlying structure for English or the determiner semantics. Errors of omission could still be partially production errors and partly faulty structures. This would be another area in which further research would need to be done.

Another Kind of Omission Error

Both native and student writers also made errors in which they omitted *a/an*. Again, the native writers did not make either over-use or submission errors. All of their errors were leaving a form out. See the examples that follow.

Examples 78: Omission of *a/an* by Student and Native Writers

- (a) Moreover, he was very strict person. (#12, Korea, IIEP)

He was a very strict person.

- (b) I gave up high salary and choose a real life what I want. (#13, Taiwan, IIEP)

a high salary

- (c) The earth become much smaller than long time ago, because of developed transportation, internet and so on. (#2, Korea, EPE)

a long time ago

- (d) Australia, Saudi Arabia, Sudan and Italy in Group 4 all have much smaller population compared to other groups... (#9, China, EPE)

a much smaller population

- (e) I strongly support continuing to require general education requirements as part of baccalaureate degree program. (#32, Native English)

a baccalaureate degree program

The error by Native 32 in (e) here again could be a production error, meaning something like the writer was writing so quickly that a word was skipped in the writing down, but not in the structure planned by the writer. For the other examples, Student 2 and Student 9 at least have many other errors of omission of determiners in noun phrases. The main interest in these examples is: (1) Ionin's Article Choice Parameter/Fluctuation Hypothesis theory cannot explain omission; and (2) In contrast, Gundel, Hedberg, and Zacharski's (1993) Givenness Hierarchy does make predictions about when language learners might omit determiners in noun phrases that are not generic. In the Givenness Hierarchy, English *a/an* is equivalent to the mental or discourse state of type identifiable or low memory accessibility. In Chinese, though, GHZ states that either *yi* or

nothing are discourse markers for type identifiable referents. If the Chinese writers were transferring structures from their L1, they may be likely to omit *a/an*. (See GHZ 1993:284.)

Summary of Omission of Definite Determiners

The table that follows shows the five students who made the highest number of errors omitting definite determiners. For each student, the NPs with omissions are classified according to whether the uniquely identifiable meaning could be calculated from the meaning of the referent, shared background knowledge, the essay context, or as a name.

This table shows that most of the omissions are connected to referents that require a great deal of calculation of meaning from various parts of the text. The indications of uniqueness of identity for these NPs are not just outside of the semantics of a particular word, but require the linking of information from several sentences or chunks of discourse. If mapping or processing problems exist as one cause of mistakes in second language production, then this might be a situation in which they are occurring. The load on memory and cognitive processing is much higher for such expressions. For example, for names, one has to consider the semantics of the name and the semantics of the determiner or the syntactic rule involved in choosing the determiner for such a name. For the unique referent/ background knowledge cases, the learner has to understand the semantics of the lexical items as with the names and the syntactic rules as with the names.

Table 96: Examples of NPs Missing Definite Determiners

<i>participant</i>	<i>unique because of meaning of referent/ shared background</i>	<i>unique because of essay/ NP context</i>	<i>names</i>
2, Korea, EPE	internet outside most significant factor	parts of wearing and eating western one appearance of western food health*	US West
3, Korea, EPE	--	development of medicine and science very high technological age population (of Japan, Italy, and Russia) population (of India, Indonesia, Nigeria, Mexico, and Saudi Arabia) first two (Australia and US) industry and economy of China	Industrial Revolution USA
4, Korea, EPE	--	expectation of future population third country next order (like the next set of countries) other countries' population[s] total population for the Selected Countries (in 2050)	US USA
6, China, EPE	slowest increase of population growth	12 countries (previously listed) population (of China)	USA US USA
9, China, EPE	--	countries studied (clear from the graph) population in US and Indonesia population in Japan other groups (essay specifically categorizes and refers to four groups)	US US

* *Health* is probably the only NP on this list for which a possessive pronoun may be more appropriate as an indicator of definiteness than *the* would be.

But, the learner also must take into account their own guesses of what information the audience shares with them. This gets even more complicated for the referents determined to be unique by the discourse. The writer must coordinate lexical meanings, syntax, shared knowledge of the world, and previously supplied information in memory. Such a possible processing problem is difficult to define let alone examine, though.

Furthermore, it is not clear why a heavy processing load would have to lead to an omission of a form rather than the selection of a default form in a large number of situations.

The next table summarizes the frequencies of choices of determiners in contexts in which a native writer might prefer *the*. As can be seen in this table, Student 2, Student 3, and Student 4 all have quite a few errors in *the* contexts, and close to even percentages of choosing *the* or nothing in these circumstances. This is a much higher percentage of mistakes than has been seen in other areas. Additionally, while there are only three instances in which *a/an* was used instead of *the*, these three contexts have something in common with the omissions. See the examples below.

Examples 79: Over-use of *a/an* in *the* Contexts

- (a) Furthermore, during the period of time, I would suspect that whether I made a correct decision or not. (#16, Taiwan, IIEP)

I would wonder if I made the correct decision or not.

- (b) I would be a dictator of my country Korea, if I could be anybody for one day, because of the following three main reasons. (#11, Korea, IIEP)

I would be the dictator of my country, Korea...

- (c) About 100 years ago, every people weared Korean traditional clothes, but in these days there are no people who wear that in a street. (#2, Korea, IIEP)

...in those days there were no people who would wear that on the street.

Table 97: Determiners used by Students in *the* Contexts

<i>participant</i>	<i>correct use of the</i>	<i>incorrect use of a/an</i>	<i>incorrect zero article</i>	<i>incorrect possessive</i>
1, EPE, K (5)	100% (5)	-- --	-- --	-- --
2, EPE, K (18)	50% (9)	6% (1)	44% (8)	-- --
3, EPE, K (15)	60% (9)	-- --	40% (6)	-- --
4, EPE, K (16)	56% (9)	-- --	44% (7)	-- --
5, EPE, K (15)	73% (11)	-- --	27% (4)	-- --
6, EPE, C (19)	79% (15)	-- --	21% (4)	-- --
7, EPE, C (9)	89% (8)	-- --	11% (1)	-- --
8, EPE, T (13)	92% (12)	-- --	8% (1)	-- --
9, EPE, C (21)	71% (15)	-- --	29% (6)	-- --
10, EPE, C (17)	88% (15)	-- --	12% (2)	-- --
11, K, IIEP (10)	90% (9)	10% (1)	-- --	-- --
12, K, IIEP (0)	-- --	-- --	-- --	-- --
13, T, IIEP (4)	75% (3)	-- --	-- --	25% (1)
14, K, IIEP (2)	100% (2)	-- --	-- --	-- --
15, K, IIEP (17)	82% (14)	-- --	18% (3)	-- --
16, T, IIEP (8)	88% (7)	13% (1)	-- --	-- --
17, T, IIEP (3)	100% (3)	-- --	-- --	-- --
18, T, IIEP (2)	50% (1)	-- --	50% (1)	-- --
19, T, IIEP (13)	85% (11)	-- --	15% (2)	-- --
20, K, IIEP (10)	70% (7)	-- --	30% (3)	-- --

All three of the referents in the previous examples are ones in which the referent can be defined as uniquely identifiable based on the meaning of the real-world entity and/or shared background knowledge. Considering (a), it seems that most people think that there is one correct or ideal decision to make in any particular situation. Therefore, *correct decision* is uniquely identifiable. Those involved in the issue may not know what this correct decision is, but they believe there is one (and only one?) to be known.

Example (b) is more obvious because the semantics of the word *dictator* implies only one person in power. Finally, example (c) is similar to the example of *bathroom* given earlier. In any situation at any time, *the street* is most likely the one that is closest to the speaker or the location in question.

So, in this section it was discussed how *the* was omitted by some students in particular contexts. With the over-use of *the* common in the writing of other students, the result was that the frequency of the use of *the* by native and student writers was not significantly different. The error analysis presented here shows that individuals develop their own idiosyncratic rules for when to use and when to omit articles and determiners. However, because the students are language learners, these rules are either not always applied, attached to certain lexical items, or affected in some way by the mental taxing of production so that the surface forms do not all appear the same. The next two sections will briefly discuss two other kinds of errors in NP structure and morphology.

Noun Number, Singular and Plural Nouns, and Count or Non-Count Nouns

English is somewhat unusual in that it has both count and non-count nouns. Count nouns are those nouns that semantically are items that can be counted, and therefore are those that appear with plural morphology. Non-count nouns are semantically not things that can be counted, such as liquids, small items not easily counted individually, categories, and abstract ideas. Because they cannot be counted, they do not occur with plural morphology or the article *a/an*. Both the plural morphology and mastering the English perspective on what can be counted and what cannot are challenging for English language learners. Furthermore, because sometimes the same lexical item is used as a count and a non-count noun (with meaning differences), the input students receive is confusing.

In regards to determiners and noun phrases, understanding when a token of a word is being used as count, non-count, singular, or plural is important because some types of nouns and certain determiners cannot co-occur. For example, part of the meaning of *a/an* is “one,” a counting number and singular, so *a/an* cannot be used with non-count or plural nouns. Another difficulty is that every singular count noun in English must follow an article. Given that for many other NPs, especially those with generic referents, no article is used, students can easily get confused.

The landscaped table in this section summarizes the mistakes that were made by the students and native writers in the number or semantic type of referent (count or non-count). As can be seen from the table, by far the most common error was to omit the plural marking in NPs referring to countable, plural real-world entities. Even two native speakers made errors in which the appropriate noun did not have a plural marker at all, or in which the plural marker was attached to the wrong noun (not the head noun). The assumption is that for the native writers these are merely surface errors. Again, these errors by native speakers show that the error rates cannot distinguish between mistakes on the surface that may have resulted from carelessness or speed, and mistakes that appear due to faulty structures or semantics in the target language.

Looking in detail for other patterns, it can be seen that nine of the writers made errors of only one type. Students 10, 11, 13, 15, and 18 all made mistakes only through omitting the plural morphology on count nouns referring to plural referents. In contrast, Students 17 and 20 only made mistakes by adding plural morphology to NPs referring to countable, singular referents. The other nine students made errors of both omitting and adding morphology inappropriately, although they most often represented countable, plural referents without plural suffixes.

The third column of this table is a somewhat strange, catch-all category. This column lists referents which readers cannot definitively state to be singular or plural, count or non-count. This ambiguity results because:

- The nouns or lexical items can be used with a count or non-count meaning. For example, *food* is non-count when it is used as a general category or abstraction describing things that people eat. However, *food* is count in questions such as, “What foods do you like to eat for breakfast?” In this case, *foods* is countable because it means something like *types of food*. The expected answer to this question would be a list of things good to eat for breakfast.
- Generic count nouns can appropriately appear with the plural noun form and no article, or else with a singular noun form and *a/an*. Since the reference is generic, there is no way to determine from the context of the essay whether the writer intended to select a plural form and mistakenly omitted the plural morphology, or if the writer intended to select a singular form and omitted a determiner.
- The lexical item is strange or was made up by the writer. This is the situation with *burgersize* and *okesize*. The meaning is clear, but this is not an actual commonly used word of English. Due to the fact that the context is an apposition, though, a native writer might choose to edit and replace this phrase either as *the size of Cokes*, or *the sizes of Cokes*.

Table 98: Number Mistakes in Noun Phrases in All Essays

<i>participant</i>	<i>singular countable referent</i>	<i>plural countable referent</i>	<i>referent that might be singular or plural, count or non-count</i>	<i>non-countable referent</i>
<i>1, Korea, EPE</i>	--	different culture their hobby their lifestyle our life the relationship of their friends our life Korean our body	good relationship (generic)	foods
<i>2, Korea, EPE</i>	--	one of the important word which express the world many cultures from other country Western clothes are much more comfortable to behave than Korean one. Korean traditional foods are totally different from western one. TGIfriday Subway sandwich Korean's bodies these part (wearing and eating)	burgersize cokesize	--

Table 98—Continued

<i>participant</i>	<i>singular countable referent</i>	<i>plural countable referent</i>	<i>referent that might be singular or plural, count or non-count</i>	<i>non-countable referent</i>
3, Korea, EPE	(all over) the worlds	the human	--	--
5, Korea, EPE	(the USA has) different situations a close positive relations	a good policy such as taxes, health insurance, educations, and so on (people tend to be reluctant to get) the baby (the people are afraid of giving birth to) baby	educations	--
7, China, EPE	--	Cultural Difference the culture difference between countries the culture from other countries	(there are) a lot of food in China today	--
8, Taiwan, EPE	each selected countrys	government event	--	--
10, China, EPE	--	problems caused by the expanding need from their people	--	--
11, Korea, IIEP	--	voters' mind	--	--
12, Korea, IIEP	--	my father's thought every father's thought a good thing and bad thing his thought	--	behaviors

Table 98—Continued

<i>participant</i>	<i>singular countable referent</i>	<i>plural countable referent</i>	<i>referent that might be singular or plural, count or non-count</i>	<i>non-countable referent</i>
<i>13, Taiwan, IIEP</i>	--	no holiday	--	--
<i>14, Korea, IIEP</i>	(I will send them) foods and clothes foods and clothes	property like building, land, or stock	--	--
<i>15, Korea, -IIEP</i>	--	human relationship and research skill (for some days, I had suffered) nightmare in the dream...	--	--
<i>17, Taiwan, IIEP</i>	(some people think) foods (different important possessions)	--	--	--
<i>18, Taiwan, IIEP</i>	--	without any relative or friend (and) other classmate or friend	--	--
<i>19, Taiwan, IIEP</i>	(to the couple it means to wish everything goes) their ways	one of special ceremony a pair of watch some beautiful pictures for wedding memory (the guests give) 'red envelop' (to the couple)	the wedding party (generic)	--

Table 98—Continued

<i>participant</i>	<i>singular countable referent</i>	<i>plural countable referent</i>	<i>referent that might be singular or plural, count or non-count</i>	<i>non-countable referent</i>
20, Korea, IIEP	baby butterflys houses (it makes) many kinds of woods	--	--	--
35, Native English	--	general educations course	--	--
40, Native English	--	the requirement (general educations requirements)	--	--

One final type of error related to the connection of number and reference has to do with pronoun selection. Certain individuals seemed to have trouble selecting pronouns that matched the number of the referent and the other NPs connected in the co-reference chain. See the examples that follow.

Examples 80: Pronoun Number and Referent Number do Not Match

- (a) Our traditional houses are not so big, but it so organizational and beautiful. It makes many kinds of woods. So It colors are fantastic! (#20, Korea, IIEP)

Our traditional houses are not so big, but they have an organized and beautiful design. They [traditional houses] are made of many kinds of wood, so their colors [traditional houses] are fantastic!

- (b) Our traditional houses are not so big, but it so organizational and beautiful. (#20, Korea, IIEP)

... but they [traditional houses] are so...

- (c) When we get known well to other persons, we often follow his speaking style. (#1, Korea, EPE)

When we get to know other people well, we often copy their [other people] speaking styles.

- (d) In fact, without this skill, this scientist (I'm an epidemiologist) would not even be able to express my opinions through this letter. (#25, Native English)

... would not be able to express his/her [this scientist] opinions through this letter.

In (a) and (b) above, Student 20 has a plural antecedent one or two clauses ahead of the uses of *it*. Gundel, Hedberg, and Zacharski (1993) describes the meaning of pronouns as indicating that a referent is 'in focus,' quickly and easily accessible in working memory, and the center of attention of the discourse. The closeness of the antecedent means that Student 20 should not have difficulty in pointing to the antecedent

with these pronouns. Notice, though, that Student 20 uses *it* for all of these plural referents. It could be that this student does not use *they*, and *it* is used for all third person pronouns. (In fact, there are no instances of *they* in Student 20's essay, but this does not sufficiently prove that Student 20 does not use this lexical form.) Student 20 is consistent in this essay, though, meaning that *it* has a special meaning in this individual's interlanguage that is not shared by native speakers of English.

Example (c) is relatively similar to what is seen in Student 20's essay. The antecedent is only one or two clauses separated from the pronoun, the referent is plural and third person, and the pronoun (in this instance a possessive pronoun) is *singular* and third person. The difference is that Student 1 is not consistent in what pronominal forms are chosen. Student 1 does use *their* correctly for a plural referent and close antecedent: "So many people is meeting by on-line chatting and sharing their own information..." This shows once again the kind of optional use of forms that is so commonly seen in language learners' production. Without more information, there is no way to make further statements about the underlying competency or decision-making process of Student 1.

Finally, in (d), a mistake by a native speaker in the selection of the appropriate number pronoun can be seen. Native 25 selects the wrong number pronoun even though the pronoun and its antecedent are in the same predicate. The interference here, though, is the aside in parentheses. *This scientist*, the subject of the main clause, is followed by a sentence in which *I* is the subject. Then, for the pronoun, the first person form is chosen, thereby matching the subject of the aside and not the actual clause. Aside from the interference of having another potential antecedent in subject position, it is one that refers to the same real-world entity as the previous subject. This makes *I* the closest co-referring topic/antecedent for the pronoun to latch onto. Furthermore, it is stylistically odd to refer to oneself in the third person in English. The intervening first person reference then has more added weight from this favoring it as the antecedent of the

pronoun. The relatively higher complexity of what Native 25 is trying to express in contrast with the students makes it less surprising that a native writer of English would make such a reference mistake. However, writing teachers may recognize that native writers often have trouble building well-flowing chains of co-reference through their essays. This is one of the skills that seems to progress as writers gain skill and maturity. Therefore, future research perhaps comparing the pronoun issues in co-reference chains of teenage native writers (pre-college so that they would have had less training in writing) with the pronoun issues in co-reference chains of advanced second language learners. It may be that the skill of referring accurately with pronouns is connected to refining the presentation of ideas in writing across an essay's discourse and that there is something in common between native writers still learning patterns of discourse and non-native writers learning similar skills. On the other hand, there may be absolutely no connection.

The discussion thus far in this section has glossed over or ignored the fact that there are many errors in these students' noun phrases other than determiner choice or noun number. The next section will present a picture of the wide array of other errors that are present.

General Description of Noun Phrase Structure

A fourth really common type of error seen in the learner corpus is simply making mistakes in the overall structure of the descriptive parts of the noun phrases that accompany the head nouns. These structural errors are things such as problematic word order, inaccurate relative clauses, and similar concerns.

In the two examples that follow, Student 1 and Native 26 have somewhat opposite problems. Student 1 is using prepositional phrases to express possession periphrastically. However, for simple possession by a person, native English speakers tend to prefer the Saxon genitive, which is a word order or form of genitive expression not common cross-linguistically. Native 26, on the other hand, has the order of a Saxon genitive without

having the genitive morpheme 's present. Furthermore, since *government* is not a person or animal, the periphrastic construction may be preferable here.

Examples 81: Word Order and Genitive Expressions in Noun Phrase Descriptions

- (a) For example, many children learn their words from their friends. That is, the relationship of their friends influence the words and behavior of them in the respect that friendship is a way to borrow another culture of my friend. (#1, Korea, EPE)

*...influence their words and behavior...my friend's culture??
another culture from my friend??*

- (b) We must understand how the three branches of government cross-balances are to function. (#26, Native English)

We must understand how the checks and balances of the three branches of government?? the three branches of government's cross-balances?? are supposed to?? function.

It is difficult to discern what would be better, though, because there are enough mistakes compounded in parts of these descriptions that what the author intended to communicate is obscured. In the second half of (a), the word order and the choice of the word *another* confound the reader. Similarly, in (b), the word order and the coining of the expression *cross-balances* muddy the meaning conveyed.

In the next example, the relative clause restricting the meaning of *everything* hinders understanding. The relative clause starts with an illegitimate complementizer for English, and then it is missing a copula or verb to link the proposition/*thinking* to what is being thought. As with the previous examples, it is extremely difficult to determine how this relative clause might be better expressed because the communicative intent of the clause is not completely clear.

Example 82: An Ill-Constructed Restrictive Relative Clause

He believed the God and he quit everything what he thought bad things. (#12, Korea)

...everything that he thought was bad?? everything he thought to be bad things??

There are many examples of such descriptions that go awry in language learners' essays. There is relatively little that can be said about such errors, though, because the mistakes and word choices are fairly individual and idiosyncratic. Word choice mistakes or misinterpretations of the meanings of particular lexical items can be particularly damaging to an essay's ability to convey what the author desires.

Conclusions about Error Analysis

Both frequency counts and error analysis have advantages and disadvantages. Frequency data is helpful in looking at the big picture and aggregate details, but obscures the patterns that may be visible by looking at a particular individual's production. Error analysis allows individuals' differences to be viewed, but it is not easy to use such information to make conclusions. This is partly because language learners rarely are consistent in the forms that they choose in particular situations. For the students in this study, the majority of the noun phrases they produced were accurate. Some of the students had a higher concentration of inaccurate noun phrases than others, though. Also, different individuals had patterns in which they either omitted or over-used forms more than others. For example, the table below shows that three of the students (#1, 12, and 19) only over-used *the*, and never omitted it. In contrast, eight of the students only omitted *the* and never over-used it. There is no known shared characteristic, though, that would indicate why any particular student was utilizing one pattern over the other. If a follow-up study were to be done, more personal information would be collected from the

students, more accurate groupings based on level and age of first exposure would be set up, and more data would be collected.

Table 99: Summary of Common Error Patterns in Student Essays

<i>participant</i>	<i>over-use of the</i>	<i>omission of the</i>	<i>noun number</i>	<i>NP structure</i>
1, EPE, Korea	3	--	10	1
2, EPE, Korea	--	8	10	6
3, EPE, Korea	3	7	3	3
4, EPE, Korea	--	7	--	9
5, EPE, Korea	6	2	6	3
6, EPE, China	--	6	--	3
7, EPE, China	2	2	5	--
8, EPE, Taiwan	1	1	2	5
9, EPE, China	--	6	--	1
10, EPE, China	3	2	1	1
11, IIEP, Korea	--	2	2	3
12, IIEP, Korea	7*	--	6	5
13, IIEP, Taiwan	--	--	1	2
14, IIEP, Korea	--	--	3	2
15, IIEP, Korea	9	3	5	3
16, IIEP, Taiwan	2	1	--	6
17, IIEP, Taiwan	--	1	1	4
18, IIEP, Taiwan	--	1	4	--
19, IIEP, Taiwan	2	--	6	4
20, IIEP, Korea	5	3	7	1

*All seven of these over-uses are with the same lexical item and referent: “the God”.

Final Conclusions

This study attempted to not only examine discourse and determiners, but also to combine some of the methods of computational linguistics and theoretical linguistics. Computational linguistics and corpus linguistics is heavily based on frequency information and gathering enough language data that commonalities and patterns emerge. One difficulty in corpus linguistics is that a small amount of the language is used a great deal, and a large percentage of what is possible is rare, even in a single essay. This was seen by the fact that on average 70% of the word types in the native essays appeared only one time. Overall, computational linguists are focused on what people typically do, and create descriptions of language use that can usually be exploited for practical use (search engines, voice recognition software, and more). Linguistic outliers and strange constructions are problems. For theoretical linguists, linguistic outliers and strange constructions are often also problems, but they in turn help determine the limits of what is possible in human language. Theoretical linguists are not typically interested in frequency data because extremely specific examples are often needed to validate a hypothesis or possible structure.

Those who study language acquisition are often stuck in the middle between looking at what learners actually do, and trying to figure out the limits of what is possible for them, or what they know, but do not know that they know. In the end, this study may have ended up stuck in the middle—not enough data or patterns for one side, and not enough theoretical explanation for the other. Even so, several possible further areas of inquiry were found, such as a more detailed examination of co-reference chain forms, or else the investigation of a corpus of writing on more controlled topics.

APPENDIX A: TEST QUESTIONS AND
INSTRUCTIONS GIVEN TO NATIVE ENGLISH-
SPEAKING WRITERS FOR CREATION OF THE
NATIVE CORPUS

Text of Test Instructions and Questions

Directions:

There are two essay questions below. Please select ONLY ONE question to write about. Plan your essay and then begin typing or writing, but do not work longer than 30 minutes from the time you start to read the questions. Please take the test all in one sitting, rather than starting and stopping. If you do not have a 30-minute block of time, then I cannot use your essay, so you should not take the test.

It is fine if you finish the test early. You may choose to quit this study at any time. Simply stop writing and discard your essay. If you want to be a part of the study, when you are done, please mail or email me your writing:

[ADDRESS]

PLEASE BEGIN TIMING YOURSELF NOW.

Question 1:

Imagine that the local government is considering making it illegal for people to enter bars unless they are over 21 years of age. They hope that this will cut down on the amount of underage drinking among local college students. Some community members think that this would hurt local businesses and negatively affect their neighborhoods by increasing the number of disruptive house parties where alcohol is served. What do you think? Give reasons for your answer.

Question 2:

The public universities in your state are considering eliminating general education requirements. General education requirements demand that students take classes in a variety of subjects (math, English, foreign language, science, and history), no matter what their interests or intended majors are. If general education requirements are removed, students could avoid taking math, science, or composition if they chose. Write a letter to the editor either supporting general education requirements or supporting the elimination of these requirements. You may use your personal experience for examples and as part of the basis of your argument.

THANK YOU!

APPENDIX B: PENNTREEBANK PART-OF-SPEECH

TAGS

The linguistic analysis program MontyLingua that was used to automatically identify the parts-of-speech of all the words in the corpora uses the PennTreebank tagset (Santorini 1990) that is listed below.

Table B1: PennTreebank Tags for Part-of-speech Tagging

<i>Part of speech</i>	<i>Tag</i>
<i>Adjectives</i>	
Adjective or ordinal number	JJ
Comparative adjective	JJR
Superlative adjective	JJS
<i>Adverbs</i>	
Adverb or negation	RB
Comparative adverb	RBR
Superlative adverb	RBS
<i>Nouns and determiners</i>	
Pre-determiner**	PDT
Article or determiner*	DT
Plural common noun	NNS
Singular or mass common noun	NN
Personal pronoun	PRP
Possessive ending	POS
Possessive pronoun	PRP\$
Plural proper noun	NNPS
Singular proper noun	NNP
<i>Verbs</i>	
Gerund or present participle	VBG
Modal verb	MD

Table B1—Continued

<i>Part of speech</i>	<i>Tag</i>
Past participle	VBN
Particle	RP
Past tense verb	VBD
Present tense verb, other than 3 rd person singular	VBP
Present tense verb, 3 rd person singular	VBZ
Base form of a verb	VB
<i>Other</i>	
Cardinal number	CD
Coordinating conjunction	CC
Exclamation or intejection	UH
Existential there	EX
Foreign word	FW
List item marker	LS
Preposition or subordinating conjunction	IN
Possessive WH-pronoun	WP\$
Symbol	SYM
to, preposition or infinitival marker	TO
WH-determiner	WDT
WH pronoun	WP
WH adverb	WRB

**Determiner—DT:* “This category includes the articles a(n), every, no, and the, the indefinite determiners another, any and some, each, either (as in either way), neither (as in neither decision), that, these, this, and those, and instances of all or both when they do not precede a determiner or possessive pronoun (as in all roads or both times). (Instances of all or both that do precede a determiner or possessive pronoun are tagged as predeterminers (PDT).) Since any noun phrase can contain at most one determiner, the fact that such can occur together with a determiner (as in the only such case) means that it should be tagged as an adjective (JJ), unless it precedes a determiner, as in such a good time, in which case it is a predeterminer (PDT).” (Santorini 1990:4)

***Predeterminer—PDT:* “This category includes the following determiner-like elements when they precede an article or possessive pronoun. Examples: all/PDT his marbles, nary/PDT a soul, both/PDT the girls, quite/PDT a mess, half/PDT his time, rather/PDT a nuisance, many/PDT a moon, such/PDT a good time” (Santorini 1990:6)

APPENDIX C: PYTHON PROGRAM USED TO
INITIATE PART-OF-SPEECH TAGGING BY
MONTYLINGUA-2.1

```
#!/usr/bin/env python

# allow for use of MontyLingua, importing and creating an
# instance of the class
import sys
import MontyLingua
ml = MontyLingua.MontyLingua()

# retrieve file names input on the command line
file1 = sys.argv[1]
file2 = sys.argv[2]

# open corpus.txt
f1 = open(file1, 'r')

# read the file
lines = f1.readlines()

# close the file
f1.close()

# open a new file into which to write the tagged lines
f2 = open(file2, 'w')

# call the MontyTagger to tokenize and tag the text
for line in lines:
    tokenized = ml.tokenize(line)
    tagged = ml.tag_tokenized(tokenized)
    chunks = ml.chunk_tagged(tagged)
    # print or write to the new file
    f2.write(chunks)

# close the second file
f2.close()
```

APPENDIX D: PERL PROGRAM FOR
 AUTOMATICALLY REMOVING NON-NOUN OR
 DETERMINER PHRASE RELATED PART-OF-SPEECH
 TAGS

```
# Use perl to read this file
#!/user/bin/perl -w

# ensure the proper number of files input to the program on the command line
if ($#ARGV != 1) {
    die("Usage: \tperl pat_match2.pl sourcefile outputfile\n");
}

# open the files to be manipulated
open(F, $ARGV[0]) or die("No source file...\n");
open(F2, ">$ARGV[1]) or die("Couldn't open new file...\n");

# This section deletes certain part-of-speech tags

# Read each line of the input file
while ($line = <F>) {

    # Find and delete non-noun related tags attached to words
    $line =~ s/([A-Z]?[a-z]+|[VIRCSMWET])[A-Z][A-Z]?/$1/g;

    # Find and delete adverb or adjective and verb phrase tags
    $line =~ s/\([AV]X//g;
    $line =~ s/[AV]X\)//g;

    # Find and delete symbol tags
    $line =~ s/[<\/]SYM/</g;
    $line =~ s/>\/SYM/>\n/g;

    # Print the altered lines to the screen and to the output file
    print("$line\n");
    print(F2 "$line\n");
}

# Close the files
close(F);
close(F2);
```

APPENDIX E: CO-REFERENCE CODING RULES

Based on the coding rules of L. Hirschman (1997), “MUC-7 Co-Reference Task Definition. Version 3.0”

Co-reference Coding Rules and Syntactic Constraints on Markables

***many examples and text appearing in quotes comes from Hirschman, L. (1997). “MUC-7 Co-reference Task Definition: Version 3.0.” Accessed April 13, 2006.

http://www.itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/co_task.html

Part 1: How to mark and determine co-referring noun phrases (NX...NX)

How to define the extent of NPs, considering complements, NPs embedded in other NPs, and syntactic position:

- Individual nouns embedded in larger noun phrases, whole larger noun phrases with complements, and pronouns (personal, possessive, and demonstrative) should all be tagged as (NX...NX) units. This means that I may mark a larger noun phrase that has contained within it other noun phrases.
- “the relationship of their friends” 001: (NX₁ the relationship of (NX₂ (NX₃ their NX₃) friends NX₂) NX₁)

Pronouns may be embedded in a relative clause, and yet should still be tagged as NX units. Hirschman (1997) gives the example “every man who knows his own mind.” In this case, ‘his’ is tagged as an NP that co-refers with the whole NP that consists of “every man who knows his own mind.”:

- See also student 004: “the countries that increase their population”: (NX₁ the countries that increase (NX₂ (NX₃ their NX₃)/ID001 population NX₂)/ID002 NX₁//ID001: The subscripts have been added to clarify the noun phrase

embeddings. The ID numbers show the entities that co-refer. NX_1 and NX_3 co-refer, and this is shown by the fact that they have the same ID. They share identity in that they both refer to the same real world object or idea.

Noun phrases that cannot be split into parts should be marked as one (NX...NX) unit. In other words, this means that **noun phrase heads** cannot be marked as their own NP. Only the full phrase can be marked. In addition, **names** cannot be split into parts. In some cases, decisions about what can be split and what cannot will be difficult to make. In these situations, the annotator's personal judgment will have to be relied on, and a list should be made of the problematic cases and tagging to increase the uniformity of the tagging. Being unable to split a larger NP into many parts may mean that certain units cannot be marked as co-referential because they will not end up as exactly the same set. However, these will usually be within the same, larger noun phrase. Co-reference to other noun phrases will still usually be able to be marked. If this is a problem with the discourse theories, this conclusion may have to be revisited later, which is true of all the coding rules.

- Hirschman (1997): “Equitable of Iowa Cos. ... located in Iowa.” (the 2 Iowas are not marked as co-referential because the first Iowa is not marked as its own NP because of how the first is embedded in the name)
- Hirschman (1997): “<COREF ID="0" MIN="building">the large strange-looking building, which is <COREF ID="1" TYPE="IDENT">Widener Library</COREF></COREF>” (building and Widener library are not marked as co-referential because Widener library is embedded in the larger NP and building is not marked as its own NP)
- Hirschman (1997): “okay then I'll take <COREF ID="0" MIN="E two">engine E two</COREF> ... so uh the plan is to take <COREF ID="1" TYPE="IDENT" REF="0" MIN="E two">engine E two</COREF> ...” (E two and engine are not

marked separately as two NPs because it is impossible to determine that they are truly being treated as two separate entities and not just one compound NP)

- Few proper names other than country names are used in the student essays, so there are no examples exactly like the ones above. This is a similar case, though: Student 007 said: “For example, there are a lot of food in China today. The Chinese people have more choices of food than before.” In “the Chinese people,” “Chinese” works like an adjective or part of a name and cannot be tagged as a separate part of the NP. Therefore, there can be no co-reference tagged between “Chinese” and “China” from the sentence before.

NPs that are **assertions** and **predicate nominals** should be tagged as (NX...NX) units. These can additionally be tagged as co-referential with subjects when it seems they refer to the same real world entity. This can happen even when the predicate is indefinite.

- “I have a machete” Hirschman (1997)
- “John is a policeman”
- Student 013: “Yes, I was confused and afraid when I decided to leave my previous job which was a sales manager position.”
- example: Ellen is the president of the Spanish Club.
- Hirschman (1997): examples: “*Mediation* is *a viable alternative to bankruptcy*.”
- Hirschman (1997): “*Farm-debt mediation* is *one of the Farm Belt's success stories*.”
- Hirschman (1997): “*ARPA program managers* are *nice people*.”

NPs that are **predicates of negative verbs** should be tagged as (NX...NX) units, and these units co-refer with the subject noun phrase.

- Hirschman (1997): “I don’t have (NX a machete NX)”
- Student 006: “They don’t need (NX more children NX).”
- Student 002: “(NX People NX) eat (NX rise NX), not (NX bread NX), and there were (NX no chocolate, pizza, soda, and coffee NX).”
- Student 011: “I would order my cabinet to make a strict law for people not to use (NX any kind of violence NX) in their daily life under any circumstances.”
- Student 002: “About 100 years ago, every people weared Korean traditional clothes, but in these days there are no (NX people who wear that in a street NX).”
- Student 002: “ (NX Korea, my country, NX) is not (NX an exception NX).”

NPs that are in **questions** should be tagged as (NX...NX) units.

- “Do you have a machete?” Hirschman (1997)
- Student 012: “Is he knew about his death?” (the only question in the student writing)

The full NP, with all complements, should be an (NX...NX), for example: DET + ADJ + ADJ + N + Relative clause.

- “the relationships of their friends” student 001
- “many cultures from other country” student 002
- Note that relative clauses may be malformed by non-native speakers: See student 012: “and bad things what everyone does”: (NX₁ bad things what (NX₂ everyone NX₂) does NX₁)
- Here is another example of a poorly formed relative clause: student 019: “the things of wedding related” This whole unit should still be tagged as an (NX...NX) unit.

Gerunds that are noun-like should be marked as NPs. These gerunds are possibly modified by other nouns or adjectives, preceded by an article, or else followed by a prepositional complement.

- Hirschman (1997) ex: program trading, excessive spending, the slowing of the economy
- “on-line chatting and sharing” student 001
- “the culture borrowing” student 001
- “dying of starving” student 014 Both gerunds in this phrase should be counted as nouns.

Possessive forms of pronouns that appear before head nouns in phrases get their own (NX...NX) marking, and they can be marked for co-reference separately from the head.

- Hirschman (1997) ex: “its chairperson”—tag both “it” and “its chairperson” as noun phrases
- “every father’s thought” student 012: (NX₁ (NX₂ every father’s NX₂) thought NX₁)

Modifiers that are nouns that appear before other nouns should be tagged as a separate (NX...NX) phrase when these nouns co-refer with a name or the head of some other noun phrase. This would imply that they are not functioning as a compound noun, but rather as a separate entity. The decision about whether a modifying noun will be referred to separately later may be difficult to make in certain situations. When in doubt, modifying nouns should be tagged as embedded noun phrases. In most cases, though, the modifying nouns will not be used on their own later.

- Hirschman (1997): “The price of *aluminum* siding has steadily increased, as the market for *aluminum* reacts to the strike in Chile.” Aluminum in the phrase “aluminum siding” would have its own NP marked here.
- Student 015: “In conclusion, the challenging spirit and real activity are pivotal to the life. From the event I got the active character and brave. My life decree is that ‘I just do it, not watch it’.” In “life decree,” the noun “life” modifies the noun “decree.” This “life,” though, is co-referential with the generic use of “the life” earlier on.

For **conjoined NPs**, the entire conjunction is tagged as a unit as one (NX...and...NX). The NPs in the conjoined construction are also marked as NPs. (This is different from MUC-7 encoding. In MUC-7 tagging, the whole conjunction or the parts are only marked based on whether they are referred to as a whole or as separate parts elsewhere in the text. The only units tagged are those referred to later—the whole if the whole is referred to, the parts if the parts are referenced separately. Due to concerns of taxing on the memory of the coder, this practice was altered.) In cases in which there is a determiner or modifier in front of the first conjunct, it will be assumed that this modifier can be read before any bare conjunct in the list, unless this seems to be incorrect due to the context. In the example below, then, it will be assumed that the student intended to also say “their foods.”

- Student 001: “their hobby, foods, and their life style”: (NX₁ (NX₂ their hobby, NX₂) (NX₃ [their] foods, NX₃) and (NX₄ their life style NX₄) NX₁) [Note: the possessive pronouns would also have to be tagged as separate noun phrases]

“**Headless**” **noun phrases**, or those in which there is a null noun are still marked as noun phrases. The “head” of the phrase is taken as the last part of the whole phrase before any prepositional phrases, relative clauses, or other complements.

- Hirschman (1997): “the six youngest” (youngest=head)
- Hirschman (1997): “the five who are the best” (five=head)
- Student 015: “the villige for the poor” (poor=head)

For **compound nouns and collocations**, the head of the NP is considered the syntactic head of the noun phrase, and not the whole compound.

- Hirschman (1997): ex: “income taxes”—head = taxes
- Student 015: “the college newspaper”—head=newspaper

Nouns in prepositional complements are tagged as NPs, and these are in turn tagged as being co-referential with the entire NP, if appropriate.

- “<COREF ID="1" MIN="job">The job of <COREF ID="2" REF="1">manager</COREF></ COREF>”—manager and the job of manager are marked as co-referring (even though this is a little weird)
- Student 002: “the number of fat people”: (NX₁ the number of (NX₂ fat people NX₂) NX₁): no co-reference is marked
- Student 009: “both of them”: (NX₁ both of (NX₂ them NX₂)/ID001 NX₁/ID001: co-reference is tagged between the NP embedded in the PP and the entire NP
- Student 009: “those in Group (3)”: (NX₁ those in (NX₂ Group (3) NX₂)/ID001 NX₁/ID001: co-reference is tagged between the NP embedded in the PP and the entire NP

Appositional phrases are tagged as separate NPs that are co-referential with the entire NP. However, there are exceptions, so please see part 2.

- “Korea, my country,” student 002: (NX₁ Korea, (NX₂ my country NX₂)/ID001 NX₁/ID001

For **functions** in chains of reference that may be collapsed, primary consideration is given to names or individuals over kinds or types.

- For example, if there is a particular individual who has held two jobs at two points in time, all three can be linked in the same co-reference chain (even though the person probably did not hold both jobs at once). On the other hand, if there is a job that was held by two individuals at two different points of time, these cannot be marked as co-referring. They would all get NP identifications, but they cannot be linked because the people cannot be conflated like the jobs can.
- Hirschman (1997): ‘Henry Higgins, who was formerly sales director for Sudsy Soaps, became president of Dreamy Detergents’ should be annotated as:
`<COREF ID="1" MIN="Henry Higgins">Henry Higgins, who was formerly
<COREF ID="2" MIN="director" REF="1" TYPE="IDENT">sales director for
Sudsy Soaps,</COREF></COREF> became <COREF ID="3" MIN="president"
REF="1" TYPE="IDENT">president of Dreamy Detergents</COREF>’`
- If one is talking about the president, and means for president to go with Bill Clinton at one point, but to go with George Bush at another point, then the two instances of “president” should be linked to their respective people, but not to each other.
- Functions can be difficult to mark for co-reference, but they are straightforward to tag as NPs because they do not have embedded elements usually.
- Hirschman (1997): For example, here everything co-refers: “GM announced *its third quarter profit*. *It* was *\$0.02*.”
- For a function, mark the co-reference with the item in the same clause, or else the most recent item to fill the function. See the other comments in Part 2.
- Here is a student example (from student 006) to consider: “In 2000, China has the most population and in 2025 China still has the largest population among these

countries. But in 2050 the population of India will become the most one and it surpasses the population of China.”

- The function here is the country with the largest population. This is filled by three different items: (a) the population of China in 2000, (b) the population of China in 2025, and (c) the population of India in 2050. These are not all co-referential, and we do not want to collapse the chains of reference.
- In this case, the names of the countries get precedence. The three instances of noun phrases referring to the largest population would not be tagged as co-referring to each other. Each instance refers to a different time or country, so they do not co-refer.

In regards to **metonymy**, the NPs should be marked and the head should be determined according to the rules listed above. If two NPs, one of which involves metonymy, seem to refer to the same real world entity, then they should be tagged as co-referring. Hirschman (1997) recommends making a list of problematic items so that some kind of consistency can be achieved.

Numbers and number phrases will not usually be tagged as (NX...NX) units, except when the number is acting as the head of a noun phrase, or similar to a modifier of another noun. In these instances, the entire NP is tagged, and then co-reference may be tagged between the entire NP and parts of any complements. (This is different than the coding rules outlined by Hirschman 1997.)

- Student 002: “one of the important word which expresses the world”: This whole phrase is a noun phrase unit.
- Student 002: “100 years ago”: This would be tagged as a noun phrase.
- Student 006: “China will increase population from 1,200,500 in 2000 to 1,400,100 million population in 2025 and then decrease the population to the

1,400,000 million population in 2050.”: Only the underlined expressions will be tagged as noun phrases.

Quantifiers before other nouns, such as “all,” “every,” and “some”:

- “such a good thing” Student 012
- “all the wealth in my society and my nation” Student 011
- “some countries’ population” Student 004
- “every culture” Student 007
- “every father’s thought” Student 012
- “Now my friends all support and encourage me.” 013

Typos and misspellings that are clearly meant to be words that are part of an NP should be tagged as an (NX...NX) unit. These decisions will have to be made on an individual basis. The reason is that the decision has to be made whether the error obscures the intended meaning and therefore the placement of the word in question. If it is not clear where the item in question is in regards to the syntax, then it should not be tagged. If there is a reasonable amount of certainty about what was meant, then it should be tagged.

- Student 015: “there events made what I am”: (NX there events NX) made what (NX I NX) am: It seems that “there” is actually a typo or misspelling for a demonstrative pronoun. Thus, it has been tagged as part of an NP.

Referential “it,” even when it only refers back to a sentence, clause, or idea, should be tagged as a noun phrase. The only issue is that the co-reference with the phrase or clause will not be indicated in the tagging.

- See student 018: “Two years ago, I traveled with my best friend in America, and at that period of time, I realized something we call ‘afraid.’ It was our first time

to come abroad without any relative or friend, and other classmate or friend.” In this case, “it” refers to the trip abroad. Therefore, it is referential and should be tagged as a noun phrase, but no co-reference will be tagged.

Part 2: What should NOT be marked as a Separate Noun
Phrase (NX...NX) or as Co-referring

Wh-noun phrases should not be tagged as (NX...NX) units, even if they participate in co-reference with another NP.

- Hirschman (1997): “Which engine would you like to use?” “Who is your boss?”—neither “which engine” nor “who” is markable
- Student 015: “What made me hard is not that I was scared but that I didn’t try to do anything to help him.”: “What” in the relative clause in subject position is not tagged as a noun phrase.

Clauses should not be tagged as (NX...NX) units, even if they participate in co-reference with another NP. This relationship will simply not be recoverable by the tagging scheme. See the section above on referential “it.”

- Student 003: “In addition China will be a developed country after 2025. It is easy to imagine because industry and economy of China is getting much bigger.”: “It” in the second sentence would be tagged as a noun phrase, but it would not be tagged as co-referring with anything, because it co-refers with the previous clause.

Names and compounds should not be split into parts. Separate parts of such noun phrases cannot truly be split up. The phrases from Hirschman given below are all examples of appositions, or else they could be called left modifying nouns. The other

examples mainly contain reduced relative clauses or compound nouns. In any case, it is not clear that they could be split into parts. There is also the problem that it would be difficult to define what the head noun was. See also the discussion under Part 1.

- Hirschman (1997): “the real estate company Century 21”
- Hirschman (1997): “the realtor Century 21”
- Hirschman (1997): “presidential advisor Joe Smarty”
- Hirschman (1997): “Treasury Secretary Bucks”
- Student 011: “all I want”: This phrase would probably also be considered to be “headless.”
- Student 009: “all the countries selected” and “countries studied”
- Student 013: “12 hours a day”
- Student 013: “my things to do list”: Note that “my” would also be tagged as a noun phrase of its own.

NP heads cannot be separately marked—they have to be marked as part of the whole NP. The head will be listed in the annotation, but it is not its own (NX...NX).

- Hirschman (1997): “Linguists are a strange bunch. Some linguists even like spinach.”-linguists and some linguists are not marked as co-referring because the head of “some linguists” is not marked on its own, and the sets are not equal.
- Hirschman (1997): “<COREF ID="0" MIN="rate">The rate, which was <COREF ID="1" REF="0">6 percent</COREF>,</COREF> was higher than that offered by the other bank.’ In this example, pronoun *that* is coreferential at the FUNCTION level with *The rate*. However, *that* occurs as the head of a noun phrase, *that offered by the other bank*, which is NOT coreferential with *The rate* and *6 percent* (indeed, it refers to a higher rate), so *that* is an instance of a pronoun that cannot be marked in our current framework, even though we lose some type coreference information by not marking it.”

- Student 015: “one of them”: (NX₁ one of (NX₂ them NX₂) NX₁): “one” cannot be tagged as a noun phrase by itself

Gerunds that are verb-like should not be marked as NPs. These gerunds often are followed by an object (instead of a prepositional complement), and can be modified by an adverb.

- Hirschman (1997): “slowing the economy”
- Student 001: “borrowing the food of another culture”

Do not tag **gaps, empty pronouns, PRO, or any other null elements** as noun phrases. This includes wh-relative words in relative clauses.

Existential “it” or “there” should not be tagged as noun phrases. They are noun phrases, but they do not refer to any real world entity and are just there for grammatical convenience. Since I am interested in reference, I am not tagging these at all. (This is different than the co-reference tagging standards of MUC-7.)

- Student 016: “It is true for many people there are many things to make them confused or afraid.”: “It” is not tagged as a noun phrase in this sentence since it does not refer.

Nouns acting as modifiers before other nouns would not be tagged as an embedded NP when they always appear as pre-nominal modifiers. This implies that they are functioning as part of a compound noun, and are contributing to one reference rather than a reference on their own.

- Hirschman (1997): “Ocean Drilling & Exploration Co. will sell its *contract drilling* business. ... Ocean Drilling said it will offer 15% to 20% of the *contract drilling* business through an initial public offering in the near future.” Contract

would not get tagged as its own NP here because it is obviously participating in the compound noun “contract drilling.”

- Student 002: “Subway sandwich”: “Subway” is never used other than as a modifier of “sandwich,” and therefore is not tagged as a separate noun phrase.
- Student 002: “burgersize” and “cokesize”: I am also including these examples in this category. The spacing used by the student demonstrates that he or she does not consider the modifying noun to be a separate entity.

Idioms should not be marked as noun phrases (even when they are). This is because they do not truly refer. Furthermore, they are set phrases, so the error type if a student used an idiom incorrectly would be different than the error using a referring expression.

Appositional phrases that are negative can be tagged as separate noun phrases, but should be considered to be co-referring.

- Hirschman (1997): “Ms. Ima Head, never a great MUC fan,”—in this case, the apposition is tagged as an embedded NP
- Hirschman (1997): “The criminals, often legal immigrants, ...”—this also has an embedded NP, and co-reference is marked.

Predicate nominals should still be tagged as noun phrases. However, they are not tagged as co-referring with subjects or other NPs when there is added probability expressed in the verb phrase, when the predicate is negative, or if the NPs do not cover exactly the same set of real world entities.

- Hirschman (1997): “Phinneas Flounder may be the dumbest man who ever lived.”
- Hirschman (1997): “Phinneas Flounder was almost the first president of the corporation.”

- Hirschman (1997): “If elected, Phinneas Flounder would be the first Californian in the Oval Office.”
- Student 002: “Korea, my country, is not an exception.”: (NX₁ Korea, (NX₂ my country NX₂)/ID001 NX₁/ID001 is not (NX₃ an exception NX₃)/ID002
- Student 002: “To Korean people, western food is not strange thing anymore.”: To (NX Korean people NX)/ID001, (NX western food NX)/ID002 is not (NX strange thing NX)/ID003 anymore.

NPs that refer to types and sets can be tagged as NPs. However, types and sets of particular real world items cannot co-refer because they are not the same. Basically, specific and generic NPs do not co-refer. When in doubt about whether NPs are types or kinds, the form of the NP (for example, bare plural), may indicate type or kind (although this will not work with all non-native speakers’ production).

In **functions**, again all the NPs are tagged as (NX...NX). However, not everything is stated to be co-referring. In addition, for functions, the item not in the same clause or that is not the most recent does not usually get marked as co-referring to previous elements. See the examples in the section on functions above.

- Hirschman (1997): For example, here the NPs between stars do not all co-refer: “*The temperature* is *90*....*The temperature* is rising.” The first “temperature” and “90” co-refer, but the second “temperature” does not because it refers to the function and not the temperature of a specific time and place.

Set phrases, transitions, and other collocations in which the noun phrases do not actually refer to real world entities should not be tagged as (NX...NX) units. These items not only do not participate in reference in the way other noun phrases do, but they are often learned as memorized units. (This is a difference similar to an irregular verb versus

the use of a productive suffix.) The errors in use will these set phrases may not be related to those in other types of noun phrases. (Note: Many of these were tagged as NPs or nouns by the MontyTagger, whether they are usually considered as such or not, and that is why they are listed here. These items need to be corrected in the text of the revised corpora.) Spelling variants of the forms of these words should also be left untagged. For example, “any more” and “anymore” may be used as interchangeable by students. Either form should not be tagged if it fits this category.

- In the respect that, in the case student 001
- for example student 002
- “First of all” student 011
- all over, in conclusion, in addition student 003
- a lot
- the fact that
- anymore
- sometime
- all over
- in addition
- in conclusion
- now
- and so on
- on the other hand
- another example
- to the contrast
- on the contrast
- in order to
- each other (see student 012 and student 018)
- of course (student 013)

- anything (student 015)
- something (student 015)
- as a result
- in favor of (student 017)
- on the one hand (student 017)
- anything (student 017)
- last but not least (student 017)
- student 005: “giving birth”: I am not tagging “birth” as a separate noun phrase.
- student 016: “I am a little confused.” This is a degree expression that has a determiner before an adjective. This will be considered an example of a set expression, and therefore will not be tagged or considered part of the data.

Dates, currency expressions, and percentages will not typically be marked as (NX...NX) units, unless there are extenuating circumstances. These items often participate in co-reference, but they do not often have determiners. If they appear in longer phrases, such as “the year 2025,” they will be tagged, but if they are solely numbers, such as “2025,” they will not be tagged. If they are necessary to determine why a particular subsequent noun phrase is structured in a certain manner, then they may be tagged, but this would be a special circumstance. (This differs from the MUC-7 tagging system.)

- Student 003: “In addition China will be a developed country after 2025.”: “2025” is not tagged.

Items whose part of speech cannot be determined, confusing typos, and unintelligible misspellings should not be tagged as NPs because it is unclear what they are. These decisions must be made on an individual basis. Sometimes it seems clear

from the context and position what the writer intended, and sometimes the problems obscure the meaning too much to be able to count the NPs.

- student 011: “to live happily and that to enjoy” What is “that”? Is it a demonstrative or a clause subordination marker?
- student 015: “I got the active character and brave.” The question here is whether “brave” is meant to be an adjective or is merely a misspelling for a noun. Since it is unclear what the student actually intended, “brave” is not tagged in this case. There is not enough information to tell whether this should be a noun. This is because ESL writers often have difficulty with parallelism.
- student 018: “I realized something we call ‘afraid.’” It is unclear if the student meant to use an adjective or a noun. The student may have mistaken “afraid” to be a noun, but on the other hand, the student may not have known that a native speaker would use a noun in this position. Therefore, this cannot be assumed to be a noun.
- student 013: “We always say talking rest is for to last longer.” The problem here is the phrase “talking rest.” It is not clear what the student means by “talking.” “Talking” could describe “rest” as an adjective. On the other hand, it could be a misspelling of taking and therefore a verb. “Talking has therefore not been included in the noun phrase, while “rest” is still tagged as a noun. It is likely that in either case “talking” would be a verb-like gerund that would not be counted.
- Student 016: “In the beginning, I would try to find out information about how to solve these problems. I become confused and afraid. However, I understood this is a normal respond. Furthermore, during the period of time, I would suspect that whether I made a correct decision or not.” The problem in this excerpt is that it is unclear what part of speech the underlined “that” is. It is clear that the student misuses certain words—such as respond instead of response. Because it is

unclear whether “that” is a demonstrative determiner or a subordinating conjunction, this item has not been tagged.

- student 017: “Troubles will make me very sad, and I won’t to do any more.” It is possible that “any more” here is a headless noun phrase (such as “any more work”), but this is unclear, so it has been left untagged.
- student 019: “Third, they have to inform their friends and relationships by phone or mail that make sure everyone be invited.” In this case, there is a weird use of “that.” It could be interpreted as either a demonstrative or a subordinator. Because it is unclear, this has not been tagged.
- student 011: “to become new people mentally” It is unclear whether “mentally” is placed within the noun phrase, or whether it is higher up in the VP. Therefore, “mentally” has not been included in the tagged NP.

Time adverbs should not be tagged as part of the noun phrase unless they clearly are referring to the NP, because otherwise they may really describe the timing of the verb in relation to the noun. See the example: “I really love my life now.” (013) Now could be describing the NP or the whole VP, but the VP is more likely.

Repetitions of the essay question should not be tagged, even if there are mistakes in the repetition. This is true for both native and non-native speakers. For example, native speaker 028 included a summary of the question. In addition, student 016 repeated the question, but made mistakes in the articles and noun phrases in part of it. Even though mistakes in repetitions are often used as diagnostic tests of students’ grammatical ability, these will not be counted. The students had the written text in front of them, so it could have been copied exactly, and it is therefore unclear what parts were exactly copied and what parts were produced from memory. Therefore, these will not be included in the consideration of noun phrase types.

Part 3: What relationships should be tagged and how

- **IDENTITY**—mark noun phrases that refer to the same real world entity as being in the same co-reference chain.
- Each NP should get its own ID number, and if appropriate, a reference chain number.
- Identity is symmetrical and transitive, so problems occur when the two sets are not exactly the same, or when there are functions functions (like temperature) that change over time
- The MUC-7 standards use SGML or XML, but this study will use an altered version of the MontyTagger tags.
- Each chain is marked with the same chain number. Each NP can only be marked as belonging to one chain.
- The ID and Co-ref chain numbers will start over for each subject. This is because they can still be referred to uniquely by listing the subject #. For each participant, then it will be obvious how many co-reference chains there are, and how many separate nouns.

APPENDIX F: STUDENT AND NATIVE WRITER

ESSAYS

L2 English Student Essays

<001>

<language=Korean>

<test=EPE>

<school level=G>

<TOEFL=217>

<classes held for=IIEP Writing, ESL Conversation, ESL Grammar, ESL Writing>

It is easy for us to contact to different culture in these days. So many people is meeting by on-line chatting and sharing their own information such as their hobby, foods and their life style. So I would like to explain about the culture borrowing in the several aspects.

First, Friendship is an good example of cuture borrowing. When we get known well to other persons, we often follow his speaking style. This is very important to our life. For example, many children learn their words from their friends That is, the relationship of their friends influence the words and behavior of them in the respect that friendship is a way to borrow another culture of my friend, Good relationship can influence to our life positively

Second, Borrowing the food of another culture is very substantial to our health. In the case of Kimchi, a traditional food of Korean, It plays a role of preventing the cancer from growing in our body. This case is benefitful for other countries' people.

</001>

<002>

<language=Korean>

<test=EPE>

<school level=UG>

<TOEFL=220>

<classes held for= ESL Grammar, ESL Reading>

In these days one of the important word which express the world is 'Globalization.' The earth become much smaller than long time ago, because of developed transportation, internet and so on. Many cultures from other country were mixed up, and Korea, my country is not an exception. Especially in parts of wearing and eating, Korea has been influenced from US. a lot.

One of the biggest changes of Korea is what we wear. About 100 years ago, every people weared Korean traditional clothes, but in these days there are no people who wear that in a street. People's wearing is almost same to western culture not only on outside. We also wear inner wear, too. There was no innerwear in traditional Korea. Western clothes are much more comfortable to behave than Korean one. It would be great advantage, but the fact that beautiful tradition is disappearing is so sad.

The other change is eating. Korean traditional foods are totally different from western one. People eat rise, not bread, and there were no chocolate, pizza, soda and

coffee. However in these days, everyone eat western food. In Korea There are T.G.Ifriday, Starbucks and Subway sandwich. To Korean people, western food is not strange thing, anymore. During the adaptation, western food was changed to be proper to Korean people. For example, the size of food, such as bugersize and cokesize is smaller and less fatty than western one. To cook a Korean food needs lots of time, so appearance of western food make people convenient, but Korean's bodies are not fit the fatty food, and Korean food is much better to health than western one, especially fast food. In these days in Korea, the number of fat people is increasing, and the reason of that must be fatty western food. It's significant disadvantage of borrowing culture.

Including housing, wearing and eating is most significant factor of one's life. In these part, Korea adapted lots of things from West. It's so

</002>

<003>

<language=Korean>

<test=EPE>

<school level=UG>

<TOEFL=247>

<classes held for=ESL Conversation, ESL Pronunciation, ESL Grammar, ESL Writing, ESL Reading>

The population in the earth, has increased gradually since the humans made the cities, nations, like that. Especially in Industrial Revolution started from Britain, the population increases dramatically all over the worlds.

But, now, in very high technological age, the situation is changed gradually. Because of development of medicine and science, the human can divide the kind of given chart by developed countries and developing countries, overall.

First, analyze the growth of developed countries such as Japan, Italy, Russia. In these countries, people can live so long and doesn't want many babies. So, population is getting smaller.

Second, the developing countries, India, Indonesia, Nigeria, Mexico, Saudi Arabia. In these countries, population will grow gradually until sometime. However, the grow rate of these countries maybe become smaller because of development.

Australia, U.S.A., China have special cases, I think. First two, are going to increase, although they are developed. Because, there are many immigrations from other countries annually. And, many Chinese goes to Korea, Japan and other countries to earn money. In addition China will be a developed coutry after 2025. It is easy to imagine because industry and economy of China is getting much bigger.

In conclusion, the population in our planet will decrease slightly after 2050 or later, according to the increasing rate of development of medical things or biology.

</003>

<004>

<language=Korean>

<test=EPE>

<school level=G>

<TOEFL=243>

<classes held for=ESL Conversation, ESL Writing, ESL Reading>

The graph shows us Today's population and expectation of future population for selected countries.

Now in 2000 among the countries, China has the greatest population, over 1.200,000 million. Second country is India, the country that have over 1.000.000 millions. third country is U.S that has over 300,000 millions. And next order is indonesia, Russia, japan, Nigeria, Mexico, Italy, Sudan, Saudi Arabia and Australia.

In 2025, the graph indicates changes. That is that some countries' population increases and others' population decreases.

For example, the countries that increase their population are, China, India, U.S, Indonesia, Nigeria, Mexico, Sudan, and Saudi Arabia. But Other countries' population decreases.

In 2050, India's population is bigger than China's population. That means India has the greatest population in the graph. The next country in china and then U.S, Nigeria, Indonesia and so on.

The specific country is Australia. The population of the country is stable.

In 2050, total population for the Selected Countries is 6,700,000 millions. In conclusion, the population of the countries increases slowly and is stable.

</004>

<005>

<language=Korean>

<test=EPE>

<school level=G>

<TOEFL=247>

<classes held for=ESL Conversation, ESL Pronunciation, ESL Grammar, ESL Writing, ESL Reading>

The below graph about projected population growth (PPG) includes many important things and then we can make use of this PPG to practice a good policy such as taxes, health insurance, educations, and so on.

First of all, the graph represents world population increases over the world in total basis. However, the population of Some developing countries such as India, Nigeria, Mexico increases greatly because of economic growth and resolution of the necessity of hungry. China is different from such developing countries. That is result of the strong policy of decreasing population by the government. Second, some advanced nations people tend to be reluctant to get the baby. They prefer to develop their identities and their dreams. Also, the other reason of the decreasing population is the weak economic growth and the people are afraid of giving birth to baby. But, the USA has different situations. The people have a strong belief on the economic growth which is followed by increasing population.

In conclusion, the graph gives us many points and a good resources to use. Also, we can realize that the population growth and the economic growth have a close positive relations.

</005>

<006>
 <language=Chinese>
 <test=EPE>
 <school level=G>
 <TOEFL=600>
 <classes held for=ESL Pronunciation>

The graph represents the projected population growth for selected countries such as Australia, Saudi Arabia, Sudan, Italy, Mexico, Nigeria, Japan, Russia, Indonesia, USA, India, and China. The graph tells us about the population number of different countries in 2000, 2025, and 2050.

From the graph, we can see that Australia has the fewest population and shows slowest increase of population growth among 12 countries. In 200, China has the most population and in 2025 China still has the largest population among these countries. But in 2050 the population of India will become the most one and it surpasses the population of China. There are seven countries showing the steadily increase of population such as Saudi Arabia, Sudan, Mexico, Nigeria, Indonesia, United States and India. Several countries shows the trend of slowly decrease of population, such as Italy, Japan and Russia. China will increase population from 1,200,500 in 2000 to 1,400,100 million population in 2025 and then decrease the population to the 1,400,000 million population in 2050.

I think due to the policy of one child of one family, China successfully control the increase of population. From 2025 to 2050, China will show the tendency of slowly decrease in population. India will keep the trend of steadily increase of population. Other countries such as USA, Australia has the steadily increase of population. I think it is because their concept. They don't need more children.

</006>

<007>
 <language=Chinese>
 <test=EPE>
 <school level=G>
 <TOEFL=227>
 <classes held for=ESL Conversation, ESL Writing, ESL Reading>
 Culture Difference

Along with the development of society, the culture difference between countries becomes smaller and smaller. The main reason of this change is our own culture has borrowed from another culture. In my opinion, it is helpful that we borrow and study the good sections of another culture.

My country is China which has a very long culture history. A long time ago, our country didn't accept the culture from other countries. Since we open the door of our country, so many people from other countries come to our country which bring all kinds of cultures. For example, there are a lot of food in China today. The Chinese people have more choices of food than before. This kind of change benefits our lives and let our food become more colorful.

On the other hand, there are more and more young people who are influenced by west culture. They become more open and do not have the ability to support family and society. It is estimated that the divorce rate in today's China is 25.9.

I think every culture has their own good thing which we need to study.

</007>

<008>

<language=Taiwanese/Chinese>

<test=EPE>

<school level=UG>

<TOEFL=220>

<classes held for=IIEP Communication, IIEP Reading, IIEP Grammar, IIEP

Writing>

This graph is showing the projected population growth for selected countries. To take the 2000 year as their basis and forecast the population of the 2025 year and 2050 year of each selected countrys.

In this graph I found that except Italy, Japan, Russia, and China, other selected countries have growth in population.

The forecast is including many elements and factors to influence the population growth result, like the development of the country, wealth of the country, geography, government event and so on.

Take Japan as an example. Japen is a developed country and their growth in population is decrease may shows that the work pressure is increasing. Another example is China, Because China has a potitical policy is that each family only has one child till 2000, so they will have a obviously growth after 2000, then, when 2050 the population get balance.

the other selected countries which is not big change in population growth may shows that their population almost balance if the economic and politic is going steadily.

</008>

<009>

<language=Chinese>

<test=EPE>

<school level=G>

<TOEFL=267>

<classes held for=none>

The bar chart below shows the expected population growth among countries studied. These countries can be divided into four groups, according totheir present populations in millions: (1) above 1,000,000 (2) above 400,000 (3) close to 200,000 (4) below 100,000.

Regarding Group (1), although in 2000 the population in China is almost 200,000 larger than that in India, India appears to have a much higher population growth rate than China. As a result, while in 2050 the population in China is expected to decrease to about 1,200,000, that in India will run up to almost 1,600,000.

Within Group (2), there is a similarly stable growth trend of population in United States and Indonesia. Both of them increase to nearly 400,000 in the year 2050, yet U.S tends to have more people all time through.

Among those in Group (3), Nigeria has an obviously rapid rate in population growth. The number of Nigerian people rises sharply from slightly above 100,000 in 2000 to nearly 400,000 in 2050. To the contrast, population in Japan decrease steadily, in the year 2050 there will only be about 100,000 people.

Australia, Saudi Arabia, Sudan and Italy in Group (4) all have much smaller population, compared to other groups, especially Australia, from 2000 to 2050, then to 2050, there will be only a mere increase in Australian people and Sudan people, Italians may even become fewer. Only in Saudi Arabia there will be a relatively higher rate of population growth.

To sum up, those nations who have a larger population are expected to gain more in future years. while those with little population will increase less. Consequently, the gap between them progressively widens. Among all the countries selected, India has the largest population growth rate, in the year 2050 its population may finally surpass that in China, to make India the nation with the largest population in the world.

</009>

<010>

<language=Chinese>

<test=EPE>

<school level=G>

<TOEFL=597>

<classes held for=ESL Grammar, ESL Writing>

The graph compares the population growth of 12 countries from different continents. According to the predicted data, the biggest country, China, will slow down its population growth after 2025 and will experience a small reduction. While the another big country, India, will replace China to be the biggest country in the world. Some other developing countries like Nigeria will also has a significant growth in their population. On the other hand, the population of those developed countries, such as Italy and Japan, will undergo a reduction in the following 50 years. The USA will keep its population steadily growing within reasonable rate.

It can be concluded, from the graph, that the developing countries will still have to face the big pressure from drastically growing population. Though in some of these countries, economy are growing rapidly, however, they still need to find the way the solve the problems caused by the expanding need from their people. According to the production, the US will keep healthily developing, since its popular growth remain in a reasonable range. The population structure of US is suitable to its developing speed, not as old as Japan and China who are getting more and more burden, and not too young, as India.

</010>

<011>

<language=Korean>

<test=IIEP>
 <writing rating=6>
 <IIEP level=P>

I would be a dictator of my country Korea, if I could be anybody for one day, because of the following three main reasons.

First of all, I'd distribute all the wealth in my society and nation equally to everyone of my country so that they could be well off irrespective of sex and age, rich and poor, noble and miserable. It's because this policy will help eradicate the sinful nature of men to steal, rob and even kill.

Secondly, I would order my cabinet to make a strict law for people not to use any kind of violence in their daily life under any circumstances. It's because every human being has a right to live happily and that to enjoy longevity enough until the day the Omnipotent had destined him or her to enjoy on the earth.

Lastly, I would set free all kinds of prisoners in jails in order to give them opportunities to become new people mentally. This action may also bring the desirable result later which will help me be re-elected President who is respected by all the people in my country. As you know, it is not easy to get voters' mind. From this perspective, I am sure much more voters will cast their polls for me, as they remember the favor I had shown to them.

In conclusion, all I want and hope is my fellow countrymen's wellbeing and my society's welfare. (Thank you for reading this funny and weird journal. :))

</011>

<012>
 <language=Korean>
 <test=IIEP>
 <writing rating=4>
 <IIEP level=P>

If I have this chance I want to be my father. Because I want to know about my father's thought and I want to know every father's thought. I think, after that, I can be a good father.

My reason is behaviors. My father did very good things and bad things what everyone does. Especially before my father gone, he did very well such as he knew his death. Why he did such a good thing and bad thing before he gone. long time ago, he didn't believe the GOD and he tried to abuse the God. He usually smoke and drink a lot. I think more than common people. Moreover, he was very strict person and he wanted be a perfect person as the God. And he usually say "No" when I wanted to do something.

However, before he die, he changed a lot. He believed the God and. he quit everything what he thought bad things. Is he knew about his death? I'm still wondering. If I know his thought, I want to explain to my mother about his behavior. Why he did like that.

In my guess, he was very smart and very proud of what he had, especially, he believed by himself. After he believed the God, he understood everything belongs to the God not to him.

Anyway, he is my pride and he is the most smart and brave person I've seen. Certainly I will see him in Heaven with my family soon. Because we love each other very much and he should see me right now with the God.

</012>

<013>

<language=Taiwanese/Chinese>

<test=IIEP>

<writing rating=5>

<IIEP level=P>

Yes, I was confused and afraid when I decided to leave my previous job which was a sales manager position. I worked for this company for eight years. I was promoted from an assistant to a manager from my hard working, 12 hours a day and 6 days a week. I enjoyed this job and actually learned a lot during the past years.

"No pay no gain", at the time I was enjoying what I've got, I found I lost more than I've got. I lost my health, family, friends.... just because I always put my job on the top priority, no holiday, no hobbies ...I lived with myself. Few months ago, I decided to leave this job and tried to relax myself for a while. Of course, learning English is my major thing in my things to do list.

We always say taking rest is for to last longer. I'm glad I made this decision. I gave up high salary and chose a real life what I want. Now my friends all support me and encourage me. I'm traveling frequently and spending more time with my family. I really love my life now.

</013>

<014>

<language=Korean>

<test=IIEP>

<writing rating=3>

<IIEP level=E>

If I can be anybody for one day, I want to be a very rich person. And I will buy everything I want and help poor people. I am going to spend all the money I have. I can buy a luxury car and a big house but I will not. Because I will become a normal person as I used to be, so they could be a burden for me. Luxury car and a big house cost a lot of money to maintain. Therefore, I am going to spend money buying property like building, land or stock. They could make me a rich after the day.

And I will help poor people. In the world, there are a lot of people dying of starving. Wherever I see them on T.V I was so sad and have pity on them. So if I am a very rich person for a one day, I will make an association that help them. And then I will send them foods and clothes, build houses, schools, and water-plants. Foods and clothes can not help them for a long time. But if they are educated and have land that they can grow crops, soon they will not be starved any more.

For these reasons, I will be a very rich person and I will

</014>

<015>

<language=Korean>

<test=IIEP>

<writing rating=4>

<IIEP level=P>

Life consists of a lot of challenges and obstacles. I always enjoy challenging the new objects. During my life, there were some events to remember. And these events made what I am. Especially one of them kept in my heart until now.

When I was a university student, I worked for the college newspaper. The experience in the newspaper gave me much such as the perspective of life, human relationship and research skill.

Some day of May, 1991, I visited the village for the poor. The village was made illegally by the poor. At that time, government tried to rebuild the new apartment. But residents of the village had no alternative, so they ardently and severely resisted the method of government. During the conflicts, one resident fired himself and the situation got worse and worse.

I was in that place. I watched him firing by himself. But I couldn't do anything to help him. I was an observer. For some days, I had suffered nightmare. In the dream he tried something to me and caught my hands, but I just stood and watched. What made me hard is not that I was scared but that I didn't try to do anything to help him.

After that event, I quit the newspaper. I thought that it was not helpful to the other people to report that it was not helpful to the other people to report the event indirectly. So I joined the social activity group for the poor labors. Real activity is really essential to the change of society.

In conclusion, the challenging spirit and real activities are pivotal to the life. From the event I got the active character and brave. My life decree is that "I just do it, not watch it". Thank you.

</015>

<016>

<language=Chinese (Taiwan)>

<writing rating=?>

<IIEP level=W>

Have you ever had an experience in which you were confused or afraid? Write about this experience and tell what learned from it.

It is true for many people there are many things to make them confused or afraid. I am the same with these people because I consider that as a human being, we always are confused and afraid when we arrived in the new environment.

For example, when I made my decision to come Iowa, I would be a little confused about some questions, such as where the Iowa is, how to go, and where to live. In the beginning, I would try to find out information about how to solve these problems. I become

confused and afraid. However, I understood this is a normal response.

Furthermore, during the period of time, I would suspect that whether I made a correct decision or not. However, I believed that I learn very much during this period of time. The feeling and participating of dealing with problems are very well.

As a result, I would be confused and afraid when I arrived in the new environment. Although I would be afraid, I still thought this duration was positive. This feeling is like finding light in the dark. I knew I had strength from my heart. Thought is the seed of action. The first step is always the difficult one. When I go through the difficult one, I will become less confused and afraid.

</016>

<017>

<language=Chinese (Taiwan)>

<test=IIEP>

<writing rating+2>

<IIEP level=E>

Many people have different important possessions. Some people think money and some people think foods...etc. In this essay, I will argue in favor of the most important possession is perseverance.

On the one hand, when I do one thing. I may have different troubles. Troubles will make me feel very sad and I won't do any more. If I have perseverance. I don't give up and I will continue again until complete. Nevertheless, I think perseverance is one especially important possession in my life.

On the other hand, perseverance will help everyone to reach success. When I have perseverance, I will encourage myself to meet anything. To help everyone with perseverance and to do things with perseverance ... etc. will help me with love and encouragement.

Last but not least, everyone has his own especially possession, but the most important possession that I have is perseverance. I will keep my perseverance to meet everything in my life.

</017>

<018>

<language=Chinese (Taiwan)>

<test=IIEP>

<writing rating=3>

<IIEP level=O>

Two years ago, I traveled with my best friend in America, and at that period of time, I realized something we call "afraid."

It was our first time to come abroad without any relative or friend, and other classmate or friend. Everything was exciting but

unknown. In that wild, strange country, we lived together, talked together, and even went toilet together.

Consequently, we got too closed, too depended to each other. Every tiny little problem became wierdly important. Our relationship got more sensitive. I started to be afraid. I was afraid to lose her, my best friend. At one night, I couldn't stand anymore and cried out loudly. She enbrased me tidly and said that it just the same thing she wanted to do. We

</018>

<019>

<language=Chinese (Taiwan)>

<test=IIEP>

<writing rating=4>

<IIEP level=P>

The wedding party is one of special ceremony in my country. In general, if a couple wanted to get marry, they have to prepare many things for their wedding party. First, the couple might discuss the wedding date with their parents. The wedding day usually be decided by the Lunar's suggestions. Then they need to book a wedding place with restrant. Second, they may need to prepare the things of wedding related such as wedding rings, formal suits, a pair of watch...and so on. The most important one is that they will to take some beautiful pictures for wedding memory and ceremony. Third, they have to inform their friends and relationships by phone or mail that make sure everyone be invited. When the wedding come, they will hold a formal dinner. All of the friends and relationshps join the dinner together. The guests will tell the couple their deeply blessings and give "red envelop" to the couple it means to wish everything goes their ways. Therefore, I always enjoy each wedding party and wish them my best.

</019>

<020>

<language=Korean>

<test=IIEP>

<writing rating=3>

<IIEP level=I>

The Korea

I'm from Korea. My country is very beautiful place. People are so kind and their pets too. We have many traditional things. First, Most famous thing is the food. Their are kimchi, bulgogi and chongkukjang---. For example, The Kimchi is most famous Korean food. It is very hot, but it is so cool. It makes red pappers and fresh vegetables. (green form) I love the kimchi. Second, I think

the clothes are famous too. It makes baby butterfly houses. This clothes are very pretty and beautiful. Many foreigners like the cloth. Oh! I miss the name of the cloth. We call it "Hanbok". 3rd, the architecture (house) is so wonderful! Our traditional houses are not so big, but it so organizational and beautiful. It makes many kinds of woods. So its colors are fantastic! I describe some of our country's traditional things. If you have a chance to go to another country, I want you to go to Korea. You are not disappointed. Thank you. :)

</020>

Native English Writer Essays

<021>

<language = English>

A Letter to the Editor

Specialization is the key to success in America today and our state's insistence on requiring undergraduates to "broaden their knowledge" by taking general education classes is sadly mistaken. Basic core classes merely introduce students to important ideas without allowing them the time needed to explore these concepts and develop a true and deep understanding. This kind of knowledge is important for appearing worldly at dinner parties or job interviews but is essentially only window-dressing. The number of classes a student can take in their four years is limited.

</021>

<022>

<language = English>

In Favor of General Education Requirements

General education requirements at a college or university, when done well, can provide a common core or base for specialized studies. They can also provide a common foundation for people studying different specialties where this foundation is a bridge between the people and their specialties. These general education requirements, especially if completed early, can also expose the students to areas of knowledge to which they would otherwise not be introduced. This exposure can result in changes in the directions of study that they pursue. All these benefits come at the relatively modest price of about twenty percent of a student's time. The benefit to broader exposure to learning and a firmer base of common foundational knowledge is much greater than investing this twenty percent in additional specialized study.

As a case study of the benefits that can be derived from general education requirements, consider their benefits in an engineering education. Fully

eighty percent of an engineering education is devoted to specialized studies in engineering disciplines, physics, chemistry, and mathematics. Yet engineers and their professional societies have found that the twenty percent of the engineering curriculum devoted to such general education topics as English, history, economics, philosophy, psychology, and foreign languages is a lifelong good investment. These subjects provide essential insights and skills to successful engineers, allowing them to interact with and understand non-engineers (who are usually the people for whom the engineers are developing things) and to work in interdisciplinary teams.

There are also non-professional benefits accruing from general education requirements. Well done, these studies can broaden the student's cultural understanding and enhance their quality of life throughout their entire lives. General education requirements can result in life long avocations.

The overall benefit of general education requirements is a broadening of perspectives and the capability of building closer ties with others and understanding them better.

</022>

<023>

<language = English>

Why General Education Requirement Should Be Retained

I strongly feel that general education requirements should be retained in all higher education institutions. Being truly well educated is a matter of knowledge on a broad scale, as well as a broad knowledge of a particular subject. Whether you choose to continue to study a particular subject or not, your life is enhanced by knowledge of many different subjects. Furthermore, it has been my experience students often change their majors as they begin to delve into a particular field of study or become aware of alternative fields available. In my particular case I majored in a subject that I took initially as a general education subject, which I found interested me and, at the same time, I discovered that I really disliked the field I had thought would be my major. Other subjects which I studied for general education requirements, such as art, geology and philosophy have greatly enhanced my life.

</023>

<024>

<language = English>

The elimination of general education requirements in public universities would be an incredible disservice to the impressionable minds seeking to expand their academic horizons.

The benefits to a well-rounded education are numerous. Often students enter college unsure of the path they would like to take but understand the importance of a college degree. Requiring that students explore all avenues of study would allow those undecided to make a better informed decision of the field of study they wish to pursue.

What about those students who are sure of their education path? Would general classes not divert this path and take time away from the main area of interest? While this point does have merit, an opposing point could be made that a student who thinks broadly about their fields is able to better interconnect to a larger issue. It is rare that one problem does not span multiple disciplines.

In addition, a more general education will allow one to be more adaptable if faced in the future with a slim market within one's major field of interest

Finally, there is the personal enrichment gained from a broad knowledge of many subjects. An appreciation of multiple disciplines can offer the student a balance when the major field of interest becomes overwhelming.

The elimination of general education requirements in today's public universities would deprive students the joy all disciplines have to offer.

</024>

<025>

<language = English>

Dear Editor,

Recently, I learned of the public universities' proposed plan to eliminate the general education requirements currently mandated for all students. These requirements include such basic courses as composition, math, and science. This proposal reflects short-sighted thinking on the part of the universities: while students in the short-term will rejoice as they abandon the "tedium" a much-loathed subject, in the long-term, these students will regret that they do not have the skills to write a letter, balance their checkbook, or make the most basic repairs to their homes.

More progressive universities have considered this short-term/long-term trade-off and have retained, and even added, general education requirements. For example, Yale Medical School instituted a humanities course for all medical school students. This course is meant to enhance students'

appreciation of what it means to be human-to introduce a personal aspect to the science of medicine. The University of Virginia requires proficiency in the sciences and humanities as well. To address the concerns of students for whom a subject might be particularly difficult, the University offers classes targeted to these students. “How things work” is physics class for students who are interested in science but not prepared for a course in quantum physics. Still, students in this course leave with an appreciation of the main physics concepts and, most importantly, analytic skills that can be applied to other disciplines.

These two examples provide two arguments in favor of general education requirements: 1) exposure to another, disparate perspective and 2) the development of analytic skills that are not addressed by other disciplines. The third justification in support of general education requirements is perhaps the strongest: the general education requirement for proficiency in composition demands that all students, regardless of the discipline that interests them, are proficient in constructing a letter, writing a proposal, or pitching a product. Without suitable writing acumen, even the most astute scientist will not be able to communicate and share his ideas. In fact, without this skill, this scientist (I’m an epidemiologist) would not even be able to express my opinions through this letter.

General education requirements are essential to the long-term success of students. Vote against their abolition. Thank you.

</025>

<026>

<language = English>

To be an educated person is to be educated in the Liberal Arts as well as the major field of study. The reasons are legion.

It is critically important if we are to retain (at this point, regain) our representative democracy. We must have a reading, critically thinking, articulate, independent citizenry. We must know what the Constitution and Bill of Rights, et al, state and how they have been interpreted. We must understand how the three branches of government cross-balances are to function. We must have history classes. Those who ignore history are doomed to repeat it.

That same argument is why we must understand world history, warning signs and what did work in the face of all the problems.

In order for us to function in this world economy, we must study their cultures, We need social studies courses. Each of us needs to understand economics. We also need that knowledge for our own personal financial responsibilities.

In today's world of corporations controlling our health care, we must understand human biology and chemistry. We have to advocate for ourselves. The hands of our physicians have been tethered to some extent. We must know how to do research and think critically. We also need to understand the world's environment, What is happening anywhere on this planet impacts each of us and our future generations. We have to figure out what we can do to impact this for the betterment of generations.

The whole person has many needs in order to understand and appreciate life. The whole person needs many ways of expressing him/her self. Each person's needs are different. We must understand the value of the arts to the human species and to the world. Understanding other cultures shows us how they have expressed themselves.

A person who is learned is a humbler person, a person who has learned to listen, far more equipped to function in this world.

</026>

<027>

<language = English>

Like most simplistic solutions, this won't work. It will also penalize people under 21 who have legitimate reasons for being in bars. It is also too vague to be enforceable unless bar is defined.

Why won't it work?

- * Most underage drinking does not take place in bars (30 minutes isn't enough time to look up references and provide documentation, but I'm confident that the statement is correct.)
- * Those who currently use fake IDs to buy drinks will use them to enter bars.
- * Establishments that don't require proof of age to buy drinks won't require proof to enter.
- * If "bar" is narrowly defined as a place that only serves alcoholic beverages i.e.-no food, no music, no other entertainment and no amusement devices (e.g. pool tables, pinball machines, chess boards), it will be circumvented by the addition of minimal food service (as simple as microwave pizza), or a pool table.
- * If "bar" is more broadly defined as a place that serves liquor, pool halls, comedy clubs and pizza parlors will be off limits to everyone under 21 [even if accompanied by a parent or spouse?] In that case local businessmen-and ordinary citizens-will ignore the law.

</027>

<028>

<language = English>

letter to the editor for retaining general education requirements

Dear Sirs:

The current proposal being considered by the state legislature to eliminate general education requirements should be rejected. The traditional arguments for such programs, based on assumptions that well rounded citizens are better for society, still make sense. Absent compelling evidence to the contrary, the current requirements should be maintained.

It is certainly understandable that at any given moment in history, proponents of teaching only that which is of direct value to the career a student intends to pursue will suggest such curriculum revisions. When national insecurity about falling behind internationally in hi-tech rises, willingness to give short-shrift to history, English, and arts will rise. But the temptation to endorse producing graduates who have not been exposed (forcibly, if necessary) to what is called a liberal education should be resisted. Our society is better off if our scientists and engineers have at least had a passing introduction to the great thoughts and ideas of past philosophers, the great deeds of past leaders, and the great works of past artists. Even if the scientists and engineers look forward to being free from such thoughts after graduation, one would hope that they would have a better perspective to assist them in their work, as well as in their functioning in society.

The requirements in the opposite case--English majors taking science, for example--are also justified, even if the traditional course requirements are not particularly challenging. Exposure to the scientific method is useful, even if its beauty is not always absorbed by the non-science student. Math instruction is useful, even if painful for many non-quantitative types.

The proposal to eliminate general education requirements should be rejected. If anything, the current requirements should be strengthened.

Sincerely,
<028>

<029>
<language=English>

Far from diminishing current general education requirements, our state universities should consider reinstating many of the gen ed requirements abolished in the late 1960s and 1970s. Students who graduated from college in the 1950s and 1960s have a much broader general knowledge base than those who graduated in the 1980s and 1990s. There is a certain level of general knowledge that every college graduate should have and general education requirements help to achieve this.

My own education highlights the limitations of recent general education requirements. I was a history major at a prestigious liberal arts college in New England and graduated in the 1990s. Western Civ was not a requirement for even history majors and based on general education and major requirements, I could have graduated from college without taking any non-American history courses. While the ability to focus on my area of interest was appealing to me, I realized that American history is best understood within the context of world history so I took one course in medieval England, one in twentieth century Chinese history and one on Europe during World War II. However without a required Bible course or a more general Western Civ course, I find that college graduates who are fifteen to twenty years older than I am know far more about general world history than I do, regardless of what their majors were. That doesn't say much for the educational requirements of the institution I attended. Our required writing course which could be in almost any subject--history, literature, science, economics, mystery writing--however did provide me, and most of my classmates, with stronger, more critical writing and analytic skills than many of my professional colleagues seem to have.

The ability to write a concise letter, essay, email, presentation outline or project proposal is a valuable skill in any field. Allowing a science or math major to graduate without learning the skills necessary to write a well-thought out and clear document short-changes the student and the future of his/her job field as well.

As important as history and writing are, math and science are equally valuable skills for students to learn. Being able to figure out the tip at a restaurant, how much you will save when something is 15% off, just how much you need to save to pay for that car, trip, or house is something that few people are comfortable with today. A knowledge of basic math is critical to understanding and managing your finances.

</029>

<30>

<language = English>

Dear Editor:

There has been a great deal of debate in recent weeks about the elimination of general education requirements at state universities. I, myself, was able to completely opt out of these requirements as an undergraduate and I loved it. It allowed me flexibility to explore courses I may not have had the chance to otherwise take and it allowed me the room in my schedule to obtain a double major. I was also able to completely avoid take a basic science, which was a bonus. Yet, I still feel strongly that I had a well rounded education. That said, I would still advocate that universities do not do away with these requirements completely, but that they be more thoughtful about the ways in which they are implemented.

Having some general education requirements helps ensure a minimal level of proficiency in certain important areas. For example, when I was a graduate teaching assistant I found that many undergraduate students did not have very basic writing skills and I often struggled to grade their work because they were not able to articulate what they wanted to say. Also, while I loved that I did not have to take classes I did not want to take, there is something to be said for making sure that students obtain breadth in their education.

The struggle for many students is that there are so many general education classes and so many required classes for one's major that few students are able to double major, and many find themselves taking expensive summer school classes just to be able to fit everything in. It then becomes difficult not to feel that the high level of required coursework is simply a way for the universities to make more money. Students also often complain they cannot get into all of their required classes and that many of the general education courses are mundane and useless.

The general education requirements should be structured in such a way that they allow for both a well rounded education and the flexibility to explore new areas of study. Students should also be able to opt out of classes that are not useful to them if there is some means for ensuring their basic proficiency in the subject. In addition, the range of courses that will allow students to fulfill these requirements should be broad so that students are still afforded the opportunity to explore. This is the difference between taking a science and an English class and taking a set of one size fits all, prescribed courses.

Sincerely,
</030>

<031>
<language=English>
Dear Sir or Madam:

I am writing in support of the current general education requirements that exist in our state schools. One of the fundamental principles of a liberal arts education is that students should receive a well-rounded education in a variety of subjects. General education requirements are the means by which this goal is accomplished. I believe a

well-rounded education is important to a well-educated and informed citizenry and is key to ensuring that our students are prepared for the dynamic and changing workforce.

Employers today require people who are capable of functioning in a variety of roles. If students are exempted from general education requirements, what means is there to ensure the talented math student also knows how to express him or herself effectively. Writing skills are essential for effective communication. In the same manner, an English student who is at a complete loss in understanding basic scientific phenomenon or mathematics may have difficulty grappling with new technology. A sudden change in their company could require a different skill set than that which they were originally hired for, and those who will prosper are those who have the ability to be flexible and adapt accordingly. In our global world, foreign language skills are important both in and outside the United States. In California where I live being bilingual is an enormous asset, no matter which career path a person chooses to pursue.

The metaphor can be extended into basic life skills. Students will encounter science and mathematics in every day life, whether it's when they're receiving medical advice or trying to finding a mortgage that is right for them. Conversely, a scientist who is at a loss when it comes to history or foreign languages will have a much poorer understanding of the world in which he or she lives and works.

The general education requirements are not onerous, nor do they demand students become specialists in fields outside their primary area of interest. They merely require students be exposed to a broader world than they may have chosen to do without this requirement. Students who take AP classes in high school are able to count these credits towards the requirements, freeing up their schedules even further.

I hope that the leaders of our university system recognize the benefits of adhering to the tradition of a liberal education and defend that tradition vigorously.

Sincerely,
</031>

<032>
<language=English>

Dear editor:

There are many benefits from having an educated society. I strongly support continuing to require general education requirements as part of baccalaureate degree program. Currently only about 1 in 4 adult Americans have earned a college or university degree. Even some of those may not have taken some general education courses. However, each adult American has the opportunity to use many of the fundamental concepts which comprise each general education course in our everyday life. Most of us take our educational accomplishments for granted as well as the concepts we learned as part of a degree program.

Core values such as responsibility, acceptance, tolerance, and many others come in part from having the knowledge and understanding which we obtained from general education courses. A few examples may illustrate how we use general education course contents daily.

All Americans must use debit or credit cards, checks, or pay with cash. The use of any of those requires basic math skills. Certainly, a bit higher skills are required to understand interest expenses, late penalties, mortgage calculations, and other such concepts. While not usually aware of it, we use math when driving as we calculate time required to/from work given distance traveled and speed limits. We use math in estimating stopping distance, time available to enter traffic flows safely, etc.

Use of the English language provides an immediate impression when first meeting and visiting with someone. Misuse of subject-verb tense, excessive or under use of articles, and similar common errors in speaking leave the impression of an educated or uneducated individual. Most of us would much prefer to leave a positive first impression rather than a negative one.

Americans are among the worst people in the world in terms of our ability to read, speak, and understand more than one language. Many international students I have advised can communicate in several languages, often a native language, a national language, and English, an international language. Rather than eliminating foreign language requirements, we need to take a cue from the rest of the world and require more foreign language courses. Of course, first we should increase our use of our native English language.

Science, too, is a necessity in our world today. Medicine and health care is a good example. Scientists continually learn more about the basic functioning of our bodies, the interaction between health and exercise, diet, and attitude. More, rather than less, science is required for us to understand and communicate clearly with health care professionals, to be part of the health care team that cares for and the treats us. Basic science courses can assist us by providing a base of understanding on which we can build as we continually read and hear more about scientific advances in the health and medicine field.

Lastly, an old adage is that history repeats itself. Only with a thorough understanding of history can we recognize the factors or patterns leading to repeating history. Certainly, there are many aspects of our history, especially international conflicts and atrocities, that many of us would hope are never repeated. How to avoid those in the future requires understanding what led to them in the past and what alternatives are or were available that could have changed history and could change the future.

Public universities are charged with providing students with an education, not just teaching an employable skill. Public universities are charged with teaching students to think and to think critically. Certainly, students need more than basic or general

education courses in order to obtain quality employment. But first we need to provide students the opportunity to have a well-rounded education, which can be accomplished by requiring a set of general education courses. From this base, students can delve further into topics of interest and we can deepen their understanding and generate new knowledge over time.

It is that creation of knowledge that will lead all of us to have a better quality of life in the future.

</032>

<033>

<language=English>

My personal feelings on the matter are thus; Lasie Fair. I have no idea how to spell that but I think it means “Hands Off”. I don’t feel this way about all government intervention but on this issue I do. It seems to me that the owner of each individual establishment has the choice to decide what is best for business and how best to serve their clientele. Clearly they are bound by the current laws of the state that dictate that no one under the age of 21 can consume alcohol in there establishment. To my knowledge there are no restrictions on members of society being present when others are consuming. I feel as if the bars that choose to allow minors to frequent them are aiming for a certain demographic and they end up getting it. This does require them to be on a higher level of vigilance to make sure that friends “of age” are not provided minors with alcohol. The bars that choose not to deal with such issues simply limit who can come and go based on age. I think it is nice that there are some places that groups can go regardless of the entire groups age. Those of age can consume alcohol and those not of age can still share in their company. It also allows for an early look at the widely varying effects alcohol can have on people.

I think there is a lot of substance to the argument that if minors are completely excluded from the bar scene that it does drive them to visit house parties where little if any supervision is taking place. If the city was looking to ease the public pressure to do something they might find that the increase in police house calls negates any pressure that might have been relived in the public’s eye.

Allowing minors to be exposed to alcohol consumption before they are of age in a controlled environment such as a public bar would be the best possible of either situation in my mind.

</033>

<034>

<language=English>

The recent information provided in this paper about the University considering eliminating the general education requirements for every major course of study is a

mistake. I, too, thought when I attended college that many courses were a waste of my time, as I majored in engineering, which is a very technical course of study.

Do engineers really need to learn grammar, punctuation and English? You bet they do!!! If an engineer cannot communicate effectively in his job, it will hinder advancement and effective project management. Engineers should especially learn history and foreign languages as well. An old saying, "History is doomed to repeat itself," comes from not knowing history.

I think a well rounded education is the best way to prepare for future success. Many people have criticized our younger generations for not knowing how to spell, properly form a sentence or use correct grammar. Is it the fault of the student or the teacher? If general education courses are removed from technical curriculums, then I am afraid American education will continue to fall behind the rest of the world. Engineering students complain about their course of study being too long and difficult, but if general education requirements were removed, the courses that replace them would be more technical and specific. This is not the way to produce a well educated person.

My only regret now in later life is not taking enough "general education" courses to learn more about life!!

</034>

<035>

<language=English>

Dear Editor,

I have read about the debate concerning the elimination of general education courses at the college level. Even though many jobs in today's society require specialization in the field of work, a base of understanding about many areas of study are necessary for individual's to be able to effectively communicate with people from various backgrounds and with various jobs. General education courses must be retained.

General education courses offer the following benefits to incoming college students. First, the exposure to areas of study in which the student may have had minimal exposure to prior to the college experience is an advantage. The basis of such exposure allows students to see connections between all aspects of life. People integrate subject area information daily and the integration is strengthened with a broader knowledge base. Secondly, As a student's knowledge base is broadened, the student may be exposed to a field of study that he/she would like to pursue further and thus might consider a focus of study and career choice that was not expected. Thirdly, many incoming college students lack study skills necessary for successful classroom experiences in college. General education courses are ideal for teaching students how to read material (e.g., textbooks, professional journal, reviews) written for a specific content area, how to learn various types of content material, and how to communicate using the language of that content area. Lastly, Students can learn to take risks and move away from their comfort zones.

They will meet people who are different from them not only culturally, but in the way they learn and approach learning. Risk taking requires a certain degree of flexibility, openness, and critical thinking.

</035>

<036>

<language=English>

Dear Editor,

I am writing to address the issue of removing general education requirements at the University of Maryland. A graduate student with a dual undergraduate degree, I am able to remember my disdain for subjects other than those I found interesting.

My college grouped subjects into three general categories, and required each student to take two courses from each category. For me, this was akin to a jail sentence: I had to sit through not one, but two semesters of dreaded science courses? Even though I was able to achieve good grades, I did not enjoy the material. How could I be forced to take math and science classes when my chosen profession would have nothing to do with those subjects? With utmost clarity I was sure that I would work squarely in the arts, never having any exposure to the sciences. I severely underestimated the way in which I would constantly face ideas and issues from across the academic disciplines.

Now, having been an editor in a government agency specializing in economics and at an association for healthcare professionals, I've frequently lamented my short-sightedness in not taking more courses in other areas that would have rounded out my general knowledge more. Economics especially is an area in which I feel lacking, because I feel at significant disadvantage in planning for my future retirement without fully understanding the market and economic forces.

Forever will I be grateful to the board at my undergraduate college for having the real-world insight to keep and indeed strengthen the requirements for cross-discipline education, for thanks to them I learned the skill to confront materials in which I had no true interest and to engage that material.

With so many subjects available at the university, it is easy to see how a student may be overwhelmed and retreat to a comfortable subject in which she is interested. It is the responsibility of the university to recognize this reality and both organize the many disciplines available and compel their students to learn from a wide variety of these disciplines. This is true of professional graduate programs, which recognize that while a student is participating in the program to learn more about a specific discipline, it is just as important in the development of new professionals to offer courses that address related issues. For example, my master's in library science program includes management and technology courses in addition to the traditional research and cataloging courses.

In the real world, the less an individual knows, the more that individual becomes disconnected from the social discourse, and the more powerless that individual becomes.

That individual has less solid knowledge on which to base analysis and decision-making of new ideas. Exposure to other disciplines makes one more able to participate in the public discourse about all manner of issues. How can those with no knowledge of the myriad issues in the world possibly be expected to make well-informed decisions?

There is a direct causal relationship between the educational level of an individual and the ability of an individual to engage different ideas and issues. There are both personal and professional reasons for an individual to be as well-rounded as possible educationally. As a society, we are responsible for making sure that the next generation is educated well enough to make knowledgeable personal and public decisions about new issues,

</036>

<037>

<language = English>

Underage drinking is a serious public health concern, associated with increased risk for mortality, academic failure, and mental health problems. I believe that making it illegal for people to enter bars unless they are over 21 years of age would be an effective way to cut down on underage drinking among college students. This would make it more difficult for underage students to gain access to alcoholic beverages. It would also cut down on number of underage individuals driving home from bars, potentially reducing drunk driving in this population. Many nightclubs are already successfully utilize this policy.

Those who are against this proposal have suggested that this action would hurt local businesses and negatively affect neighborhoods by increasing the number of disruptive house parties where alcohol is served. First, this action will not significantly impact law-abiding local businesses, as they should not be making very much money from individuals under 21 if they are not serving them liquor. They may experience a slight decrease in revenue from food sales. However, the benefits to society of reducing underage drinking outweigh this concern. Also, the many nightclubs that only allow individuals over age 21, have managed to remain profitable despite this policy.

Secondly, I believe that disruptive house parties will remain common in areas occupied by college students, but will probably not become significantly more frequent. Furthermore, underage drinkers will be just as disruptive in bars (and on the streets on their way home from the bars), as in neighborhood homes. In fact, the risk for drunk driving is much more disruptive and potentially fatal than loud noise in the neighborhood.

Although the policy would be inconvenient for many people under 21 who wish to go out with their older friends, I believe it would be a beneficial policy from a public health perspective.

</037>

<038>

<language = English>

Recently, the state has been considering eliminating general education requirements at the university. Some time ago I was an undergraduate at the state university and didn't like being required to take math and P.E. class; however, with more experience and perspective I have come to realize that these requirements serve a valuable and fundamental purpose. The reason for these requirements goes to the very heart of the purpose for public higher education. While the purpose of higher education has been disputed, I still hold that it should accomplish two goals: (1) to provide a well-rounded education which will make students better citizens and (2) to give students specialized knowledge to contribute in some field.

These two purposes are served by the requirement. First of all, to be well-rounded students should know more about math, foreign languages etc. then they have learned in high school. Having these courses required also students more choice that they had in high school and they may take classes in East Asian History or logic. This broadening of basic knowledge can make students more open-minded, more flexible, and better informed to participate in the "workforce" and the world community.

Furthermore, being specialized in a field of study should include the broad base on which general education courses are built. While I wondered often why I had to take calculus as a psychology major, I learned later that the logic and problem solving strategies I practiced in proofs helped me when explaining difficult grammatical structures to my students as an ESL teacher. Certainly a foreign language such as Spanish can be helpful in many fields because of it's base in Latin; for example, medical students learning anatomy may be aided in having this language base.

Another reason for supporting these classes is to assure all students equal footing. Since high school students do not all have equal quality and quantity in coursework, the university can provide what these students may be lacking. As the state schools are public institutions, they should consider the articulation they provide for students post high school. Related to this, students at the university may still be exploring what to specialize in. If their high school did not offer quality science courses, maybe because of limited funding for labs, then taking science in college could help them find something which interests them more than the realize.

A last point, I would add it that general education requirements include a lot of choices and options. Many classes can count for social science credit from developmental psychology to art of Africa. So, there may be a misunderstanding in this debate that general education requirements force everybody to take one physics or algebra class.

In sum, the state university should continue to require general education courses. It serves the very purpose of public higher education.

</038>

<039>

<language = English>

I have read with chagrin that the public universities in this state are planning to eliminate general education requirements. I think that this would be a major mistake because it would, for many students, turn higher education into nothing more than a vocational school. Students would go to school to learn a marketable skill and nothing more. The value of college for young people is that it puts them in touch with the world at large and gives them at least some knowledge of various fields of human endeavor. An enlightened human being must have a broad knowledge of many fields in order to make wise decisions in his or her own field.

Not to be partisan, but the current president has little or no intellectual curiosity and has only a limited knowledge of many of the issues he must make decisions about. Because of his limitations, he often decides with his "instincts" or his "gut", not through careful analysis. We can see the catastrophic result of this way of interacting with the world.

Students in college need to be exposed to the great ideas that man has developed over the years. If they are unaware of the seminal works of philosophy, political science, art, literature, etc., they are diminished as human beings. The legacy of these giants that came before them is essentially wasted. In order to reach his or her potential, a person must have a broad knowledge of the world, not just a narrow knowledge of one vocation. We need to have more Renaissance men and women in this country and fewer technocrats.

</039>

<040>

<language = English>

Dear Editor,

I am writing in support of preserving the practice of requiring undergraduate students to study a variety of subjects at the college level beyond their personal interests or intended majors. A specialized knowledge of a particular field will ultimately benefit students professionally. Luckily, the system of higher education already fulfills that need. A society which wishes to identify problems and investigate solutions needs thinkers with general reasoning and creative problem solving skills. This sort of development comes from a diverse education where students are stretched beyond their personal interests and talents. General education requirements are not the perfect solution; many students are not engaged in these courses they view as a continuation of high school. Instead of eliminating the requirement, I propose that general education courses be more tailored towards current world issues (i.e. science focused on environmental change instead of chemistry).

Sincerely,
Concerned Citizen

</040>

REFERENCES

- Adger, D. (2003). *Core Syntax: A Minimalist Approach*. New York, NY: Oxford University Press.
- Al-Raheb, Y. (2006). "Pragmatic Discourse Representation Theory". *Proceedings of the Analyzing Conversations in Text and Speech (ACTS) Workshop at HLT-NAACL 2006*. New York, NY: Association for Computational Linguistics. 58-61.
- Anderson, R. and Souto, S. (2005). "The use of articles by monolingual Puerto Rican Spanish-speaking children with specific language impairment". *Applied Psycholinguistics*. Vol. 26. 621-647.
- Ariel, M. (2004). "Accessibility Marking: Discourse Functions, Discourse Profiles, and Processing Cues". *Discourse Processes*. Vol. 37, No. 2. 91-116.
- Ariel, M. (1990). *Accessing Noun Phrase Antecedents*. New York, NY: Routledge, Chapman & Hall.
- Baauw, S., de Roo, E., and Avrutin, S. (2002). "Determiner Omission in Language Acquisition and Language Impairment: Syntactic and Discourse Factors". Ed. by Skarabela, B., Fish, S., and Do, A. *Proceedings of the 26th Annual Boston University Conference on Language Development*. Vol. 1. Somerville, MA: Cascadilla Press. 24-35.
- Baker, M. (2003). *Lexical Categories: Verbs, Nouns, and Adjectives*. New York, NY: Cambridge University Press.
- Bamberg, M. (1987). *The Acquisition of Narratives: Learning to Use Language*. New York, NY: Mouton de Gruyter.
- Barnett, R. et al. (1996). "EAGLES Preliminary Recommendations on Subcategorization: EAG-CLWG-SYNLEX/P". Ed. by Antonio Zampolli Institute of Computational Linguistics. <http://www.ilc.cnr.it/EAGLES96/synlex/synlex.html> Accessed January 14, 2009. Pisa, Italy:
- Beakley, B. (2007). "Grice's Conversational Maxims". <http://www.ux1.eiu.edu/~cfbxb/class/1900/prag/grice.htm> Accessed July 12, 2008. Charleston, IL: Eastern Illinois University.
- Beck, M. (1998). "L2 Acquisition and Obligatory Head Movement: English-Speaking Learners of German and the Local Impairment Hypothesis". *Studies in Second Language Acquisition*. Vol. 20. 311-348.
- Bernardo, J.M. (2003). "Bayesian Statistics". Ed. by Viertl, R. *Encyclopedia of Life Support Systems*. Volume: Probability and Statistics. Oxford, UK: UNESCO. 1-46.
- Berwick, R.C. (2003). "6.863J Natural Language Processing, Lecture 4: From finite state machines to part-of-speech tagging". *February lecture notes*.
- Biber, D., Conrad, S., and Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. New York, NY: Cambridge University Press.

- Boland, J. and Blodgett, A. (2001). "Understanding the Constraints on Syntactic Generation: Lexical Bias and Discourse Contingency Effects on Eye Movements". *Journal of Memory and Language*. Vol. 45. 391-411.
- Bowen, R. (2005). *Noun Complementation in English: A corpus-based study of structural types and patterns*. Goteborg, Sweden: Acta Universitatis Gothoburgensis.
- Brill, E. (1994). "Some Advances in Transformation-Based Part of Speech Tagging". *AAAI Proceedings*. <http://www.informatik.uni-trier.de/~ley/db/conf/aaai/aaai94-1.html#Brill94> 722-727.
- Brill, E. and Mooney, R. (1997). "An Overview of Empirical Natural Language Processing". *AI Magazine*. Vol. 18, No. 4. 13-24.
- Bruhn de Garavito, J., and White, L. (2000). "L2 Acquisition of Spanish DPs: the Status of Grammatical Features". Ed. by Howell, S.C. et al. *Proceedings of the 24th Annual Boston University Conference on Language Development*. Vol. 1. Somerville, MA: Cascadilla Press. 164-175.
- Charniak, E., Hendrickson, C., Jacobson, N., and Perkowski, M. (1993). "Equations for Part-of-Speech Tagging". *Proceedings of the Eleventh National Conference on Artificial Intelligence*. <http://crl.nmsu.edu/~raz/Ling5801/papers/pos/eqfuo> Providence, RI: Department of Computer Science, Brown University.
- Chen, P. (2003). "Indefinite determiner introducing definite referent: a special use of 'yi 'one' + classifier' in Chinese". *Lingua*. Vol. 113. 1169-1184.
- Chinchor, N. (1998). "Overview of MUC-7/MET-2". http://www-nlpir.nist.gov/related_projects/muc/proceedings/muc_7_proceedings/overview.html Accessed August 24, 2008. Science Applications International Corporation.
- Chomsky, N. (1993). "On the Nature, Use, and Acquisition of Language". Ed. by Goldman, A.I. *Readings in Philosophy and Cognitive Science*. Boston, MA: MIT Press. 511-534.
- Conrad, S. (2002). "Corpus Linguistic Approaches for Discourse Analysis". *Annual Review of Applied Linguistics*. Vol. 22. 75-95.
- DeRose, S. (1988). "Grammatical category disambiguation by statistical optimization". *Computational Linguistics*. Vol. 14. 31-39.
- Delmonte, R. (2000). "How to Annotate Linguistic Information in FILES and SCAT". http://www.ilc.cnr.it/tressi_prg/papers/delmonte.p Venice, IT: Universita' "Ca Foscari". 1-10.
- Diaconescu, R.C. and Goodluck, H. (2004). "The Pronoun Attraction Effect for D(iscourse)-linked Phrases: Evidence from Speakers of a Null Subject Language". *Journal of Psycholinguistic Research*. Vol. 33, No. 4. 303-319.
- Dressler, W.; Stark, H.; Vassilakou, M.; Rauchensteiner, D.; Tosic, J.; Weitzenauer, S.; Wasner, P.; Pons, C.; Stark, J.; Brunner, G. (2004). "Textpragmatic impairments of figure-ground distinction in right-brain damaged stroke patients compared with aphasics and healthy controls". *Journal of Pragmatics*. Vol. 36, No. 2. 207-235.

- Eckardt, R. (2006). "Discourse structure, pronouns and more".
<http://wwwuser.gwdg.de/~reckard/> Gottingen, Germany: Department of English Language and Literature, University of Gottingen. 1-6.
- Engdahl, E. (2001). "The Role of Syntactic Features in the Analysis of Dialogue". Ed. by Rohrer, C., Rossdeutscher, A., and Kamp, H. *Linguistic Form and its Computation*. Stanford, CA: Center for the Study of Language and Information. 111-142.
- Garrette, D.H. (2008). "Discourse Representation Theory". Ed. by Bird, Klein, and Loper. *Natural Language Processing (Draft)*. 1-8.
- Goad, H., and White, L. (2004). "(Non)native-like Ultimate Attainment: The Influence of L1 Prosodic Structure on L2 Morphology". Ed. by Brugos, A., Micciulla, L., and Smith, C.E. *Proceedings of the 28th Annual Boston University Conference on Language Development*. Vol. 1. Somerville, MA: Cascadilla Press. 177-188.
- Goddard, C., editor. (2008). *Cross-linguistic semantics: explications, metalanguage, grammar*. Philadelphia, PA: John Benjamins Publishing Co.
- Granger, S. (2002). "A Bird's-eye View of Learner Corpus Research". Ed. by Granger, S., Hung, J., and Petch-Tyson, S. *Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching*. Philadelphia, PA: John Benjamins Publishing Co. 3-33.
- Grosz, B., Joshi, A., and Weinstein, S. (1995). "Centering: A Framework for Modeling the Local Coherence of Discourse". *Computational Linguistics*. Vol. 21, No. 2. 203-226.
- Guerts, B. and Beaver, D. (2007). "Discourse Representation Theory". *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/discourse-representation-theory/> Accessed October 22, 2009.
- Gundel, J. (1998). "Centering Theory and the Givenness Hierarchy: Towards a Synthesis". Ed. by Walker, M., Joshi, A., and Prince, E. *Centering Theory in Discourse*. New York, NY: Oxford University Press. 183-198.
- Gundel, J., Hedberg, N., and Zacharski, R. (2001). "Definite descriptions and cognitive status in English: why accommodation is unnecessary". *English Language and Linguistics*. Vol. 52. 273-295.
- Gundel, J., Hegarty, M., and Borthen, K. (2003). "Cognitive Status, Information Structure, and Pronominal Reference to Clausally Introduced Entities". *Journal of Logic, Language and Information*. Vol. 12. 281-299.
- Haegeman, L. and Gueron, J. (1999). *English Grammar: A Generative Perspective*. Malden, MA: Blackwell Publishers.
- Halle, M. and Marantz, A. (1994). "Some Key Features of Distributed Morphology". *MIT Working Papers in Linguistics*. Vol. 21. 275-288.
- Hammond, M. (2003). *Programming for Linguists: Perl for Language Researchers*. Malden, MA: Blackwell.

- Hatch, E. and Lazarson, A. (1991). *The Research Manual: Design and Statistics for Applied Linguistics*. New York, NY: Newbury House Publishers.
- Hawkins, J.A. (1978). *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. London, UK: Croom Helm.
- Hawkins, R. (2000). "Persistent Selective Fossilization in Second Language Acquisition and the Optimal Design of the Language Faculty". *Essex Research Reports in Linguistics*. Vol. 34. 75-90.
- Hawkins, R., and Chan, Y. (1997). "The Partial Availability of Universal Grammar in Second Language Acquisition: The 'Failed Functional Features Hypothesis'". *Second Language Research*. Vol. 13, No. 3. 187-226.
- Hawkins, R., and Liszka, S. (2003). "Locating the Source of Defective Past Tense Marking in Advanced L2 English Speakers". Ed. by Van Hout, R., Aafke, H., Kuiken, F., and Towell, R. *The Interface between Syntax and Lexicon in Second Language Acquisition*. Philadelphia, PA: John Benjamins Publishing Co. 21-44.
- Hays, W.L. (1994). *Statistics (Fifth Edition)*. Belmont, CA: Thomson/Wadsworth.
- Hegarty, M., Gundel, J., and Borthen, K. (2001). "Information Structure and the Accessibility of Clausally Introduced Referents". *Theoretical Linguistics*. Vol. 27, No. 2-3. 163-186.
- Heim, I. (1983). "File Change Semantics and the Familiarity Theory of Definiteness". Ed. by Bauerle, R. et al. *Meaning, Use, and Interpretation of Language*. Berlin, Germany: Mouton de Gruyter. 164-189.
- Hendriks, Henriette. (2003). "Using nouns for reference maintenance: A seeming contradiction in L2 discourse". Ed. by Ramat, A. Giacalone. *Typology and Second Language Acquisition*. New York, NY: Mouton de Gruyter. 291-326.
- Hess, M. (1987). "Descriptive Anaphora in Discourse Representational Theory". <http://www ldc.upenn.edu/acl/E/E87/E87-1025.pdf> Zurich, Switzerland: Department of Computer Sciences, University of Zurich. 148-155.
- Heycock, C. and Zamparelli, R. (2005). "Friends and Colleagues: Plurality, Coordination, and the Structure of DP". *Natural Language Semantics*. Vol. 13. 201-270.
- Hirschman, L. (1997). "MUC-7 Coreference Task Definition: Version 3.0". http://www.itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/co_task.html Accessed April 13, 2006.
- Horowitz, R. (2007). "Creating Discourse and Mind: How Talk, Text, and Meaning Evolve". Ed. by Horowitz, R. *Talking Texts: How Speech and Writing Interact in School Learning*. Mahwah, NJ: Lawrence Erlbaum Associates. 3-54.
- Horowitz, R. (2007). *Talking Texts: How Speech and Writing Interact in School Learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Hudson-D'Zmura, S. (1990). *The Structure of Discourse and Anaphor Resolution: The Discourse Center and the Roles of Nouns and Pronouns*. Dissertation: Department of Linguistics, Stanford University.
- Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karona.
- Ionin, T. (2003). "The Interpretation of 'the': A New Look at Articles in L2 English". *Proceedings of the 27th Annual Boston University Conference on Language Development*. Vol. 1. 346-357.
- Juffs, A. (2007). "Second Language Acquisition of Relative Clauses in the Languages of East Asia". *Studies in Second Language Acquisition*. Vol. 29. 361-365.
- Jurafsky, D. and Martin, J. (2000). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Upper Saddle River, NJ: Prentice Hall.
- Kamp, H. (2001). "The Importance of Presupposition". Ed. by Rohrer, C., Rossdeutscher, A., and Kamp, H. *Linguistic Form and its Computation*. Stanford, CA: Center for the Study of Language and Information. 207-254.
- Kamp, H. and Partee, B.H. (2004). *Context-Dependence in the Analysis of Linguistic Meaning*. New York, NY: Elsevier.
- Kamp, H. and Reyle, U. (1993). *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory, Parts 1 and 2*. Boston, MA: Kluwer Academic Publishers.
- Kehler, A. (2004). "Discourse topics, sentence topics, and coherence". *Theoretical Linguistics*. Vol. 30. 227-240.
- Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. Stanford, CA: Center for the Study of Language and Information.
- Kempson, R. (1996). "Semantics, Pragmatics, and Natural-Language Interpretation". Ed. by Lappin, S. *The Handbook of Contemporary Semantic Theory*. Cambridge, MA: Blackwell Publishers. 561-598.
- Kirk, R.E. (1999). *Statistics: An Introduction (Fourth Edition)*. Belmont, CA: Thomson/Wadsworth.
- Klein, D. (2005). "Lecture 8: CS 294-5: Statistical Natural Language Processing, POS Tagging II". <http://www.cs.berkeley.edu/~klein/cs294-5/index.html> Berkeley, CA: Natural Language Processing, University of California at Berkeley.
- Knott, A. (2004). "Introduction to Discourse Representation Theory". <http://www.cs.otago.ac.nz/staffpriv/alik/> University of Otago, New Zealand: Department of Computer Science.
- Kohlhase, M. (2000). "Model Generation for Discourse Representation Theory". <http://kwarc.info/kohlhase/papers/ecai00.pdf> Saarbrücken, Germany: Department of Computer Science, Universität des Saarlandes. 1-5.

- Kupisch, T. (2003). "Cross-Linguistic Influence in the Acquisition of Determiners in German-Italian Bilinguals". Ed. by Beachley, B. et al. *Proceedings of the 27th Annual Boston University Conference on Language Development*. Vol. 2. Somerville, MA: Cascadilla Press. 461-472.
- Lardiere, D. (1998). "Case and Tense in the 'Fossilized' Steady State". *Second Language Research*. Vol. 14, No. 1. 1-26.
- Lardiere, D. (2000). "Mapping Features to Forms in Second Language Acquisition". Ed by Archibald, J. *Second Language Acquisition and Linguistic Theory*. Malden, MA: Blackwell Publishers. 102-129.
- Lardiere, D. (2004). "Knowledge of Definiteness despite Variable Article Omission in Second Language Acquisition". Ed. by Brugos, A., Micciulla, L., and Smith, C.E. *Proceedings of the 28th Annual Boston University Conference on Language Development*. Vol. 1. Somerville, MA: Cascadilla Press. 328-339.
- Larocca Neto, J., Santos, A.D., Kaestner, C., Freitas, A., and Nievola, J. (2000). "A trainable algorithm for summarizing news stories". <http://www.ppgia.pucpr.br/~alex> Curitiba, Brazil: PUC-PR. 1-12.
- Lascarides, A. and Asher, N. (2003). "Segmented Discourse Representation Theory: Dynamic Semantics with Discourse Structure". <http://homepages.inf.ed.ac.uk/alex/pubs/iwcs4.html> Edinburgh, UK: Division of Informatics, University of Edinburgh. Austin, Texas: Department of Philosophy, University of Texas at Austin. 1-31.
- Levelt, W. (1989). *Speaking*. Cambridge, MA: MIT Press.
- Li, Wendan. (2004). "Topic Chains in Chinese Discourse". *Discourse Processes*. Vol. 37, No.1. 25-45.
- Liu, H. (2004). "MontyLingua: An end-to-end natural language processor with common sense". <http://web.media.mit.edu/~hugo/montylingua> Cambridge, MA: MIT Media Lab.
- Lleo, C., and Demuth, K. (1999). "Prosodic Constraints on the Emergence of Grammatical Morphemes: Crosslinguistic Evidence from Germanic and Romance Languages". Ed. by Greenhill, A. et al. *Proceedings of the 23rd Annual Boston University Conference on Language Development*. Vol. 2. Somerville, MA: Cascadilla Press. 407-418.
- McCallum, A. and Wellner, B. (2003). "Toward Conditional Models of Identity Uncertainty with Application to Proper Noun Coreference". *IJCAI Workshop on Information Integration and the Web*. <http://www.cs.brandeis.edu/~wellner/pubs.html> Massachusetts: Department of Computer Science, University of Massachusetts, Amherst and the MITRE Corporation. 1-6.
- MIT Speech Communication Group. (1999). *The American English Map Task*. <http://dspace.mit.edu/handle/1721.1/18239/browse?type=title>
- Napoli, D.J. (1993). *Syntax: Theory and Problems*. New York, NY: Oxford University Press.

- Nicol, J. and Swinney, D. (1989). "The Role of Structure in Coreference Assignment During Sentence Comprehension". *Journal of Psycholinguistic Research*. Vol. 18, No. 1. 5-19.
- Nistov, Ingrid. (2001). "Reference continuation in L2 narratives of Turkish adolescents in Norway". Ed. by Verhoeven, L. and Stromqvist, S. *Narrative Development in a Multilingual Context*. Philadelphia, PA: John Benjamins Publishing Co. 51-85.
- Oakes, M. (1998). *Statistics for Corpus Linguistics*. Edinburgh, UK: Edinburgh University Press.
- Perez-Leroux, A., and Glass, W. (1999). "Null Anaphora in Spanish Second Language Acquisition: Probabilistic versus Generative Approaches". *Second Language Research*. Vol. 15, No. 2. 220-249.
- Pinkal, M. and Thater, S. (2007). "Semantic Theory: Discourse Representation Theory I". http://www.coli.uni-saarland.de/courses/semantics-08/lectures/lecture_7.pdf
- Poesio, M. (1999). "Coreference". http://www.ims.uni-stuttgart.de/projekte/mate/mdag/cr/cr_1.html Accessed April 1, 2006.
- Poesio, M. (1999). "Part of Speech Tagging". <http://cswww.essex.ac.uk/poesio/papers.html> Presentation made October 1, 2003. Venice, IT: Universita' di Venezia.
- Poesio, M. (1999). "Visual situation uses of definite descriptions and focus shifts: A summary of Massimo's interests". Ed. by Andre, E., Poesio, M., and Rieser, H. *Proceedings of the Workshop on Deixis, Demonstration and Deictic Belief at ESSLLI XI*. HCRC, University of Edinburgh. 6-8.
- Poesio, M., Vieira, R., and Teufel, S. (1997). "Resolving Bridging References in Unrestricted Text". http://www.cl.cam.ac.uk/~sht25/papers/poesio_vieira_teufel.pdf Edinburgh, UK: Center for Cognitive Science, University of Edinburgh. 1-6.
- Pretorius, E. (2005). "English as a second language learner differences in anaphoric resolution: Reading to learn in the academic context". *Applied Psycholinguistics*. Vol. 26. 521-539.
- Rabiner, L.R. (1989). "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition". *Proceedings of the IEEE*. Vol. 77, No.2. 257-286.
- Reyle, U. and Rossdeutscher, A. (2001). "Temporal Underspecification in Discourse". Ed. by Rohrer, C., Rossdeutscher, A., and Kamp, H. *Linguistic Form and its Computation*. Stanford, CA: Center for the Study of Language and Information. 255-284.
- Robertson, D. (2000). "Variability in the use of the English article system by Chinese learners of English". *Second Language Research*. Vol. 16, No. 2. 135-172.
- Rohrer, C., Rossdeutscher, A., and Kamp, H. (2001). *Linguistic Form and its Computation*. Stanford, CA: Center for the Study of Language and Information.

- Rossen-Knill, D., Spejewski, B., Hockey, B.A., Isard, S., and Stone, M. (1997). "Yes/No Questions and Answers in the Map Task Corpus". *University of Pennsylvania Institute for Research in Cognitive Science Technical Report No. IRCS-97-11*. <http://repository.upenn.edu/ircs-reports/83> 1-19.
- Sanchez, L., and Jose Gimenez, M. (1998). "The L2 Acquisition of Definite Determiners: from Null to Overt". Ed. by Greenhill, A. et al. *Proceedings of the 22nd Annual Boston University Conference on Language Development*. Vol. 2. Somerville, MA: Cascadilla Press. 640-650.
- Sanders, T.J.M., Gernsbacher, M.A. (2004). "Accessibility in Text and Discourse Processing". *Discourse Processes*. Vol. 37, No. 2. 79-89.
- Santorini, B. (1990). "Part-of-Speech Tagging Guidelines for the Penn Treebank Project (3rd Revision, 2nd Printing)". <http://www.cis.upenn.edu/~treebank>
- Saunders, J. K. (1999). *Null and Overt References in Spanish Second Language Acquisition: A Discourse Perspective*. Dissertation: University of Texas, Austin.
- Schmitt, B., Meyer, A., and Levelt, W. (1999). "Lexical access in the production of pronouns". *Cognition*. Vol. 69. 313-335.
- Schnedecker, C. (2003). "The Issue of Proper Noun Repetition in Centering Theory and Its Problems". *Journal of French Language Studies*. Vol. 13, No. 1. 105-134.
- Schnell de Acedo, B. (1994). "Early Morphological Development: The Acquisition of Articles in Spanish". Ed. by Sokolov, J. and Snow, C. *Handbook of Research in Language Development*. Hillsdale, NJ: Lawrence Erlbaum Associates. 210-52.
- Smith, C.S. (2003). *Modes of Discourse: The Local Structure of Texts*. New York, NY: Cambridge University Press.
- Solt, S. et al. (2004). "L2 Perception and Production of the English Regular Past: Evidence of Phonological Effects". Ed by Brugos, A., Micciulla, L., and Smith, C.E. *Proceedings of the 28th Annual Boston University Conference on Language Development*. Vol. 2. Somerville, MA: Cascadilla Press. 553-564.
- Song, H. and Fisher, C. (2005). "Who's 'she'? Discourse prominence influences preschoolers' comprehension of pronouns". *Journal of Memory and Language*. Vol. 52. 29-57.
- Spalek, K. and Schriefers, H. (2004). "Dominance affects determiner selection in language production". *Journal of Memory and Language*. Vol. 52. 103-119.
- Sternfeld, W. (2001). "Semantic vs. Syntactic Reconstruction". Ed. by Rohrer, C., Rossdeutscher, A., and Kamp, H. *Linguistic Form and its Computation*. Stanford, CA: Center for the Study of Language and Information. 145-182.
- Stromqvist, S. and Day, D. (1993). "On the development of narrative structure in child L1 and adult L2 acquisition". *Applied Psycholinguistics*. Vol. 14. 135-158.
- Swales, J. and Feak, C. (2004). *Academic Writing for Graduate Students: Essential Tasks and Skills, Second Edition*. Ann Arbor, MI: University of Michigan Press.

- Tantos, A. (2005). "Segmented Discourse Representation Theory: A Theory of Discourse Interpretation". *Computerlinguistik II*. <http://ling.uni-konstanz.de/pages/home/butt/teaching/rhet-relations-cl2.pdf> Konstanz, Germany: Universitaet Konstanz.
- Ter Meulen, A. and Abraham, W. (2004). *The Composition of Meaning: From lexeme to discourse*. Philadelphia, PA: John Benjamins Publishing Co.
- Thompson, H., Anderson, A., Gurman Bard, E., Doherty-Sneddon, G., Newlands, A., and Sotillo, C. (1993). "The HCRC Map Task Corpus: Natural Dialogue for Speech Recognition". <http://www.aclweb.org/anthology/H/H93/H93-1005.pdf> University of Edinburgh. 25-30.
- Traum, D., Schubert, L., Poesio, M., Martin, N., Light, M., Hwang, C.H., Heeman, P., Ferguson, G., and Allen, J.F. (1996). "Knowledge Representation in the TRAINS-93 Conversation System". *Technical Report 633 and TRAINS Technical Note 96-4*. Rochester, NY: University of Rochester Computer Science Department. 1-52.
- Van Eijck, J. (2005). "Discourse Representation Theory". *Amsterdam, The Netherlands: CWI*. Utrecht, The Netherlands: OTS.
- Von Stutterheim, C.; Mangold-Allwinn, R.; Barattelli, S.; Kohlmann, U.; Kolbing, H-G. (1993). "Reference to Objects in Text Production". *Belgian Journal of Linguistics*. Vol. 8. 99-125.
- Walker, M., Joshi, A., and Prince, E. (1998). *Centering Theory in Discourse*. New York, NY: Oxford University Press.
- White, L. (2002). "Morphological Variability in Endstate L2 Grammars: The Question of L1 Influence". Ed. by Skarabela, B., Fish, S., and Do, A. *Proceedings of the 26th Annual Boston University Conference on Language Development*. Vol. 2. Somerville, MA: Cascadilla Press. 758-768.
- White, L. (2003). "Fossilization in Steady State L2 Grammars: Persistent Problems with Inflectional Morphology". *Bilingualism: Language and Cognition*. Vol. 6, No. 2. 129-141.
- White, L. (2003). *Second Language Acquisition and Universal Grammar*. New York, NY: Cambridge University Press.
- Wittek, A. and Tomasello, M. (2005). "Young children's sensitivity to listener knowledge and perceptual context in choosing referring expressions". *Applied Psycholinguistics*. Vol. 26. 541-558.