Training the Perception and Production of English Vowels /ɪ/-/iː/, /e/-/æ / and /ʊ/-/uː/ by Cantonese ESL Learners in Hong Kong

WONG, Wing Sze

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in
English (Applied English Linguistics)

The Chinese University of Hong Kong

September 2013

Thesis/Assessment Committee

Professor Danny Chung Hong LEUNG (Chair)

Professor Jette Gjaldbaek HANSEN EDWARDS (Thesis Supervisor)

Professor Gwendolyn GONG (Committee Member)

Dr. Paola ESCUDERO (External Examiner)

ACKNOWLEDGEMENTS

My first and foremost gratitude goes to my Heavenly Father who makes me who I am today. Thank God for granting me the love, wisdom and perseverance I need so that I can get through all the stages. I thank God also for letting me know the following fine individuals. It is not the first time for me to say that this degree belongs to them instead of me. Words cannot convey and extend fully my thankfulness to them; I only wish I can do more in the future to pay back for their engagement, support and guidance throughout the past years. While they should take the credits of the strength of this dissertation, all faults and errors remain mine.

I am deeply indebted to my supervisor, Prof. Jette G. Hansen Edwards. The eight-year period that we have gone through will certainly be an indelible memory in my lifetime. Thank you for being a great scholar, advisor and mentor who guides me through every milestone in my academic pursuit. Your wisdom and insights in research and teaching have always enlightened me. Thank you for recognizing and trusting me and I look forward to our collaborations in the near future.

I am also very grateful for having two great thesis committee members. Prof. Gwendolyn Gong is a teacher and mother with a caring heart. I am truly blessed to have met you during my university life. Thank you for always giving professional advice on editing and writing. Your meticulous comments were always of enormous help. Your encouragements have been even more important as you have let me know how capable I am and I can always step forward for something more challenging. Prof. Danny Leung is always my problem-solver and listener. Thank you for giving me so much useful advice and support. Bombarding you with trivial questions from time to time makes me feel abashed, but your constructive answers and humour always give me a relief. I am so blessed to have a friend like you.

I particularly thank Dr. Paola Escudero, my external examiner, for being so willing to assess my work. I apologize for having you read the work when you were in such a tight schedule and condition, but I am grateful that you were still extremely helpful and encouraging in offering a huge amount of useful comments and insights. Your knowledge and enthusiasm in speech research have always been one of my best examples which I admire wholeheartedly and aspire to accomplish. I sincerely hope that I can learn more from you in the future. Thank you for recognizing me and giving me so much encouragement!

My deepest heartfelt appreciation goes to my sister's secondary school teacher, Mr. Andy Wu who served as a volunteer computer engineer in this research. Your contribution is of immense importance to my study. The computer program that you wrote is marvellous and I am sure this will be useful to a lot of ESL learners. Thank you for sparing so much time and effort in helping someone who cannot return much to you.

Never had I imagined that I would be able to consult three doctors during my data analysis stage. Thank you, Dr. Patrick Chu, Dr. Jan Wang and Dr. Gavin Bei, for your professional advice on SPSS and statistical computing. Without your advice, I would not have solved the problems I encountered and finished the analysis within a short period of time. Thank you for being so willing to answer all the questions I raised, Patrick. Thank you for spending a lot of time on the analysis even when you were working hard on your newly-started business. I promise I will be one of the investors! Thank you, Jan, for your super long email replies despite that you were busily taking care of your new-born baby. I just can't imagine that you still replied me when you were just one week away from the delivery! Your kind and helpful heart is certainly a great blessing to your son. Gavin, you have always been a good elder brother of mine and I do not just like to hear your advice but I also treasure our friendship. Thank you for caring so much about my publications as well as what I wish to embark on after my graduation.

I must acknowledge my secondary school's principal, Mrs. Choi Lam Sui Mei, for giving me the permission to conduct the experiment and to use all the equipment to complete the research. I also thank Mr. Jerry Chui, Mr. Howard Chau, Miss Katherine Fok and the two technicians in the laboratory, Mr. Lai and Mr. Hung, for all the great help and support during my data collection period.

You may not expect that you are mentioned here, Mr. Peter Wong. You are the one that I owe much; I am sure you were sent by God to teach me English some ten years ago. Without you, I would never be able to come up to this stage. You certainly share a large portion of credits!

All student participants from CDGFSS who were willing to be my subjects in the experiment were extremely important to me. I would like to express the deepest appreciation to all of you for sparing so much of your time to participate in the lengthy experiment. I sincerely hope that our friendship can last long (despite our age difference!) and that you all can still keep the kind, pure and helpful hearts in the coming days. I particularly thank the four little "laboratory assistants," Cheuk Ting Chan, Katherine Chan, Tim Sheung and Tony Lee for their extra help in the coordination. I am sure I will see you all shine wherever you go.

Maggie Cheng, my good studying partner since my undergraduate studies, has offered continuous support not only to my thesis editing part, but also to my future career path. Thank you for being a good listener and a supportive friend ever since we have known. Thank you for sparing your precious time to offer solid help in the "space mission" as well!

I am grateful for having Ruby Wong, Jessica Hui and Kit Chio for collecting a large number of audio samples for my training programme. I am particularly thankful for those native speakers, particularly my sister's secondary school teacher, Mr. Gilbert Wong, for taking up all the hard work of producing the stimuli and verifying their intelligibility. In addition, I would like to thank my CUHK fellow friends and students, Dennis Chau, Arlynn Alarcon and Kit Chio for serving as the inter-raters in the phonetic transcriptions. Thanks are also due to Tracy Liang and Carly Chan, the administrative staff of Department of English at CUHK, for their excellent administrative support.

I owe a great debt to my “research assistants” Agathos Lam, Naiym Kwong, Priscilla Siu and Zoe Chau who solved many technical problems during my data analysis stage, and the always-available volunteer, Faith Cheuk. Thank you so much for all of your kind hearts, generous help and solid support. All of you have been so helpful that I am always moved. I also thank my student, Arlynn Alarcon, for the good advice during the editing stage; and my editor, Derrick Leung, for sparing much time looking over my work.

Special gratitude goes to my cell group leader, Simeon Tang – a father of kind heart, an instructor with enlightening wisdom and a prophet showing promising directions; and my spiritual mother, Tehillah Chan – a potter with unfailing love and perseverance, a coach with life-changing power and a tour guide in my life journey. The stimulating guidance and unswerving support from you both have moulded me who I am today. My deepest and whole-hearted gratitude can only be realized through taking actions and fulfilling promises I have been making.

My seven “warring pearl” friends and Team D members should not be forgotten. Some of you may not understand what I have been doing, but your kind words, sincere understanding or helping of any kind have always been my great comfort and support which have taken me to the completion of this dissertation. Your company and prayers are also highly appreciated. Particular thanks also go to Dawn Tong and Kaphar Yeung for the professional advice in the printing stage.

I am thankful to my family as well. Thank God for giving me the greatest and most caring parents in the world. You have given me all the support and love without which I cannot grow up to this stage. Even though I cannot become the *doctor* (the medical one) you wish, I know being another type of *doctor* is still a heartening status that you feel proud of always. Even though we are not well-off, the love and security that I can get both physically and spiritually are so abundant that I know I am truly blessed. I am who I am now because of the love and education you have been providing. I owe you my whole life. Special thanks go to my lovely sister, Ruby. Without you, I would not have such a joyful and meaningful life. Thank you so much for your continuous help, love and support of any kind and in any form. You seldom say no to my requests; I know you are helping simply because you love me. I love you too.

The Chinese University of Hong Kong and the Research Grants Council have also made this research possible through the “Hong Kong PhD Fellowship Scheme 2010.” This dissertation would not have been completed without the support of this three-year award.

May the Lord’s name be praised and honoured in my life always! *Soli Deo Gloria*.

ABSTRACT

This current study investigates and compares the effectiveness of High Variability Phonetic Training (HVPT), Low Variability Phonetic Training (LVPT), production training, and a combination of either HVPT or LVPT plus production training on improving both the perception and production of three English vowel pairs, /ɪ/-/i:/, /e/-/æ/ and /ʊ/-/u:/, by Cantonese-speaking ESL learners. HVPT is a perception training paradigm which provides the subjects with training stimuli under various phonetic environments produced by various speakers, whereas LVPT offers stimuli produced by just a single speaker. With a 3-phased "pretest-treatment-posttest" structure, the current study recruited 85 participants and employed an experimental design in which 9 groups of the target participants took part in different training sessions in over a period of time, with also a control group. Perceptual training intensity (intensive – 10 training sessions per day; or standard – 2 training sessions per day) was also one factor under investigation.

Analyses showed that the subjects generally had difficulties in both the perception and production of the three vowel pairs before training. While the production training was found to have only mild benefits to the perception learning, both HVPT and LVPT were effective in improving the subjects' perception of the three vowel pairs. Groups receiving both HVPT/LVPT and the production training improved as well, but their improvement was not as great as the groups receiving only the perceptual training. Perceptual learning could also be generalized to new words and new speakers, but only the HVPT group demonstrated significantly better performance than other groups. Perceptual training intensity did not have significant influence on the subjects' performance.

As for the learning in the production domain, all subjects with training improved significantly after the intervention. The HVPT group outperformed the LVPT and the control groups with significance whereas the LVPT group did not display robust difference with the control group. The production training was also beneficial to the subjects' improvement of the production accuracy. Even different training groups improved significantly in the posttest; however, only the HVPT with production training group outscored the other groups with significance. Again, perceptual training intensity was not an influential factor. The formant frequencies and durations of the vowel pairs produced by the trained groups after the training became not as conflated as before and were closer to native productions. However, although the number of target productions in the contextualization test of production at the sentence-level was high, only the HVPT group with production training outperformed the other groups with significance.

The above findings show that training in perception alone can be sufficient for significant improvement in both the perception and production domains, with the HVPT being a more effective paradigm. Production training can facilitate more learning in production than in perception. Future studies can investigate the acoustic cues that these learners rely on, as well as how different training methods can benefit L2 learners.

Keywords: speech perception and production, High Variability Phonetic Training (HVPT), L2 vowel perception, non-native phonemic contrast, production training, perceptual training intensity, Cantonese ESL learners

摘要

本研究探討並比較──「高度變異語音訓練法」、「低度變異語音訓練法」、發音訓練，以及當「高度變異語音訓練法」或「低度變異語音訓練法」結合發音訓練時──這些方法於改善以廣東話為第一語言的英語學習者對三組英語元音/⋂/-/i:/、/ɛ/-/æ/ 和 /⋃/-/u:/ 之感知與發音方面的效益。「高度變異語音訓練法」是一感知訓練方法，讓研究對象透過聽取高變異的感知訓練刺激物（刺激物分別由 6 位以英語為母語人士錄製），從而提高他們對相似元音的感知。至於「低度變異語音訓練法」的感知訓練刺激物則由 1 位英語為母語人士錄製。本研究共招募 85 名參加者，採納「前測—訓練—後測」步驟。按照實驗設計，9 組研究對象在一段時間裡參加了不同訓練課節;研究同時設有對照實驗者。感知訓練的密度（頻密──每日 10 訓練課節；一般──每日 2 訓練課節）亦是本研究其中一個探討因素。

研究分析顯示，所有研究對象於接受訓練前，對於該三組元音在感知方面和發音方面都遇上困難。就不同訓練方法的成效而言，研究發現發音訓練只對感知學習帶來輕微幫助。至於「高度變異語音訓練法」和「低度變異語音訓練法」兩者皆能有效地改善研究對象對該三組元音的感知。接受結合「高度變異語音訓練法」或「低度變異語音訓練法」結合發音訓練的研究對象同樣有改善，但改善程度並不及只感受感知訓練的組別。感知學習也同時能夠類推至新刺激物及新發音者，但只有接受「高度變異語音訓練法」的組別對比其他組別呈現明顯較佳的表現。然而，感知訓練密度並沒有對研究對象帶來顯著影響。

在發音範疇學習方面，所有研究對象於接受訓練後均有明顯改善。接受「高度變異語音訓練法」組別的改善程度比接受「低度變異語音訓練法」的組別和對照實驗者組別為高。然而，接受「低度變異語音訓練法」的組別與對照實驗者組別相比，卻沒有明顯變化。發音訓練亦對研究對象於改善發音準確度有所幫助。雖然各訓練組別於後測時有明顯的改善，只有結合「高度變異語音訓練法」和發音訓練的組別比其他組別優勝。同樣，感知訓練密度並沒有對研究對象的表現帶來明顯影響。接受過訓練的各組別，在讀出該三組元音時，共振峰頻率和元音延續時間變得接近於以英語為母語者的發音。然而，雖然研究對象在雖然在泛化發音測試（文章朗讀）中，發音準確度已改善，但亦只有接受「高度變異語音訓練法」的組別比其他組別優勝。

實驗結果顯示，單是感知訓練本身已足以對感知和發音兩個範疇帶來明顯改善，並以「高度變異語音訓練法」為較有效益的訓練方式。而發音訓練能促進發音學習過於感知學習。日後研究可探討學習者所用的感知線索，以及不同訓練方法如何能使第二語言學習者獲益等。

*關鍵字：*語音感知及發音、高度變異語音訓練法、第二語言之元音感知、非第一語言之音素對比、發音訓練、感知訓練密度、以廣東話為第一語言的英語學習者

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.0 INTRODUCTION

The inability for L2 learners to ultimately develop an L2 phonological system that resembles that of the native speakers is well-known. Despite how advanced, proficient or linguistically experienced an adult second language (L2) learner is in the target language – be it syntax or vocabulary, etc. – he/she is usually characterized by accented L2 phonological patterns (Flege, Mackay, & Meador, 1999; Munro, Flege, & Mackay, 1996; Wang, 2002). In particular, many L2 learners have immense difficulties in the perception and/or production of some non-native phonemic contrasts. The barrier stems from the efforts in re-tuning and shifting from the prosodic monolingual system of an L1 to a two-way interactive phonological system of the L1 and L2, since the acquisition processes of adult L2 learners have become less plastic; additionally, the acoustic patterns of the categories in the first language (L1) dominate so the L2 learners' phonetic systems are said to have been "fossilized" (Best, 1995; Bohn, 1998; Bradlow, 2008; Flege, 1989; Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann & Siebert, 2003; MacKain, Best & Strange, 1981; Major, 1987; Polka, 1992; Strange, 1995).

Early on, studies attributed learners' L2 learning difficulties to the differences between the L1 and L2 phonological systems (e.g. Eckman, 1977; Lado, 1957), which however, cannot account for all the complications and results observed in a number of empirical studies (Bohn, 1995; Nemser, 1971). More recent works (e.g. Best, McRoberts, & Sithole, 1988; Best & Tyler, 2007; Bradlow, 1997; Lambercher, Martens, Kakehi, Marasinghe & Molholt, 2005; MacLeod, Stoel-Gammon, & Wassink, 2009) related L2 learners' difficulties to the assimilation of L2 segments to their existing L1 categories, stating that it is the language-specific nature of perception among adult L2 learners that underlies the acquisition difficulty. One prominent example is the learning of English /ɹ/ and /l/ among Japanese learners of English, who are notorious for confusing the two sounds in terms of perception and production. Difficult as it may seem, perceiving non-native contrasts among adult learners has been found possible under some training conditions.

Various phonetic training paradigms (e.g. Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999; Iverson & Evans, 2009; Logan, Lively & Pisoni, 1991; Logan & Pruitt, 1995; Pisoni & Lively, 1995) have been devised by researchers to examine the effects of phonetic training on the modification of perception and/or production of non-native contrasts by subjects with different language backgrounds. These studies have given evidence that, through using auditory training approaches under laboratory conditions, there exists the possibility of the re-matching of the phonological system of the L2 learners from a monolingual to a bilingual system. One paradigm called High Variability Phonetic Training (HVPT) approach, which advocates the use of natural and highly variable stimuli with various phonetic contexts produced by multiple speakers, was found to be an effective approach in improving L2 learners' perception and production of non-native sounds, such as the production of non-native contrast in consonants or voice onset time duration (e.g. Lively, Logan & Pisoni, 1993; Lively, Pisoni, Yamada, Tokhura & Yamada, 1994; Logan et al., 1991; Rochet, 1995) or vowels that pose difficulties for the L2 learners (e.g. Lambacher et al., 2005; Wang, 2002; Wong, 2010).

Hong Kong Cantonese learners with English as an L2 (ESL) also have difficulties in the acquisition of some vowels that appear only in English (Chan, 2012; Chan & Li, 2000; Sewell, 2009). These sounds are often assimilated to some similar Cantonese counterparts (e.g. Cantonese [œ] for English /ɜː/; Cantonese [e] for English /æ/; Cantonese [i] for English /iː/ and /ɪ/). In this dissertation, I seek to expand and enrich the current understanding of L2 speech learning by addressing the practicality of different vowel training approaches through investigating the perception and production performance of Hong Kong Cantonese post-puberty teenagers. I also aim to explore ways to a more successful perceptual and productive learning of three non-native vowel contrasts, namely /ɪ/-/iː/ (as in *ship* vs. *sheep*), /e/-/æ/ (as in *bed* vs. *bad*) and /ʊ/-/uː/ (as in *soot* vs. *suit*), which are found to have posed difficulties among those L2 learners. I attempt to extend and modify previous experimental methodologies by comparing the efficacy and effects of different training approaches on the vowel learning, and link the results of the present investigation with the conceptual and theoretical frameworks.

Before proceeding to a more thorough exposition of relevant literature, I will first delineate the history of the development of L2 acquisition in phonology in the next section. This foregrounds how the study of acquisition of the L2 sound system has been viewed by researchers and how the relevant theories help shape our knowledge of the factors influencing L2 phonological acquisition. Following that, an overview of the whole dissertation will be given.

1.1 An Outline of the Development of L2 Phonological Acquisition

The Critical Period Hypothesis (CPH) hypothesizes why many L2 learners encounter continuous frustrations and failure in the acquisition process of the L2. Initially proposed by neurologists Penfield and Roberts in 1959 and later popularized by Lenneberg in 1967 to explain L1 acquisition, the CPH states that the maturation and lateralization[1] of the brain results in fossilization in acquiring a language since the ideal and crucial “window” of time for language acquisition has nearly ended (Lenneberg, 1967). This ideal period of time for language acquisition is called the critical period, “a period during which language can be acquired more easily than at any other time” (Richards, Platt, & Platt, 1992, p.92). Lenneberg theorized that it is the biological factors which exert constraints in late acquisition.

In the mid 1970s, this hypothesis was further extended to the investigation of L2 acquisition, especially in L2 phonological acquisition, an area that shows the most evident differences in the level of success and is the most related to the age factor (Pennington, 1998). As hypothesized in the CPH, due to the critical period, most adult L2 learners retain an obvious and noticeable accent despite having acquired a reasonable proficiency level of the language in general (Oyama, 1976). The CPH can

[1] Lateralization refers to the idea that the brain is divided into two halves, i.e. the left and the right hemispheres, and they execute different functions. Research (cf. Taylor & Taylor, 1990) has found that the left hemisphere is specialized in the language function.

itself be the appealing explanation as to why early learners are usually without a detectable foreign accent that most late and adult learners are not able to get rid of. Learners who can successfully acquire an L2 phonology with native or near-native-like performance will merely be deemed as exceptional cases or simply linguistic deviants.

The fact that no explanation can be given to answer these successful yet deviated cases has given rise to some postulations of a milder version of the CPH, which is not biologically based, but cognitively founded (Birdsong, 1992). Viewed as one kind of cognitive ability[2], language acquisition can be successful if the general cognitive flexibility of the learners can be maintained up to a certain degree. This ability is however unavailable to many learners when they have grown up and become cognitively mature.

The CPH received challenges and arguments from its opponents not just because successful L2 learners are strong counter-arguments to the hypothesized sensitive period (cf. Birdsong, 2007), but also because the CPH lacks a reasonable account and explanation of L2 learners' difficulties in acquiring the sound system (e.g. Flege, 1992; Long, 1990). Long (1990) pointed out that language interference or negative L1 transfer[3] is a more plausible reason for learners' problems, as the loss of plasticity in language acquisition will not be a sudden and one-time effect on the barrier of learning

---

[2] Cognitive abilities are underlying brain-based skills that we need when learning, remembering or thinking. They are mental processes that allow us to perform the simplest tasks to the most complex ones. Attention, memory, problem-solving, preference changing, decision making are some examples of cognitive abilities. Some researchers regarded language acquisition as a cognitive ability like learning mathematics or talking on phones.

[3] Language transfer has been considered as having a dominant influence in SLA, particularly in phonological acquisition. If the elements of the native language are similar to those in the L2, learners do not have to learn something new, i.e. elements in L1 can readily be transferred to the L2, and it is known as *positive transfer.* If learners apply their knowledge of their L1 to the L2 but they then produce deviated forms, it will be known as *negative transfer* or *interference*.

phonology. Previous research findings (e.g. Bialystok & Hakuta, 1994; Flege, Munro & MacKay, 1995; Munro & Mann, 2005; Patkoski, 1990) have supported that there is a continuous decline in the ability to learn a language with age, making the L2 phonological acquisition more difficult for adult L2 learners.

Flege and colleagues' (1995) research has provided strong evidence against the CPH, as 240 Italian learners of English who immigrated to Canada who are at a critical age boundary of around 15 years old were found to have no sudden loss of ability to produce more native-like speech; instead, the researchers discovered a linear relation between the subjects' age of learning and the degree of foreign accent. In addition, even if the subjects had an early onset of exposure to the L2, their speech was also characterized by varying degrees of perceived foreign accent. Flege, instead of totally denying or affirming the existence of a critical period, hypothesized a model called Speech Learning Model (SLM; Flege, 1991, 1995b, 1999, 2002, 2003) which provides a more comprehensive view of the factors that affect L2 speech learning, one of which is L1 transfer. Based on the premise that the perceptual categories of the learners' L1 will influence L2 perception and that the degree of L1-L2 similarity and use can predict learning difficulties, the SLM goes beyond the CPH and places L1 use or disuse alongside the age of learning the L2 to explain the ultimate success or failure in achieving native-like pronunciation competency. Details of the SLM will be discussed in Chapter 2.

Having the concept of language interference in mind, some linguists adopted a more systematic study of two languages in terms of their linguistic differences and similarities – Contrastive Analysis Hypothesis (CAH) – to explain the difficulties in mastering some structures in an L2. The CAH was most popular in the 1960s and early 1970s, influencing the focus of SLA research, especially L2 phonological research. It suggests that L2 learners' difficulties and errors are solely attributed to L1 transfer, meaning that the extent of successful L2 learning is highly dependent on the structural similarity of the L1 and L2 (Lado, 1957; Lee, 1968). At the time when the CAH was put forward to explain L2 learning phenomenon, the structuralist and behaviourist views of language learning were widely acknowledged, imposing important implications for the CAH: language learning is a set of habits that can be fostered through conditioning. L1 transfer is the interference of the old habits (learner's native language) with the new habit (the L2) formation. With transfer as a weight tenant, the CAH predicted that aspects of L2 would be easily acquired if the L1 was similar (positive transfer), whereas those L2 targets that were very different from the L1 would be learning difficulties (negative transfer). Yet, the CAH failed to account for and predict some errors found in some studies (cf. Whitman and Jackson, 1972) investigating various linguistic aspects. The increasing criticisms on the CAH in the late 1970s notwithstanding due to the lack of empirical evidence and the shift of attention from behaviorist to cognitive approaches, L1 transfer has still remained as an explanatory variable and played a dominant role in SLA nowadays, especially in the area of L2 phonology.

However, other theories and propositions, such as the notion of universal development, which concerns the concept of markedness in the study of phonology, also drew linguists' attention. The concept of markedness was firstly pioneered by

Trubetzkoy (1939) and Jakcobson (1941), two leading linguists in Prague School of Linguistics. The underlying idea of this construct is that one member of a binary opposition of a linguistic representation which has a wider distribution (either across languages or within the language) than its counterpart will be known as unmarked. These unmarked items are inherently easier to produce, more natural and common to be found, and are more basic and normal in human languages. One important formulation invoking the notion of markedness to explain L2 phonology is Eckman's (1977) Markedness Differential Hypothesis (MDH).

In the MDH, L1 transfer is not the sole factor affecting the acquisition process, but the features of the language itself and the universals can be attributive. The MDH, as a revision of the CAH, incorporated the central concept of typological markedness to fill the gap brought by the weaknesses of the CAH. While language learning difficulty can still be predicted by means of contrastive analysis of the L1 and L2, the MDH postulates further that sounds are difficult to learn only if they are typologically more marked than other sounds – this is the sufficient factor that explains L2 learning difficulty. This implies that not all differences found in the linguistic representations between the L1 and L2 will cause equal problems to a learner, but only the differences that have a relatively higher degree of markedness will be the learning obstacles. Early studies such as Carlisle (1991) provided evidence that learners' difficulty can be predicted and explained only by considering the markedness relationship in the differences between the L1 and L2. Recruiting a group of Spanish learners of English, Carlisle (1991) probed into the learners' English complex onset production performance and discovered that the subjects inserted an epenthetic vowel to break up complex consonant clusters which are not present in Spanish. As consonant clusters are relatively more marked structures, the learners adopted the strategy of inserting a vowel into the clusters to make them into two singleton consonants, e.g. from CC to CVC, so as to reduce the markedness of the structure. Other research such as Anderson (1987) on consonant clusters, Eckman (1981a, 1981b) on voiced coda obstruents among subjects with different native L1s, Major and Faudree (1996) on voicing contrast, Bhatia (1995) on aspiration, Cichocki, House and Lister (1997) on vowels also showed similar positive results supporting the MDH. However, not all evidence supported the hypothesis. A more recent study by Lo (2007) on a group of Hong Kong Cantonese speakers of English learning coda clusters found that some unmarked structures (e.g. /ts/ as a coda) were more difficult to acquire than the relatively marked ones (e.g. /fs/ as a coda). The MDH was not adequate in explicating the acquisition, since factors other than markedness such as native language phonotactics and visual salience of consonants also influenced the learning. Still, the MDH should be credited as it has improved beyond the simplistic approach of the CAH in determining L2 learners' difficulty.

Be it the CAH or the MDH, these hypotheses both highlight the role of transfer in L2 phonological acquisition. Some other scholars (Hacin-Bhatt & Bhatt, 1997; Major, 1998) also regarded transfer as a major construct, but they considered further the effect of interaction between language transfer and developmental factors on the influence of L2 phonology. No matter what foci the above theories and propositions adopted, they all attempted to find out the underlying reasons behind the difficulty in SLA. Yet, they were based on the results of learners' production solely, without focusing on the interaction between speech perception and production and their effects on L2 speech learning.

The relationship between speech perception and production is always intricate and has attracted attention in L2 phonological studies. Many researchers (e.g. Ferguson & Macken, 1980; Hewlett, 1990) have noted that speech production is a process (rather than a product) comprising different phases: perceptual, programming, processing and execution levels. Empirical studies of cross-linguistic perception and production in recent decades have brought about more research insights and understanding of adult L2 speech learning problems. Some of them focused on non-native consonant contrasts (e.g. Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999; Bradlow, Pisoni, Yamada & Tohkura, 1997; Lively, Logan & Pisoni, 1993; Lively, Pisoni, Yamada, Tokhura & Yamada, 1994; Logan, Lively & Pisoni, 1991; Yamada, 1999) or non-native place of articulation contrasts (e.g. Polka, 1991, 1992; Werker & Tees, 1983, 1984). Some were about L2 vowel perception (e.g. Escudero, 2005; Ingram & Park, 1997; Lambacher, Martens, Kakehi, Marasinghe & Molholt, 2005; Makarova, 2010; Pallier, Colomé & Sebastián-Gallés, 2001; Polka, 1995; Rochet, 1995). Relevant studies to the present investigation will be discussed in detail in Chapter 2.

From these studies, a number of factors are believed to be influential to L2 speech learning, for instance, assimilation of L2 segments to L1, age of learning, L2 experience, etc. The outset of the next chapter will attempt to delineate how some studies have shown that these factors may influence L2 speech learning among which a great body of them (e.g. Best & Tyler, 2007; Flege, 2003; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola, & Nelson, 2008) examined the nature of perceptual assimilation. With these studies and expositions, one can learn about the possible causes of L2 speech learning problems. They will lay the groundwork for the understanding of the causes of the difficulties in the perception and production of some English vowel contrasts by Cantonese ESL learners in the present study, which motivates the current training study aiming to improve their L2 speech problems.

1.2 Dissertation Overview

To end this introductory chapter, I will present an overview of the whole thesis. The thesis consists of six chapters and is structured as follows.

Chapter 1 gives an introduction of the history of L2 phonological acquisition and theories which state the complications and factors that influence the acquisition process at the outset of the thesis.

Chapter 2 is an overview of the current theories and empirical studies on L2 speech learning, particularly in speech perception and production. It starts with outlining the studies in both L1 and L2 speech perception, followed by discussing the factors that affect L2 speech learning and previous studies that investigated them. Results in cross-language studies using different phonetic training approaches in modifying the perception and/or production of non-native consonant and vowel contrasts are summarized. A short description and comparison of the vowel systems of English and Hong Kong Cantonese is followed to give background information to the present research target sounds. Relevant studies which investigated the L2 vowel system of some Cantonese ESL learners are also presented before closing the chapter by presenting a relevant study conducted by the researcher.

Based on the soundness of the theories and results in previous training studies, Chapter 3 presents the research methodology adopted in the current investigation. It includes the recruitment procedures, the training methods adopted, the structure of the experiment, the production process of the training stimuli and the test tokens. The data transcription method, reliability checking and procedures utilized in the acoustic measurement and native speaker’s judgment on the production data are also presented.

Chapter 4 reports on all the data obtained from the experiment. Besides giving straightforward scores and percentages of the subjects’ performance, different statistical analyses were also run to verify the validity and significance of the research findings. The analyses are displayed and elucidated with extensive illustration of tables and figures. A report on the acoustic analysis on the production data of the subjects is also given, with the support of various F1-F2 space plots.

Chapter 5 discusses all the research findings in all perception and production tests, in turn answering all the research questions set out in Chapter 2 one by one. Thorough discussions and justifications of the subjects’ performance are presented with support and reference to earlier studies and theories.

Chapter 6 concludes the whole thesis by giving an overview of the present study. Contributions and implications of the study are also pointed out. Limitations of the study are also discussed, followed by the suggestions of future research directions.

CHAPTER 2

LITERATURE REVIEW

2.0 INTRODUCTION

As the foundation of the major exposition of the literature that follows, this chapter will cover a review of the studies on speech perception and production as well as the theoretical accounts that are central to the understanding of the current research. It opens up with Section 2.1 outlining the well-documented factors that affect L2 speech learning and how similar or different L1 and L2 perception are by drawing examples from relevant empirical studies, such as infant speech perception, cross-linguistic perception and L2 speech perception which then lead to theories that foster more understanding underlying L2 phonological acquisition Some fundamental models explaining speech perception and production will then be introduced in Section 2.2 and 2.3 to lay a theoretical background. Section 2.4 provides an overview of L2 perceptual training studies in recent decades and some methodological issues that are central to the construction of the present research. Research gaps observed in previous studies which motivate the current research will also be highlighted. After having the understanding of the subject matter, readers will be presented with Section 2.5 which spells out the current research background. A prior full-scale study conducted by me which motivates the present study will also be introduced in Section 2.6. This chapter ends in Section 2.7 by listing the research questions of this study.

2.1 Studies of L1 & L2 Speech Perception and Factors Affecting L2 Speech Learning

2.1.1 Introduction

As noted in the previous chapter, language transfer, as a major factor that influences L2 phonological acquisition, has played a significant role in the construction of different speech learning models or theories (e.g. Best, 1994; Flege, 1995; Kuhl, 2001). Even with different emphasis and characterizations, an L2 speaker's L1 categories would usually be addressed as a possible assimilation target of the newly-encountered L2 segments in these models. This highlights that empirical studies of infant and adult monolingual or cross-linguistic speech perception are worth being discussed at the outset to deepen our understanding of L2 speech learning problems. In addition, L2 speech learning is also believed to be attributed to other factors, such as age and experience of learning the L2, phonetic salience, L2 instructional intervention and attention reorganization. This entire section is devoted to surveying and discussing some empirical studies of L1 and L2 speech perception, the role of L1 in L2 speech learning and factors affecting this.

2.1.2 Infant L1 & Non-Native Speech Perception

Since the early 1970s when new techniques were advanced enough to scrutinize infants' speech perception capabilities, research began to emerge and offered insightful understanding to researchers not only about the early tuning process of human auditory system, but also the possible causes of adult L2 learners' speech problems.

Infant speech perception studies can let researchers learn about how human beings categorize speech sounds during the early months of life. From the findings of infant perception studies so far examined, researchers have a general consensus that infants are endowed with the ability to detect acoustic-phonetic properties crucial for discriminating all sound contrasts in any language (e.g. Aslin, Pisoni, Henessy, & Perey, 1981; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Lasky, Syrdal-Lasky, & Klein, 1975; Streeter, 1976; Trehub, 1976). In other words, infants are able to perceive speech sounds categorically, regardless of their ambient language. Nevertheless, by the second half of the first year of life, when their phonological categories are forming through experience with language input, infants appear to be less sensitive to contrasts that are not relevant to their ambient language (Best, 1994a; Gerken & Aslin, 2005; Jusczyk, 1997; Kuhl & Iverson, 1995; Polka & Bohn, 1996; Strange, 1995; Werker, 1994; Werker, 1995; Werker & Curtin, 2005; Werker & Polka, 1993). Studies have shown that this loss of non-native contrast discrimination ability applies to both consonants (e.g. Best, McRoberts, & Sithole, 1988; Werker & Tees, 1983, 1984) and vowels (e.g. Bosch & Sebastian-Galles, 2003; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Kuhl & Iverson, 1995; Polka & Werker, 1994; Werker, 1995), though in different periods, to different degrees and in different manners (Houston, 2011).

Research by Eimas et al. (1971) which gauged one- and four-month-old infants' discrimination ability of consonants with voice onset times (VOTs) along a continuum has shown that infants perceived VOTs categorically; whereas follow-up studies which tested other elements such as voicing, place of articulation or manner of articulation (e.g.

Eimas, 1974, 1975; Eimas & Miller, 1980a, 1980b) also suggested that infants are born with the ability to categorically discriminate consonants during the first six months of life in general. Possibly susceptible to the surrounding input, infants' discrimination, however, is not categorically rigid and fixed as it may seem. Maye, Werker, and Gerken (2002) found that when infants are attended to different stimuli conditions, they may not be able to discriminate VOT contrasts across category boundaries; McMurray & Aslin (2005) showed that infants are sensitive to within-category variation, that is, the prototypicality of phonemes. In conclusion, the above studies generally concluded that non-native consonant contrasts are discriminated categorically.

The decline of sensitivity to non-native consonant contrasts appears to begin late in the later six months of the infant's first year, as infants are exposed to more language input from the ambient language. For instance, several Hindi contrasts were readily discernible by 6 to 8 months' old English learning infants but not the 10 to 12 months' old (Werker, Gillbert, Humphrey, & Tees, 1981). When were asked to discriminate some consonantal contrasts, infants from different language backgrounds with younger ages (6-10 months but not 10-12 months) still were found to still possess the discrimination ability (Werker & Tees, 1984). Some other studies (e.g. Best et al., 1988; Trehub, 1976; Tsushim, Takizawa, Sasaki, Shiraki, Nishi, Kohno, Menyuk, & Best, 1994) obtained similar findings, suggesting that age and language experience have an effect on the loss of universal perceptual patterns. With the ambient language, infants become more sensitive to sounds that are more linguistically relevant to them and show a decline in the ability in discerning contrasts that are not their native language. Tsao, Lui and Kuhl

(2006) found that both English- and Mandarin-speaking infants between 6 and 12 months old showed a significant amelioration in their listening to their native language contrast whereas they had a decrease in the sensitivity of non-native contrasts; Sundra, Polka and Genesee (2006) found that among English-speaking infants and children, the discrimination ability of the /d/-/ð/ contrast was enhanced with language experience, while French-speaking children showed no improvement. Taken together, many studies have given a strong indication that the language-general sensitivity begins to decline only around ten to twelve months of age, meaning that language experience affects infants' perception of consonant contrasts.

For developmental changes in vowel perception among infants, there have been, however, relatively fewer studies. Studies in general show that infants at around six to eight months of age begin to lose language-general sensitivities toward vowel contrasts (Bosch & Sebastian-Galles, 2003; Kuhl et al., 1992; Polka & Werker, 1994). One of the earliest works by Trehub (1973) demonstrated that the perceptual differentiation was preserved for some native vowel contrasts (/a/-/i/ and /a/-/u/) and non-native Polish/French oral nasal vowel contrast [pa]-[pã] among one- to four-month-old English-speaking infants. Kuhl and colleagues (1992) discovered that six-month-old English and Swedish infants, but not the four-month-olds, tended to generalize around prototypical vowel contrast exemplars (/i/ and /y/ with different prototypicality values) despite the fact that non-prototypical ones could still be discriminated. This gave the evidence of a "perceptual magnet effect" in native-language vowel categories, which will be further elaborated in Native Language Magnet Model in Section 2.3.4. Another study by Best (1999) showed that only three- to five-month-old English learning infants could discriminate some Norwegian vowel contrasts while the ten-month-olds failed to have such discernment. Polka and Werker (1994) also found that four- to six-month-old English infants could discriminate German vowel contrasts /u/-/y/ and /ʊ/-/ʏ/.

Nevertheless, the results in one study by Polka and Bohn (1996) were found to lack consistency with those of the consonant studies. The English- and German-learning infants' differentiation of non-native vowel contrasts was found to have no evidence of perceptual magnet effect or any influence by age or language influence. Hence, it was argued that even by the end of the first year of life, some non-native contrasts may have the immunity to language-specific processing as the sounds may be inherently easier to discriminate. In addition, vowel perception is usually more continuous but not as categorical as consonant perception, meaning that the perceptual process is different (Polka & Bohn, 1996; Stevens, Liberman, Studdert-Kennedy, & Ohman, 1969; Strange, 1995). The variations in results thus suggest that further investigation on what factors may facilitate or hinder vowel perception in early infancy is deserved before making more solid conclusions.

2.1.3 Adult Monolinguals' Non-Native Speech Perception

While infants are believed to be language-general perceivers at the phonetic level who can differentiate speech sounds from any of the world's languages, speech perception among monolingual adults is viewed as a language-specific processing (Bohn, 1998; Polka, 1992; Strange, 1995). Studies (e.g. Abramson & Lisker, 1970; Best

& Strange, 1992; Goto, 1971; Lisker & Abramson, 1967; Polka, 1992; Polka & Werker, 1994; Rochet, 1995; Strange, Akahane-Yamada, Kubo, Trent, & Nishi, 2001; Werker & Lalonde, 1988; Werker & Logan, 1985; Werker & Tees, 1984) have provided evidence that adult monolingual speakers' perception and labeling of non-native sound categories are usually a function of language-specific experience; hence they have notable difficulty when differentiating non-native contrasts.

For instance, Best and Strange (1992) tested the perception of English approximants among monolingual Japanese and English adult listeners. The Japanese subjects were found to identify more /w/ stimuli in an English /w/-/j/ synthesized continuum than English speakers as the Japanese /w/ is phonetically closer to /j/. The position of the perceptual boundary in the categorization of speech is reported to have been influenced by the subjects' exposure to the ambient language. Rochet (1995) also reported that monolingual Portuguese and English listeners showed different assimilation patterns of /i/ and /u/ categories to French /i/-/y/-/u/ continuum due to the unique way that those vowel contrasts are phonetically perceived and produced in their native language, meaning that their native language experience has affected the non-native sound categorization process.

Language experience, however, is not the sole factor that determines the ease of non-native sound contrast discrimination by adult monolinguals (Best & Tyler, 2007). More recent work (e.g. Best, Hallé, Bohn, & Faber, 2003; Best, McRoberts, & Goodell, 2001; Best, Traill, Carter, Harrison, & Faber, 2003; Kochetov, 2004) has reported that not all non-native segments are equally difficult to discriminate; some monolingual listeners can even demonstrate nearly native level discrimination of some non-native contrasts. The ease or difficulty of their perception of non-native phonetic contrasts can be attributed to a range of factors, such as the listener's native language (e.g. Best, Hallé et al., 2003; Best, Traill et al., 2003), coarticulatory patterns (e.g. Beddor, Harnsberger, & Lindemann, 2002; Bohn & Steinlen, 2003) and allophonic or phonetic variations (e.g. Flege & Hillenbrand, 1987; Gottfried & Beddor, 1988; Harnsberger, 2000, 2001).

A totally different picture was reported in some studies (e.g. Grieser & Kuhl, 1989; Iverson & Kuhl, 1995, 1996; Kuhl, 1993) relevant to discrimination of exemplars. Sounds that are not exactly contrastive but gradient-like within-category consonant or vowel variations, i.e. good or poor exemplars of the sound (in Grieser & Kuhl, 1989, [i] along a prototypical continuum was used), were found to display asymmetrical discrimination results: listeners have a better performance in discriminating poor exemplars than good ones. Language experience hence was claimed to impinge on listeners' identification of exemplars with good prototypicality. Added to this complication are the universal (rather than experience-tuned) phonetic sensitivity biases observed (e.g. Polka & Bohn, 2003) among both native and non-native listeners.

Evidence has supported that adult monolinguals are language-specific perceivers; yet the complications of language experience and other influences have not been fully disentangled. As previously stated, although both infants' early perception and phonological development research plus adult monolinguals' perception studies have laid a critical ground for our understanding of subsequent language acquisition including that of the L2, more research is required.

2.1.4 L2 Learners' Speech Perception

L2 speech learning, in particular the perception of non-native L2 phonetic contrasts, is usually effortful and difficult for many learners. A number of studies (e.g. Bradlow et al., 1997; Lively et al., 1994; Logan et al., 1991) in previous several decades have reported that many L2 learners have problems in learning different types of non-native contrasts. Problems in the perception of non-native place of articulation contrasts such as Japanese learners' perception of English /w/-/j/ and /w/-/ɹ/ (Best & Strange, 1992) or English speakers' perception of Hindi dental vs. retroflex stops /t̪/-/ʈ/ and/or Thompson velar vs. uvular ejectives /k'/-/q'/ (Polka, 1991, 1992; Werker, Gillbert, Humphrey & Tees, 1981; Werker & Tees, 1983, 1984) have been found in previous investigation. Other non-native contrasts such as voicing (e.g. Rochet, 1995; Rogers, 1997; Werker et al., 1981) or vowel contrasts (e.g. Baker & Trofimovich, 2000; Bohn & Flege, 1990; Escudero, 2005; Famoso, Schwartz, & DaSilva, 1998; Flege, 1997; Gottfried & Beddor, 1988; Ingram & Park, 1997; Lambacher et al., 2005; Munro, 1993; Munro, Flege & MacKay, 1996, Pallier, Colomé & Sebastián-Gallés, 2001; Polka, 1995; Strange, Yamada, Kub, Trent, Nishi, & Jenkins, 1998; Wang, 1997, 1998; Wang & Munro, 1998, 1999; Wang, 2002; Wong, 2010, 2012, 2013) in different languages have also been identified in L2 perception studies. Japanese speakers' learning of English /ɹ/-/l/ distinction is also one particular issue that has raised a lot of attention in the field (e.g. Bradlow et al., 1997; Lively et al., 1994; Logan et al., 1991; Sheldon & Strange, 1982; Yamada, 1993; Yamada, 1999).

All these studies, together with the results from infant and monolingual adult speech perception studies, have led to our understanding that the reduction in language-general perception sensitivity, attunement of phonetic perceptual processes through language experience and a wider variety of factors pertaining to the learner's traits, learning background and L2 experience, etc. are some underlying causes of the L2 learning conundrum. In the following section, before revisiting some current models and theories on L2 speech perception and/or production, factors that have mainly been focused on in the literature of L2 speech perception learning, or processing in general will be explored first.

2.1.4.1 Major Factors Affecting L2 Speech Learning

2.1.4.1.1 The Role of L1 in L2 Speech Learning and Their Interaction

One fundamental difference between L1 and L2 speech learning is due to the fact that L1 learners begin with a blank slate, which can never been recreated by any L2 learners as they have a linguistic system in place, or compensated by any natural (even ideal) learning condition with an optimal learning environment. This difference is thus widely recognized as the major and largest contributor precluding successful L2 speech learning, particularly for older learners whose L2 has always been called accented or fossilized. Influential as it is, L1's role in SLA is worth a focus and how significant it is on L2 speech learning in particular will be expanded here.

Second language speech perception research (e.g. Baker & Trofimovich, 2000; Morrison, 2006; Wang, 1997, 1998) has revealed that L2 learners tend to assimilate L2 phones to existing L1 speech sound categories, leading to the notion that learners' linguistic experience with the L1 has a tremendous effect on the ultimate attainment of L2 phonology. As briefly illustrated in Section 2.1, early attempts explaining L2 speech problems first began with a focus on errors discovered in L2 speech samples. Lado's (1957) Contrastive Analysis Hypothesis (CAH), which gave straightforward and explicit predictions of L2 speech learning difficulties via comparing and contrasting the similarities and differences between the learners' L1 and L2, suggested that any difference between the L1 and L2 would contribute to the failure of L2 phonological acquisition due to negative transfer; positive transfer will only occur when a corresponding category in L1 can be mapped to by the L2 sound. As Lado (1957) argued, "those elements that are similar to [the learner's] native language will be simple for him, and those elements that are different will be difficult." Some studies have supported this claim (p. 2). Wang (1997, 1998), who scrutinized Mandarin speakers' perception and production of Canadian English vowels, discovered that learners suffered negative transfer by substituting some L1 vowels for the English ones as no similar counterparts could be found in their native language. Bongaerts, van Summeren, Planken and Schils (1997) found that late Dutch learners of English could speak with no detectable accent as determined by a group of native English speakers. The authors attributed the finding to the typological similarity of Dutch and English as the two systems have shared many similar categories, causing less negative transfer. Flege, Bohn and Jang (1997) and Escudero (2005) in two different studies also found that Spanish learner who did not have /iː/-/ɪ/ contrast in their native language had problems in learning the English contrast while Hojen and Flege (2006) found similar difficulties among Spanish learners learning other English vowel contrasts. The role of L1 is highlighted in the CAH as the degree of typological similarity between the L1 and L2 phonological systems is directly associated with the level of success of L2 speech acquisition: the more similar the two systems, the less detectable accent the learners will attain.

Nevertheless, neither can all difficulties or success, nor the order and rate of learning can be explained simply by the CAH. In learning Salish glottalized velar and uvular stop contrasts, positive transfer could not be found for Farsi speakers who were expected to have an advantage over English speakers as Farsi speakers shared the same place feature contrast in their native language with Salish (Polka, 1991). Yet, the results did not match with the prediction that positive transfer would occur. Similarly, Jamieson and Morosan's (1986) study researching native Canadian French speaker's acquisition of the English dentals /θ/-/ð/ showed that voicing feature shared by L1 and L2 segments may not necessarily benefit the learners by positive transfer. These studies suggested that solely comparison between L1 and L2 cannot explain all the errors, but SLA as a developmental process may resemble children acquiring their L1 (Corder, 1971; Selinker, 1972). Also, as stated in the reformulated version of the CAH, i.e. Eckman's (1977) Markedness Differential Hypothesis, some sounds that are less marked may just be universally easier for acquisition. Some studies even showed an opposite prediction of the CAH, that is, L2 categories which are the least similar to L1 categories would facilitate, rather than hinder, L2 speech learning due to the saliency between the two sounds. Bohn and Flege (1992) found that German learners of English produced the dissimilar English /æ/ more authentically; Major (1987) showed that Portuguese learners of English also had more difficulty in learning similar sounds than dissimilar ones; Major and Kim's (1996) study testing Korean learners of English also found that dissimilar phenomena were acquired faster than similar ones. That the rate of acquisition is faster for dissimilar sounds than similar sounds in these studies can be summarized in one claim postulated by Major and Kim (1996) called the Similarity Differential Rate Hypothesis.

As also examined with L2 learners residing in the L2 country, the role of L1, in particular its degree and frequency of use in L2 learning environment, is identified as crucial in L2 perception and production. Not focusing only on the effect of age of arrival in L2 speech learning, Flege and MacKay (2004) tested the influence of high and low L1 use in L2 learning among four groups of Italian learners of English in the discrimination of English vowels; Flege, Frieda and Nozawa (1997) also found high correlation between the use of L1 and L2 and the degree of accent. Subjects with low L1 use outperformed those with high L1 use. Studies testing other sound segments such as English consonant perception (MacKay, Meador & Flege, 2001), English vowel production (Piske, 2002) or the degree of foreign accent (MacKay, Flege & Imai, 2006; Piske, MacKay & Flege, 2001) in general reported similar results about the relation between ongoing L1 use and L2 speech learning: the lower the use of L1, the more successful the acquisition would be. In fact, L2 learners were also found to rely on L1 phonetic cues to discriminate L2 sounds differences (Bohn, 1998; Gottfried & Beddor, 1988; Munro, 1993; Strange et al., 1998).

The abovementioned literature gives evidence that the L1 has a major role in L2 speech learning. Following the same vein of understanding, several researchers have hypothesized models of L2 speech perception and production. Many of them have received extensive research focus and four most frequently cited frameworks of L2 speech perception (and production) namely the Speech Learning Model (Flege, 1991, 1995b, 1999, 2002, 2003), the Perceptual Assimilation Model (Best, 1994a, 1995; Best, et al.,1988; Best et al., 2001; Best & Strange, 1992; and Best & Tyler, 2007), the Native Language Magnet Model and its expanded version (Grieser & Kuhl, 1989; Kuhl, 1991, 1992, 2000b; Kuhl & Iverson, 1995; and Kuhl et al., 2008) and Second Language Linguistic Perception model (Escudero, 2005) are of high relevance to the present research. These models all embrace the discussion of the effects of L1 tuning on L2 perception and the developmental nature of L2 learning. Since there also exists an intricate relationship between perception and production, I intend to isolate here the detailed presentation of the three perception models; Section 2.3 will continue with the line of reasoning to foreground the claims, foci and empirical supports of these models.

As the above empirical studies have shown, the role of L1 system is of high influence to L2 speech learning. Apart from this factor, learners' L2 experience and instructional intervention leading to a change of attention may also be other causes of the problem, which are to be explicated in the following parts.

2.1.4.1.2 Age and Learners' L2 experience: Quantity and Quality

Having established that L1 system affects L2 speech perception and production, a number of researchers explored further to what extent L2 experience influences L2 phonological development. How the correlation between age and language experience affects L2 learning and the quality of L2 exposure will be discussed in the following parts.

Even after years of L2 learning, late L2 learners retaining a detectable accent and finding it difficult to perceive some L2 sounds are still commonplace phenomenon. The age of L2 exposure indexed by the age of arrival (AOA) of immigrants in a predominantly L2-speaking country or the age of L2 learning (AOL) has hence become by far the most researched experience-related factor to the success of L2 speech learning (some studies like Flege and colleagues in a 1995 study even found it has powerful effects; MacKay and colleagues in a 2006 study suggested it as the strongest factor). The idea that language acquisition is biologically constrained and influenced by age was centered on the Critical Period Hypothesis (CPH) proposed by Lenneberg in 1967 in relation to a critical period within which learners can have the first language developed readily while beyond which language acquisition will be effortful. The strongest instantiation of this claim says that when learners grow up to around puberty, their neural plasticity or cognitive ability will suffer a sharp loss which hinders language acquisition. This concept of the CPH has been extended to second language acquisition as it can provide appealing explanations to L2 speech learning problems, particularly on why younger learners are better than older ones and why the later the AOA, the stronger the foreign L2 accents. This is why age-related factors have received much focus in different aspects of SLA studies such as in syntax (e.g. DeKeyser, 2003; Patkowski, 1980), grammar (e.g. Johnson & Newport, 1989), or more evidently, pronunciation in general (e.g. Patkowski, 1990) and the perception and production of L2 segments (e.g. Flege et al.,1999; MacKay et al., 2001; Piske et al. 2001).

Although there have been significant empirical observations showing that age is influential to L2 speech learning and give support to the CPH, a great number of studies have used “exceptional cases” to challenge the biological age effects on SLA. They include: 1) the success of developing authentic or native-like L2 by postpubertal learners, as judged by native speakers of the language (e.g. Birdsong, 1992; Birdsong & Molis, 2001; Bongaerts, 1999; Moyer, 1999, 2004); 2) late L2 learners defying the odds performing equally well with early L2 learners to achieve native-like level whereas early bilinguals not performing as well as the late ones (e.g.; Bongaerts, Planken, & Schils, 1995; Bongaerts et al., 1997); 3) possible modification of the ability in non-native contrast perception discrimination through laboratory setting (e.g. Lambacher et al., 2005; Iverson & Evans, 2007). These studies give evidence that there is no clear drop-off of the L2 learning ability when the learner has gone through the sensitive period as the system still remains plastic; it just declines gradually and continuously with age (Flege, Munro, and MacKay, 1995; Hakuta, Bialystok, and Wiley, 2003). To date, empirical evidence is still insufficient to verify the nature of the role of age. Yet, since age is inherently confounded with other factors, such as the learners’ L2 experience, many researchers are still trying to investigate this intricate issue.

The effects of age on L2 learners’ speech learning cannot be rejected even though it has been testified that biological factors as stated in CPH are not the primary and direct cause of L2 acquisition difference, because age still highly correlates with the amount of time that one can have contact with the L2 speaking community. As Flege et al. (1995) and MacKay et al. (2006) have pointed out in their two different studies, the amount of L2 experience that L2 learners receive is one factor that exerts a high effect on the degree of accent. They found that the amount of L2 experience would be a more robust predictor of the degree of success of L2 speech learning than maturational constrains. This partly explains why L2 learners who start earlier are judged as having a milder degree of accent than late learners. This is also linked to the role of the L1 in L2, because the use of the two languages is constantly in competition during the course of SLA. Provided that they have the same amount of L2 experience, younger learners are hence, in theory, able to better overcome negative transfer brought by their L1 system to the L2 as they have less L1 experience than older learners. Some recent work suggests that early L1 language experience just impedes acquisition of non-native sounds during adulthood (Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann & Siebert, 2003; Kuhl et al., 2008).

Length of residence (LOR) in the L2 setting, as one common measurement of L2 experience, has also been examined and has supporting evidence of its effects on L2 perception and production of both segmental and suprasegmental patterns (Asher & García, 1969; Flege & Fletcher, 1992; Flege, Bohn & Jang, 1997; Flege & Liu, 2001; Jia, Strange, Wu, Collado, & Guan, 2006; Piske et al., 2001; Trofimovich & Baker, 2006; Trofimovich, Baker, & Mack, 2001). Learner’s perception and production of new sounds improved when they spent more time in the L2 (Flege, MacKay, & Piske, 2002;

Flege, MacKay & Meador, 1999). In sum, it is the quantity of L2 experience (or the relative amount of L1 to L2 use) which is naturally related to age, rather than the biological neural constrains, that correlates with the ultimate attainment of L2 speech learning.

Yet, not all studies reported the effects of L2 experience indexed by LOR (e.g. Oyama,1976; Flege, 1988, 1993; Flege, Munro, & Skelton, 1992; Tahta, Wood, & Loewenthal, 1981; Piper & Cansin, 1988); some even argued that the role of L2 experience had been overstated due to the difficulty in precisely quantifying the exposure (McAllister, 2001) and the difference in the quality of L2 input (Flege & Liu, 2001). Gauging L2 exposure by using weeks, months or years of AOA, AOL or LOR as units is not precise enough since the labels do not represent leaners' authentic and meaningful contact with the ambient language, leading to inaccurate or over-generalized predictions. Piske and colleagues (2001) called it "a rough index" of L2 experience and they found that the effect of LOR will only be significant between groups with relatively large difference in years of residence. McAllister (2001) even directly pointed out that the role of experience indexed by LOR has been exaggerated, as he found tenuous correlation between LOR and success of L2 speech learning among his Swedish subjects learning English as an L2. It is true that learners have their own ways of engaging in and experiencing with the L2 community, of which a precise quantifying process is difficult. Studies investigating language experience by LOR indictor should ascertain a better definition of the amount of exposure.

Even if a better way to quantify L2 exposure is found, the quality of each input that each subject receives will not be identical. More consensus have been reached among researchers that it is the quality of L2 experience and the use of it that determine long-term gain in the phonological aspect (Flege, Birdsong, Bialystok, Mack, Sung, & Tsukada, 2006; Jia & Aaronson, 2003; Moyer, 2004, 2009). Only active use and practice of the L2 language contribute authentic and meaningful input to the learners. Flege and Liu's (2001) investigation discovered that LOR was a significant factor affecting phonological perception only for students, but not for non-students. They attributed this finding to the quality of input that the student group can receive as students can enjoy language practice with native teachers or schoolmates, meaning that language contact with active use will make a difference. Moyer (2004) further suggested only LOR with meaningful input like interactive contact with native speakers or formal instructions can one see its significance in L2 learning. Learners who spent more time in various learning modes (even if they were informal interactions with the L2, such as watching television or movies, talking with native speakers of the L2, listening to radio programs) were found to have become more native-like in terms of their speech fluency and automaticity (Derwing, Thomson & Munro, 2006). A study by Flege et al. (1999) even indicated that non-interactive activities such as watching television or movies are good predictors of accent. Thus, only active language practice and use, an interaction between the quantity and quality of L2 experience, is contributing to the L2 speech learning.

To conclude, the degree of L2 experience, whether it is measured by AOL or LOR or other interactive ways, has an obvious effect on L2 speech learning. As seen from some of the studies above, both the quantity and quality of the L2 exposure also have a demonstrable influence. One of the steps taken during SLA to improve the quality of L2 exposure is to direct learners' attention to the learning targets by offering them pedagogical interventions. The details of the effects of selective attention will be discussed in the following section.

2.1.4.1.3 The Role of Selective Attention in L2 Speech Learning

Be it a natural learning environment or an L2 classroom, it appears inconceivable for one to concentrate on every detail during learning. Learners can only select, though not necessarily consciously, to pay attention to some parts of the learning circumstance. That is what is referred to selective attention. Infants acquiring their L1s are believed to benefit from infant-directed speech produced by caretakers who deliberately shorten and simplify utterances, exaggerate pitch changes, lengthen vowels and elongate pauses with a view to catching infants' attention to the speech components (Liu, Kuhl & Tsao, 2003; Kuhl & Iverson, 1995). The enhanced phonetic saliency and noticeability of the input facilitate the acquisition of L1 among children, showing a good example of drawing one's attention during language acquisition. It leads to researchers' speculation that SLA might also have some similarities.

In fact, many researchers have theorized that attention is important during SLA (Flege, 1995; Leow, 1997; 1998; Schmidt, 2001; Simard & Wong, 2001; Thomson, 2003; Tomlin & Villa, 1994). For example, Flege's (1995) study showed that in some contexts, L2 learners were able to make use of important phonetic information in a discrimination tasks only when their attention was captured more directly to the targets.

Thomson (2003) also pointed out that in naturalistic environments in which the phonetic learning contrasts were not made salient, learners usually overlooked the distinction by simply assimilating L2 sounds to the existing L1 categories, implying the need to enhance the saliency of acoustic cues to capture learners' attention.

The concept of attention has been addressed in the field of SLA for long as it has an associated role with many aspects of theories such as input, variation, instruction, etc. Given that both the quantity and quality of L2 exposure are important in SLA, it is logical that more meaningful input of the target brought by raising learners' attention is likely to lead to learning; instructions given to learners also draw learners' attention to important cues or targets of the learning. Attention is closely related to terms such as consciousness, noticing, awareness and understanding, and is highly associated with concepts in cognitive psychology such as memory, learnability, and connectionism. Schmidt (1994a) identified attention – the detection of a stimulus – as one of four dimensions of consciousness. In his Noticing Hypothesis, which states that noticing is "the necessary and sufficient condition for converting input into intake" (Schmidt, 1990, p. 129) and "the registration [detection] of the occurrence of a stimulus event in conscious awareness and subsequent storage in long term memory…" (Schmidt, 1994b, p.179), he emphasized the importance of noticing and paying attention to L2 forms.

Posner and Petersen (1990) also suggested three networks of attention: alertness, orientation and detection. Alertness concerns the state in which learners select input for processing: the higher the level of alertness, the faster the speed of processing. Learners should focus on the current task but not the other distracting ones. Orientation means learners align their attention to a specific stimulus from a pool of available resources. When attention is oriented to that stimulus, the processing of the stimulus will be facilitated. At the end, it is only when the stimulus is cognitively registered, i.e. being detected, that it becomes available for processing. In sum, these three networks constitute the multi-faceted concept of attention, of which learners should get ready to learn (alertness), be directed to certain form or meaning (orientation) and know what critical information to focus on (detection). Research (cf. Lyster and Ranta, 1997) examining some types of second language instruction in naturalist classroom settings found that only those target forms which are made more salient by corrective feedback or other ways are better detected by learners; some studies (e.g. Leow, 2001; VanPattern, 1996, 2002) also reported positive learning of orienting learners' attention to the target forms and that the rate of acquisition of some synthetic or morphological structures increased through instruction which direct learners' attention (Ellis, 1990).

The role of selective attention in phonetic learning can also be seen in some studies. A study by Guion and Pederson (2007) who probed into the role of attention in L2 phonetic training among two groups of English monolinguals learning five Hindi consonantal contrasts ([b]-[t̪], [k]-[ɡ], [b]-[b^{ɦ}], [k]-[k^{h}] and [t̪^{h}]-[ʈ^{h}]) also demonstrated that orienting learners' attention by instruction to particular phonetic forms can facilitate speech perception learning. One group was trained to pay attention to the difference of the stimuli in the onset sounds (sound-attending group) and the other to focus on the meanings of the word pairs (meaning-attending group). The sound-attending group was found to have more improvements in discriminating some of the sound contrasts than the meaning-attending group, whereas the meaning-attending group performed better than the sound-attending group in the semantic test. This can be attributed to the attention orientation effects brought by the experiment. Another study by the two researchers experimenting English speakers learning Hindi word contrasts with onset consonants and medial vowels as the attending target also obtained comparable results, showing that orienting attention to the learning targets through instructions can enhance the learning of that target (Pederson and Guion-Anderson, 2010).

Various phonetic training studies have also demonstrated that cue enhancement which orients learners' attention have positive effects on L2 speech learning. As early as Jamieson and Morosan's (1986) study which employed the perceptual fading technique to exaggerate fricative cues in synthetic exemplars of English /θ/-/ð/ pair, it is attested that the modified training stimuli can improve the accuracy in identifying the contrast among adult francophones. Since their attention was focused towards the target cues, learning was successful. Not only did they improve in identifying the synthetic contrast, but transfer of learning to natural stimuli was also one significant result of the study. Jamieson and Morosan (1989) replicated the study and further experimented if learning could be transferred to stimuli produced by new speakers and in new contexts. Significant improvement was obtained although generalization to new phonetic environments was not found. This result supports the claim that capturing learners' attention by using prototypical exemplars can facilitate L2 speech learning. In Kondaurova's study (2008), the role of attention in phonetic learning is stressed by comparing several training methods: adaptive training in which cues are enhanced to capture attention, inhibition training in which learners are trained to ignore the cues, and normal training with no change in cue weighting schemes. Native Spanish listeners were trained to distinguish English [iː]-[ɪ] pair by shifting their attention to spectrum rather than duration as cues that define a non-native contrast. Native speakers rely predominantly on spectral properties to distinguish the two vowels, thus it is hoped that the subjects could redirect their attention from duration to spectrum for a more native way of perception. One of the training methods, the adaptive training which trained the subjects to redirect their attention to the spectral cues, was a more successful means to help learners' distinguish the vowels. These positive results imply that orienting learners' attention by modifying learners' cue-weighting schemes or enhancing the target saliency is important in L2 speech learning.

To make good use of attention for successful speech learning does not only rely on enhancing cues or targets, but enriching the variability of stimuli has also been one way that researchers adopted for better L2 speech training. One training paradigm called the High Variability Phonetic Training (hereafter, HVPT) approach has become popular in training the perception of non-native L2 segmental contrast, which was originally a follow-up action taken in response to the failure of generalization found in Jamieson and Morosan's (1989) study. HVPT has utilized natural training stimuli produced by several native speakers of the target language to enrich the stimuli variability and has been found to have better results in learning, even in new contexts . When learners are provided with sufficient variations of the target segments, they are also provided with more opportunities to attend to, detect or notice the crucial acoustic information of the sounds. Higher stimulus variability appears to give a larger amount of target tokens for learners to focus on, rather than only exemplars or models which are well recognized by native speakers of the language but easily confused or assimilated by L2 learners with L1 perceptual magnets near those exemplars. This is not to downplay the role of prototype training since it can be very beneficial: as also pointed out in Jongman and Wade's research (2007), high variability training promotes learning for easily separable contrasts, but still for some difficult distinctions, learners may require training on prototypes with minimal variability so as to learn. Minimal variability of the exemplars allows learners to turn their attention to focus on the best or model examples of the sounds which is sufficiently focused for learning contrasts that have overlapping parameters. Taken together, selective attention, either by orienting learners' attention to prototypical examples with minimal variability or by providing learners with a wider exposure of target token variability for a presumably higher chance of detection, is playing a weight role in L2 speech learning. In later sections, how stimuli variability and HVPT paradigm are playing a role in perception modification will be discussed in detail.

2.1.5 Summary

The degree of L1 and L2 similarity, the amount and quality of L2 exposure, age and length of L2 learning as well as the role of instructional intervention have received the most attention in the literature of factors influencing L2 speech perception and production. The results of the research in this area have led researchers to propose a number of speech learning models and theories which have laid the ground for explicating the difficulties in L2 perception and production. In the following sections, I will turn to revisit some influential speech learning models before investigating the relationship between speech perception and production.

2.2 Theories of L2 Speech Perception

2.2.1 Introduction

Four models that are closely related to speech perception are Speech Learning Model, the Perceptual Assimilation Model, the Native Language Magnet Model, and the Second Language Linguistic Perception model, all of which will be introduced here. They hypothesized that difficulties in L2 speech learning are due to assimilation of L2 sounds to L1 categories and attempted to account for L2 speech learning difficulties.

2.2.2 Speech Learning Model

Flege's Speech Learning Model (SLM; Flege, 1991, 1995b, 1999, 2002, 2003) is based on the systematic similarities and differences found between the sounds in the L2 and L1 systems. The purpose of this model is to account for how L2 learners, particularly the experienced ones, fail or succeed in learning to perceive and produce L2 phonetic segments. The SLM describes the phonological acquisition of an L1 as a bottom-up process and L2 as a top-down process, which predicts that foreign sounds will be poorly perceived and produced if the phonetic differences that distinguish contrasting foreign sounds do not distinguish contrasting native sounds, implying that

L2 speech accents are perceptually-based and that perceptual accuracy can limit the production performance. As a result, successful L2 phonological acquisition can be obtained only when correspondences between the phonetic systems of the L1 and L2 have been established. A new L2-category will be formed only when the learner can detect sufficient perceived differences between the L2 sound and a closest L1 sound; otherwise, a single phonetic category called *diaphone* that bears the properties of both the L1 and L2 sounds will be resulted. Learners will process the perceptually linked L1 and L2 sounds (diaphones) and eventually resemble one another in production (Flege, 1995). Also, it is believed that the phonetic segments that make up both the L1 and L2 phonetic subsystems exist in a common phonological space, thus the phonetic categories of the L1 and L2 interact and interference are bidirectional (Flege, 2005; Flege & MacKay, 2004). The SLM, as a dynamic model, also points out that L2 speakers will become more able to perceive and produce a foreign phonetic difference when they encounter it for a long time and begin early in life. Therefore, speech-learning mechanisms are posited as intact across the life span and that early L2 learners have advantages over late learners due to the fact that L2 categories are affected when L1 categories are developing with age (Flege & MacKay, 2004; Hazan & Barrett, 2000; Lee, Potamianos & Narayanan, 1999). Changes in the perception and production of native sound differences may also result if the encounter with foreign sound contrasts is long and early in life.

There are in total four hypotheses stated in the SLM. They are:

Hypothesis 1. The greater the perceived dissimilarity of an L2 sound from the closest L1 sounds, the more likely a new category will be formed for the L2 sounds. These L2 sounds are called “new” sounds.

Hypothesis 2. Category formation for an L2 sound becomes less likely through childhood as representations for neighboring L1 sounds develop.

Hypothesis 3. When a category is not formed for an L2 sound because it is too similar to an L1 counterpart, the L1 and L2 categories will assimilate, leading to a “merged” L1-L2. These L2 sounds are called “similar” to L1 categories and will pose more difficulties because “equivalence classification” interferences with the establishment for similar phones.

Hypothesis 4. When a new category is established for an L2 sound, it may dissimilate from a neighboring L1 and/or L2 sound – and vice versa – to preserve the phonetic contrast.

The 1994 version of the SLM posits that perceived L1-L2 phonetic dissimilarity should be regarded as a continuum but not strict tripartite identical-similar-new division, hence it also helps predict L2 learners’ difficulties in three more general ways:

i When category formation is blocked, L1 and L2 categories will assimilate: L2 sounds will continue to resemble L1 sounds whereas L1 sounds will begin to resemble L2 sounds.

ii When a new category is formed for an L2 sound, it and/or the nearest L1 sound may dissimilate.

iii Children are more likely to form phonetic categories for L2 sounds than adults. However, even adults retain the capacity to form new categories.

Several studies (e.g. Flege, 1987; Bohn & Flege, 1990, 1992) have provided empirical support for the SLM. Flege (1987) investigated native English speakers’ production of French phones. It was found that French phone /y/ which was more dissimilar (“new”) from English phones was more accurately produced than “similar” phones such as /u/ and /t/, in line with the SLM’s hypothesis that “similar” L2 sounds are more difficult to be developed into a novel phonetic category. Also, Bohn and Flege (1990, 1992) showed that experienced L2 learners had developed more native-like perceptual patterns of L2 vowels than less experienced ones, verifying that earlier experience of L2 benefits the learners in terms of perceptual category development.

Nevertheless, the SLM was challenged and criticized due to the lack of clear distinctions for the classification of “similar” and “new” sounds: how “similar” and “new” a sound in the L2 is to the L1 is not clear (Ingram & Park, 1997). Also, no explanation could be offered for some studies (e.g. Guion, Flege, Akahane-Yamada & Pruitt, 2000; Munro et al., 1996) which totally or partly disagreed with the SLM’s prediction of the ease of L2 segment production. Therefore, Flege (1995b) used the term “perceived phonetic distance” to signify that the degree of success listeners have in perceiving L2 sounds is based on how much the learner can perceive the phonetic difference between L1 and L2 sounds. The emphasis on the learner can offer better explanations for individual learner differences among the same L1 group.

It is worth noting that the SLM mainly predicts and accounts for the difficulties in L2 speech learning across life and especially for experienced L2 learners. Although some studies (e.g. Guion et al., 2000; Lengeris, 2009) reported that the SLM cannot account for the difficulties for L2 learners at early acquisition stage without any revision, its role as a dynamic model that focuses on L2 learning and the relationship between perception and production has already provided a good theoretical ground for predicting and explaining many L2 speech problems. Such predictions of a link between speech perception and production have already distinguished the SLM from a similar perceptual assimilation model which is a pure perception model, Best’s Perceptual Assimilation Model.

2.2.3 Perceptual Assimilation Model

Originally proposed by Best and her associates to explain L1 listeners' perception of non-native sounds, the Perceptual Assimilation Model (PAM; Best, 1994a, 1995; Best et al., 2001; Best et al., 1988; Best & Strange, 1992; Best & Tyler, 2007) was extended to predict and justify the levels of difficulty that L2 learners have in differentiating L2 segments, based on the assimilation of L2 segments to L1 segments. The PAM is proposed on the basis of direct realism (see Section 2.4.2 for details) and perception of L2 non-native contrasts is based on the degree of articulatory or gestural similarity of the L2 sounds to L1 phonetic categories. Best used specific taxonomies to explicate and predict how well listeners will discriminate different foreign sounds from one another, going beyond simple one-to-one comparison of the L1 and L2 sounds.

These are the possible assimilation types and their specific predictions: a) Two-Category assimilation (TC), meaning that the assimilation of two L2 sounds is in contrast to two distinct native sounds, in which the discrimination is excellent; b) Single-Category assimilation (SC), meaning that the two L2 sounds in contrast are only assimilated to just one native sound, either equally well or poor, and hence the discrimination is predicted to be poor; c) SC assimilation is further categorized as sounds assimilating equally to the single native category and those in which one assimilates far more than the other. These two types of assimilations are said to differ in Category Goodness (CG) with reference to the native category, and discrimination is predicted to range from moderate to good; d) Uncategorized-Categorized assimilation (UC) means that one counterpart of the L2 contrast is uncategorized while the other is categorized and is predicted to have good discrimination; e) Uncategorized-Uncategorized assimilation (UU) means the two L2 sounds in contrast cannot be categorized and discrimination varies from fair to good based on how similar the two L2 sounds are to each other and the native sounds; f) Non-Assimilable (NA) pattern is found when non-native sounds are perceived as non-speech sounds and the discrimination is predicted to be very good.

Concerning only the three assimilation patterns where the non-native contrasts can assimilate to native phonemes, i.e. TC, CG, and SC, it is predicted that L2 learners' success in distinguishing different L2 sounds will be ranked TC > CG > SC, with CG performance varying between TC and SC performance depending on the degree of Category Goodness between the L2 sounds perceived.

Some studies (e.g. Best et al., 2001; Best & Strange, 1992; Nagao, Lim and de Jong, 2003) have reported evidence that supports the PAM. For instance, Best and Strange (1992) investigated the assimilation patterns of a group of Japanese speakers of English. A case of TC assimilation was found, where American English consonant /w/ and /j/ could be assimilated to Japanese /w/ and /j/ accordingly. Best et al. (2001) examined a group of American English speakers' discrimination and identification of several Zulu contrasts, such as ejective velar stops with plosive. It was found that the subjects' identification of contrasts followed the predicted assimilation pattern order: TC > CG > SC. Nagao, Lim and de Jong (2003) tested a group of Japanese listeners' identification of English syllable structures and voicing. The PAM accurately predicted their performance and was said to be able to apply in SLA at the prosodic level.

Both the SLM and the PAM are models that hypothesize that perceptual assimilation of L2 segments to L1 sounds plays the main role in accounting for L2 speech problems. The last model in discussion of this part will be Kuhl's Natural Language Magnet Model (and its expanded version), which differs from both the SLM and the PAM in terms of the representation of a phonetic space.

2.2.4 Native Language Magnet Model (and its expanded version)

The Native Language Magnet Model (NLM; Grieser & Kuhl, 1989; Kuhl, 1991, 1992, 2000b; Kuhl & Iverson, 1995; Kuhl et al., 2008) was proposed by Kuhl and her colleagues. The NLM posits that human's speech perception develops from a language-general perceiver (infancy) to a language-specific perceiver (adulthood). In this model, it is stated that native language categories are prototypes and infants develop these prototypes for native categories. The phonetic properties that define a certain class of category will be an essential "space" in the NLM, as they will be the dimensions that each prototype will occupy. For instance, the vowel space is defined by the vowel's formant frequencies. The prototypes are analogized as "magnets" since tokens near a prototype will be drawn perceptually to it, similar to the magnetic effects. The prototypes' locations are also posited as fixed and the L2 and L1 sounds are all drawn to these prototypes as a function of their distance from them in the phonetic space. More distant L2 or foreign sounds will assimilate to another prototype if they are closer to it, or do not assimilate at all if there is no nearby prototype.

Based on a number of empirical evidence, a recent revised version of the NLM

(Kuhl et al., 2008), called the Native Language Magnetic-expanded model (NLM-e), provides five principles that guide the model. They are:

i. *Distributional patterns and infant-directed speech are agents of change.* This principle points out how exaggeration of relevant phonetic differences in infant-directed speech facilitates infants' speech learning, as compared to adult-directed speech (Liu et al., 2003).
ii. *Language exposure produces neural commitment that affects future learning.* The concept of native language neural commitment (NLNC) is proposed in this principle. It argues that a learner's neural networks for language encoding become committed to the L1 patterns and this commitment affects the learner's ability to learn the phonetic schemes of a new language, due to the physical changes of the neural tissue (Kuhl, 2000a, b, 2004).
iii. *Social interaction influences early language learning at the phonetic level.* Social interaction and contact promote and enhance infant's speech learning, especially in complex language learning situation (Kuhl, Tsao & Liu, 2003).
iv. *The perception-production link is forged developmentally.* Through experience with language and vocal imitations, strong linkages between the perceptual representations and production can be built. This connection is developmental in nature and is formed based on perceptual experience and a learned mapping between the two domains (Kuhl & Meltzoff, 1982, 1996).
v. *Early speech perception predicts language growth.* This principle states that infants' native and non-native perceptual abilities in discriminating phonetic contrasts are early predictors of future language development (Tsao, Liu & Kuhl, 2004).

The NLM-e also made several predictions of phonetic development: a) Early bilingual language experience will not have as much as interference on L2 phonetic learning than those learning the L2 late (Kuhl et al., 2003); b) Social contexts and interaction in natural settings promote more durable, robust and potent L2 speech learning (Kuhl et al., 2003; Kuhl et al., 2008); c) The role of neural commitment to the L1 phonetic categories is a potential mechanistic cause of the critical period phenomenon and experience rather than time which determines the phonetic learning and perception of an L2 (Kuhl, Conboy, Padden, Nelson & Pruitt, 2005).

The NLM (and the NLM-e) differs from the SLM and the PAM in the conceptual representation that the locations of prototypes stated in the NLM are fixed in the phonetic space. Also, in the SLM and the PAM, sounds are said to be different from one another in one or more constituent gestures or phonetic properties, and they do not rely on any spatial representation. The NLM also predicts learners' difficulties by stating that there can be discriminability differences between different instances of the same foreign contrast. The expanded version of the NLM also highlights the importance of social interaction and language experience in successful L2 speech learning and proposes that the perception-production link is developmental, but the SLM and the PAM do not emphasize these aspects.

2.2.5 Second Language Linguistic Perception Model

Based on the framework of Stochastic Optimality Theory (Boersma, 1998), the Second Language Linguistic Perception model (L2LP; Escudero, 2005) provides a theoretical and methodological framework that describes, explains and predicts the acquisition of L2 sound perception. The Linguistic Perception model (LP), of which the learning mechanisms and developmental path act as the foundation of the L2LP model, is a phonological proposal that explains how adult perceive sounds and it particularly accounts for L1 sound acquisition (Boersma, Escudero, & Hayes, 2003). It is proposed that a child acquires the sound categories through creating language-specific perceptual mappings. The child's gradual learning device, namely the Gradual Learning Algorithm

(GLA)[2], will allow him to automatically match, identify and map an incoming auditory value (e.g. F1 value) to a perceived auditory category, i.e. an abstract arbitrary label. More arbitrary features will be mapped onto the categories when the lexicon is acquired, as the categories become more optimized to deal with the lexical information. Thus, L1 acquisition of sound perception is fundamentally a process of forming language specific abstract representations.

The L2LP is composed of five theoretical ingredients that predict, explain and describe L2 sound perception. The figure below shows how the five ingredients are connected and sequenced (Escudero, 2005, p.95):

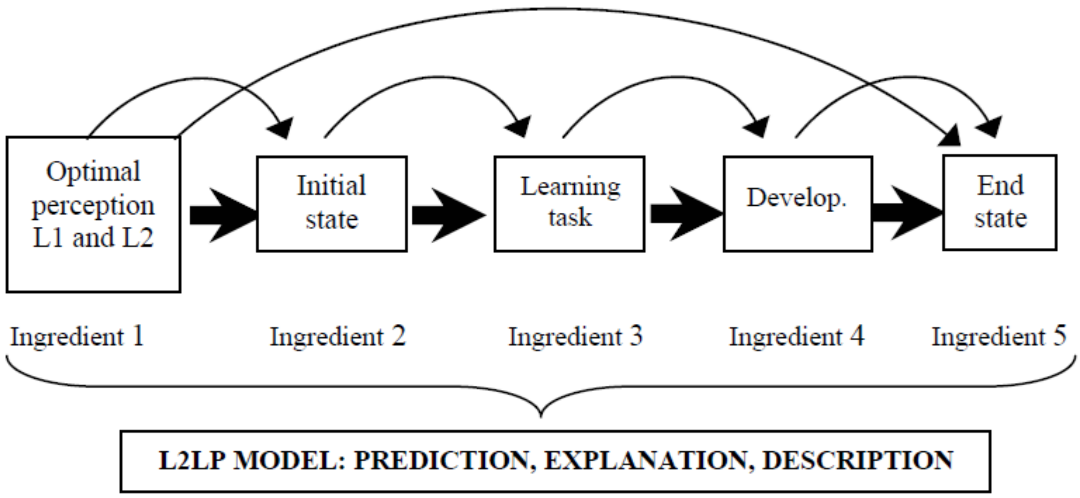


Figure 2.1. The Five L2LP Theoretical Ingredients

Ingredient 1 sets out at the description of optimal perception of the L1 and L2 of a learner, which connects the general LP model with its L2 version and allows further prediction and explanation of three aspects of L2 sound perception. These three aspects

[2] The GLA, proposed by Boersma (1998), performs auditory-driven learning by presenting as an innate identity-matching and distributional learning device that underlies the acquisition of L1 sound perception. It changes the constraint rankings based on the frequency distributions of the auditory input events, hence the most frequent values along the auditory continuum will become the preferred categories. Later, when an abstract form-meaning pair is established for a lexicon, the GLA will become a lexicon-driven learning device that reranks the perception grammar constraints when mismatches occur between the perception output and the lexicon.

are the next three ingredients in the L2LP model: *Ingredient 2* is about the initial state of L2 learning process, that is proposed to be equal to cross-language perception; *Ingredient 3* refers to the learning tasks that allow optimal and targeted L2 sound perception to occur, which is predicted to be equal to the difference between L1 optimal perception and target L2 perception; and *Ingredient 4* is the L2 developmental state that needs to occur in order to achieve optimal L2 perception, which is the same as the learning mechanism used in optimal L2 sound perception. Finally, *Ingredient 5* marks the L2 end state, with both optimal L1 and L2 sound perception. To link all the ingredients together, it means that L2 perception starts at a copy of their optimal L1 perception. When L2 inputs come in, the same L1 learning mechanism which remains accessible to a learner throughout his life will be utilized to adjust the category boundaries so as to conform to the L2 targets with the help of lexical representations, i.e. through auditory-driven category formation and lexical-driven category boundary shifting. Thus, the L1 and L2 learners can be said as not differing in the way they acquire phonology. They will finally reach the end state of L2 sound perception: be optimal in both L1 and L2 perception.

The L2LP model deals with new and similar L2 speed sound learning scenarios like the SLM, but it focuses more on the relative proximity of the boundary between speech sound categories in both L1 and L2 when L2 learning takes place. The three scenarios, the *new* scenario, the *similar* scenario and the *subset* scenario are briefly introduced here:

1) A *New* scenario: It occurs when there are more L2 phonological/phonetic categories than in the learner's L1. The L2 learner will firstly assimilate two

different L2 categories as one single L1 category. Based on the proposition that L2 speech perception utilizes the same statistical learning mechanism used in L1 perception, the learner will start to develop a new category boundary on the basis of the statistical distribution of the input. Only when the learner notices the difference between the two categories will he be able to proceed to the lexical level. The GLA will help adjust the category boundaries. The L2LP model also posits that if the new categories have dimensions that are not used in the L1, the formation of the new category will become easier. For instance, since duration is one contrastive dimension that is unused in their L1 Spanish, when Spanish learners of English are expected to perceive the difference between L2 English /i/-/ɪ/ contrast, they are encountering the *new* scenario. Among the three scenarios, the new scenario is hypothesized as the most difficult one because creation and integration of new categories are relatively more difficult than those involved in the other two scenarios.

2) A *Similar* scenario: When L1 and L2 categories are phonologically equivalent, i.e. the number of categories in L1 is the same as that in L2 along a particular dimension, the learners face a *similar* scenario. It is when slight differences are observed in the sound categorization of the L1 and L2. At first, mismatching may occur; yet, the L2 category boundaries will be adjusted to optimal locations if the leaners are exposed to ample L2 input over time with the help of lexical and context information. One example is when Canadian English speakers are learning Canadian French where both have the /æ / and /ɛ/ categories, though they are with different boundary locations, the learners have to shift the boundaries so as to achieve optimal perception. Based on the nature and number of learning tasks, the *similar* scenario is comparatively less difficult than the other two, for it involves only a single category boundary shifting task.

3) A *Subset* scenario: When the L1 perception grammar outputs more perceptual categories than the L2, making the L2 categories become a subset of the L1 categories, a *subset* scenario occurs. They have multiple-category assimilation where they will assume the existence of some phonological contrasts that in fact does not exist in the L2. In this scenario, the lexical information is crucial for the learners as it helps them reduce lexical and perceived boundaries. L1 English speakers learning Spanish as an L2 will be one example. Sometimes the Spanish /e/ will be assimilated to English /ɛ/ or /ɪ/, while Spanish /i/ will sometimes be perceived as English /ɪ/ or /i/. As the categories already exist, the L2 learners have to deploy the lexical information and conflate the two English categories into one Spanish category. This is a new scenario raised in the L2LP that has not

been considered in other previous L2 perception models. Relatively speaking, this scenario is of medium difficulty among the three.

Escudero (2005) has given empirical support to all the predications made in the three scenarios in the L2LP model. The predictions are in line with previous studies about some L2 learners' overreliance on duration when their L1s do not have this dimension (e.g. Bohn, 1995; Wang & Munro, 1999). Escudero and Boersma (2004) also found that Spanish learners of L2 Standard Southern British English (SSBE) had a different perception pattern than native SSBE listeners when perceiving /i/ and /ɪ/ since the L2 learners will map the two contrasts into their native /i/, which supports the assimilation pattern, the *new scenario*, in the L2LP model. The finding that Spanish learners of Standard Scottish English (SSE) assimilated /i/ and /ɪ/ to separate Spanish categories /i/ and /e/, which is a similar perception pattern with SSE native listeners, further confirmed the prediction of the *similar scenario* in the L2LP model. Some other studies who aimed to investigate the perceptual mapping of L2 sounds to L1 categories such as Russian listeners of American English (AE) vowels (Gilichinskaya & Strange, 2010), Peruvian Spanish (PS) listeners of AE and SSBE vowels (Escudero & Chládková, 2010) , PS listeners of Dutch vowels or English listeners (Escudero & Williams, 2011), PS listeners of /ɛ/-/æ/ contrast in Canadian English and Canadian French (Escudero & Vasiliev, 2011) or English listeners of German rounded vowels (Mayr & Escudero, 2010), all lend empirical support to the hypothesis stated in the L2LP model. Yet, this model cannot explain the perception of contrastive duration in Swedish by L1 speakers of English, Spanish and Estonian (McAllister, Flege, & Piske, 2002) or Italian consonantal length contrast by L1 German listeners who mapped the consonants with different durations to one single L1 category (Altmann, Berger, & Braun, 2012) and further studies are required to look at the complexities underlying the learners and the languages.

The L2LP model is similar to the SLM (Flege, 1987) in the way that both posit that the initial state for the L2 system begins at the L1 perception system, but is different from the SLM about the phonological space that the sounds are in: the SLM posits a single system whereas the L2LP proposes two separate grammars for L1 and L2 sounds. Thus, the SLM predicts that an L2 learner will eventually be an optimal perceiver of the entire set of L1 and L2 sounds but not an optimal perceiver of both the L1 and L2, as proposed in the L2LP model.

2.3 Linking L2 Speech Perception and Production

The SLM, the PAM, the NLM and the L2LP are influential models which put emphasis on speech perception. Yet, L2 learners expect to gain success in both perceiving and producing the phonological segments of L2 but not just in the perceptual aspect. Hence, the relationship between speech perception and production is also of high interests among linguists in the investigation of psycholinguistics and L2 speech learning. Empirical studies which have explored the relationship between the two parameters are fairly limited, and the results are usually inconsistent (e.g. Bradlow et al., 1997; Bradlow et al., 1999; Frieda, Walley, Flege & Sloane, 2000; Sheldon & Strange, 1982; Wang, 1997; Wang & Munro, 1999). However, although the direct link between speech perception and production is not clear, results of previous studies still imply a subtle link between the two parameters, and it is that draws the interest and attention of a lot of researchers.

Theories describing the relationship between speech perception and production have also been proposed, among which the revised version of the motor theory (Liberman, 1991; Liberman & Mattingly, 1985), the Direct Realist approach to speech perception (e.g. Best, 1995; Fowler, 1986) and general auditory and learning approaches to speech perception (e.g. Kingston & Diehl, 1994, 1995; Lotto, 2000; Ohala, 1996; Sussman, Fruchter, Hilbert & Sirosh, 1998) have been more influential and widely-studied. They are briefly introduced as follows.

2.3.1 The Motor Theory

Proposed by Liberman (1991), the revised version of motor theory (thereafter, the motor theory or the MT) posits that there exists a direct link between speech perception and production. Besides claiming that the link is innate and there is a common locus for speech perception and production, the MT also proposes that the objects of speech perception are articulatory events but not acoustic or auditory events. These objects are referred to as “intended phonetic gestures” which will be recovered by listeners as neuromotor commands to the articulators but not peripheral events such as actual articulatory movements (Liberman, 1991; Liberman & Mattingly, 1985). These gestures will be later references of sounds when a listener tries to encode them during perception, by the same processes involved during production. Liberman (1991) also hypothesized that speech perception cannot be attributed to general audition mechanisms and learning, but instead a speech-specific language module that consists of the phonetic system as one of the components.

The MT also claims that speech production, as a process, is linked causally by the following levels as shown in the flowchart (Diehl, Lotto & Holt, 2004):

Figure 2.2. Causal links in speech production as posited in the motor theory

One-to-one mapping correspondence is assumed to be with neuromotor commands and with muscle contractions. While the mapping between muscle contractions and vocal tract shapes is said to be highly complex due to the coarticulatory nature of speech sounds, the relation between vocal tract shapes and acoustic signals is hence non-linear (see also Stevens, 1972 on Quantal Theory of Speech), implying that speech production is a complex process mainly by virtue of the coarticulation effects of adjacent consonants and vowels.

Apart from the MT, another approach that posits the existence of a direct link between speech perception and production is the direct realist approach to speech perception or more specifically the Direct Realist Theory, which however differs from the MT in the perceptual primitives and mechanisms (Best, 1995; Wang, 2002).

2.3.2 THE DIRECT REALIST APPROACH TO SPEECH PERCEPTION

Direct Realism is originally a philosophical theory developed by Gibson (1966, 1979) and Gibson (1969, 1991) regarding the origins of perceptual knowledge. It claims that the perceiver directly apprehends the external world and the perceptual object directly and with awareness. Built on empirical research on perception, the fundamental premises of this philosophy have been applied to speech perception and thus the Direct Realist view to speech perception was developed (Best, 1984, 1993, 1994a, 1994b; Fowler, 1986, 1989, 1990). This approach to speech perception posits that the actual gestures produced by the speaker's vocal tract are directly perceived. The gestural knowledge is available in any speech for a perceiver to detect directly. Note that gestural information is not founded on acoustic features and that neural and/or cognitive mechanisms are not needed to decode inferences from speech signals. When it comes to production, the speaker aims to produce the targets based on the perceived gestures. It means that the link between speech perception and production lies in the "common communicative goal" of the language user (Bradlow et al., 1997, p.2299).

Also, perceptual learning plays a crucial role in the Direct Realist view to perception, especially when Direct Realist Model (DRT) is considered. The DRT assumes that the detection and discovery of higher-order invariants are entailed in the perceivers' attunement to native speech. The invariants are specifications of the gestural knowledge perceived, which together make up the native phonological inventory (Best, 1995). Inasmuch as perceivers become language-specific, i.e. they become attuned to their native speech, detecting crucial elements that distinguish the linguistic relevant contrast is said to be efficient. This suggests that experience with the native language will affect perception of non-native speech.

Besides the difference in the objects of perceptual gestures (intended in the MT, actual in direct realist approach), the MT and the Direct Realist approach distinguish most from each other is that no neural innate module is assumed to mediate the link between speech perception and production in the Direct Realist view. However, based on a host of studies (e.g. Diehl & Kluender, 1987; Fowler, 1986; Liberman, 1996), there is no serious disagreement among theorists that both the MT and the Direct Realist approach assure that "regularities of speech production… [is] highly correlated with listener's perceptual judgments" (Diehl et al., 2004, p.167)

Both the MT and the Direct Realist approach to speech perception however received challenges because of some new empirical findings (e.g. Miller, Wier, Pastore, Kelly & Dooling, 1976; Pisoni, 1977; Stevens & Klatt, 1974) which had led to the exploration of alternatives to the MT and the Direct Realist approach. It is referred to as the general approach to speech perception and is introduced in the next section.

2.3.3 The General Approach to Speech Perception

The general approach (GA) to speech perception was proposed based on a number of speech investigations (e.g. Kingston & Diehl, 1994; Lotto, 2000; Ohala, 1996; Sussman et al., 1998). The GA, in contrast to the MT, does not entail any specialized or human-specific mechanisms or modules to explain speech perception. Instead, speech perception is hypothesized as involving the use of the same mechanisms of audition and perceptual learning to any sounds from the environment. Also, the GA does not assume the recovery of speech from the acoustic signals as mediated by the perception of gestures (Diehl et al., 2004). The GA is also regarded as an approach rather than a theory owing to its abstract nature in defining itself mainly as an opposition to the MT and the direct realist approach. It also differs from the MT and the Direct Realist approach in terms of how speech perception and production are linked.

The GA posits two general accounts for the connection between speech perception and production. They are a) production follows perception, and b) perception follows production. For the first account, the GA explains that productions can be made only when the auditory distinctiveness of phonemes is maximized to promote intelligibility of a sound, and hence the entire sound system. Thus, when a listener can perceive a sound that is acoustically and so auditorily most distant from other sounds, productions can be made. The second claim of the GA proposed that during speech production, any regularity in the production such as context dependencies will be reflected in the acoustical signal which a listener can make use of during general mechanisms of speech perception. Thus, as posited in the GA, since a listener perceives only the acoustical consequences of articulatory gestures, he/she can correlate those production regularities that have been perceived as acoustic signals in the speech perception.

The above two sections have introduced several current models, theories and approaches that explain speech perception and/or production. Over the past several decades, evidence for a link between speech perception and production has begun to surface in some training studies carried out in laboratories. Some of them are perceptual training studies on the perception and production of non-native contrasts which pose great difficulties on the L2 learners (e.g. Bradlow et al., 1997; Hazan, Sennema, Iba & Faulkner, 2005; Iverson & Evans, 2009; Lambacher et al., 2005). These cross-language studies and training experiments will be introduced in the following sections as reinforcement of the aforementioned theories and to set the background for the current research study.

2.4 Review of L2 Perceptual Training in the Laboratory

Difficulties in L2 speech perception and production among adult learners are well-known, as the retuning of the phonetic system of these language-specific learners is usually more effortful and challenging (e.g. Best, 1995; Iverson et al., 2003; MacKain, Best & Strange, 1981). Hypotheses and theories have also been proposed for explaining these immense difficulties in L2 speech learning, as stated in Section 2.1. Nonetheless, some research (e.g. Bradlow et al., 1997; Hazan et al., 2005; Iverson & Evans, 2009; Lambacher et al., 2005) conducted in the last few decades have shown that L2 speech learning difficulties can be ameliorated by intensive laboratory training. These studies have shown that the perception and production of L2 sounds can be modified through a short period of intensive laboratory training while the learning can also be generalized to new words that the subjects have not encountered during training and to tokens that are produced by new speakers whose voices have also not been heard during the training. A long-term retention effect was also reported. The results of these training studies not only give a new way out for learners who cannot afford the time for long immersion in the ambient language environment to improve their L2 speech learning, but they also shed light on how malleable and plastic adults' speech learning systems can be. The following sections will discuss some methodologies and variables that are more commonly adopted or concerned in cross-language training studies, and present the results of some studies that will shed light on the current study. One of the variables, training intensity, which has been always assumed as a constant in previous training studies, will also be scrutinized in particular to set the stage for the present investigation.

2.4.1 Methodologies Adopted in Previous Training Studies

Researchers have designed different methodologies and set different target variables to encapsulate the relationship between the choice and design of tasks and the efficacy of the training method. Interactions and combinations of training task types, stimulus variability and training duration have been tested under laboratory conditions in some studies (e.g. Bradlow et al., 1997; Hazan et al., 2005; Iverson & Evans, 2009; Jamieson & Morosan ,1986; Lambacher et al., 2005; Logan & Pruitt, 1995; Wong, 2010, 2012, 2013a, 2013b). Issues addressing these factors will be introduced in the following parts. One often ignored factor, training intensity, will also be introduced to prepare the ground for the current research.

2.4.1.1 Task Types: Discrimination vs. Identification Training

Choosing appropriate methodologies to adopt in perceptual training studies is crucial. A fundamental distinction in training studies in literature is usually between the adoption of discrimination and identification tasks. In discrimination tasks, subjects have to determine whether the stimuli are the same or different in a paradigm, which can be with the AX paradigm (to decide whether two tokens, A and X, are the same or different) or the ABX paradigm (to decide whether X is the same as A or B, also called an *oddity task*). While in identification training, subjects will hear only one token in one trial. They have to identify the stimulus presented to them with the respective label or choice given. Immediate feedback is usually provided in this type of perceptual training. Thus, discrimination task is often used in infant speech perception studies because a listener must have the knowledge of the language well to respond to the label in an identification task but discrimination task requires no knowledge of the language.

The earliest perceptual training studies under laboratory conditions (e.g. Carney, Widin & Viemeister, 1977; McClaskey, Pisoni & Carrell, 1983; Pisoni, Aslin, Perey & Hennessy, 1982; Strange & Dittmann, 1986) adopted the use of discrimination tasks or a combination of discrimination and identification tasks in training. Carney et al. (1977) and Pisoni et al. (1982) found that American English speakers could successfully discriminate within-category differences along the English voice onset time (VOT) /b/–/p/–/p^h/ continuum after a short period of training with discrimination and identification tasks. McClaskey et al. (1983) replicated and extended the study of Pisoni et al. (1982) and indicated that even no further training was provided, the subjects can transfer the perceptual learning of a third voicing category to a new place of articulation. Although these three studies investigated only the perceptual improvement of monolingual speakers' L1 categories under laboratory conditions, the positive results motivated Strange and Dittmann (1984) to start adopting a similar protocol to evaluate the effect of perceptual learning on non-native sounds.

Strange and Dittmann (1984) gauged the production of /ɹ/ and /l/ contrast among a group of adult Japanese speakers of English with varied L2 experience. The training adopted the use of AX discrimination on a synthesized /ɹ/ and /l/ continuum (by using "rock" vs. "lock") with immediate feedback. The subjects also had to participate in a pretest and a posttest with natural tokens contrasting the /ɹ/ and /l/ pairs as well as identification and discrimination tasks on two synthetic /ɹ/ and /l/ continuum, the "rock-lock" series and a new "rake-lake" series. Learning was found to have transferred to the new continuum and more demanding ABX discrimination tasks, but not to naturally produced words.

Strange and Dittmann's (1984) investigation, besides being the pioneering research in training cross-language perceptual performance, also marked the change of research from using both discrimination and identification tasks to the sole adoption of identification tasks in the training paradigm. The lack of transfer of learning to natural tokens in the study was said to be partly attributed to the solely adoption of discrimination tasks. It is claimed that discrimination tasks cannot promote effective and generalizable perceptual learning and these tasks do not provide an optimal training ground for inducing changes in phonetic categorization (Logan & Pruitt, 1995). Also, discrimination training tends to draw the subjects' attention to within-category differences and hence the subjects may ignore essential acoustical properties for category formation (Jamieson & Morosan, 1986, 1989). It is probably the reason for the decrease in popularity of discrimination training. In contrast, identification tasks are more widely adopted in more recent training studies (e.g. Bradlow et al., 1997; Lively et al. 1993; Lively et al., 1994; Logan et al., 1991; Logan & Pruitt, 1995) because they were believed to be more effective in significantly improving the subjects' perceptual performance in non-native phonemic contrasts, especially when the stimuli used in these tasks have some token and speaker variability. Nonetheless, the above studies which focused on binary non-native contrasts seemed to have ignored the possibility that giving the subjects only two-alternative forced choices would yield a chance level as high as 50% for the subjects. Since the subjects were required to choose their answer out of the given two choices, even if they could not identify the stimuli, they still had to give an answer which could merely be based on speculation. Hence, the results may not be reflecting the genuine perceptual ability of the subjects. This possible shortcoming in design was, however, overlooked and not mentioned as a limitation. Recent vowel training studies (e.g. Lambacher et al., 2005) tended to use multiple-choice identification format in which they trained the participants a set of vowels which were not binary contrasts. This may be one adoption of identification task that could yield more convincing and true results of the subjects.

Although only limited evidence has shown that identification tasks are superior to discrimination training (Flege, 1995), identification tasks with high stimulus variability, named in previous literature as the High Variability Phonetic Training (HVPT) approach, started to become the preferred and dominant phonetic training approach in the past two decades (Bradlow et al., 1997; Bradlow et al., 1999; Hazan et al., 2005; Lambacher et al., 2005; Lively et al., 1993; Lively et al., 1994; Logan et al., 1991; Iverson & Evans, 2007a, 2009; Nishi & Kewley-Port, 2007b, 2008). The details and relevant studies will be explicated in the Section 2.5.2. Before proceeding to how HVPT has been administered in previous research, how stimulus variability and the development of the adoption of stimuli have been taken as a crucial factor that influences the training effectiveness will be presented first.

2.4.1.2 Training Stimuli: Nature and Variability

Training stimuli used in previous studies can be divided into two types: the nature and the variability of the stimuli. Concerning the nature of the training stimuli, both synthetic (sound tokens manipulated by the researcher using digital software) and the naturally produced tokens, are both common in perceptual training studies. For naturally-produced tokens, variability can be set to enhance the training effect.

Synthetic tokens allow the researchers to manipulate the properties of the sounds, in particular the vowel qualities or acoustic properties, so as to suit the research purpose. Perceptual training research in the early 1970s to 1990s tended to make use of this advantage to investigate the perception of sounds along a continuum. This is an aforementioned technique called the fading technique which was well-adopted in some perception training studies (e.g. Jamieson and Morosan, 1986, 1989). With synthetic stimuli, the researchers could present these tokens even with extreme acoustic or phonetic properties that may not exist or be noted in the real life to the participants, aiming to train the attention of participants towards more accurate category distinction even for ambiguous sounds in the fussy area along the manipulated continuum. Albeit the freely-altered nature of the synthetic stimuli, they are still not genuine tokens that represent the real-world utterances with the auditory properties in the ambient sounds. It may not help the accomplishment of the common goal of linguists and teachers as learners should attend to more natural tokens that can enable them to transfer their learning to real life applications. Hence, natural tokens have also been adopted in some studies to achieve this purpose.

Using natural tokens produced by just one single speaker or by multiple speakers in a training program is common in the training literature. With more talker variability, subjects can be exposed to a wider range of inter-speaker differences to promote perceptual learning. Added to this is the variability in phonetic environments. In daily connected speech, subjects are not only perceiving or producing a set of tokens controlled within confined environments. A more extensive use of stimuli with various contexts is thus believed to promote learning and generalization, for it resembles real-world communication with the native speakers.

The concept of using higher variability stimuli in talkers and phonetic contexts has been widely accepted and adopted in some training studies in recent decades of which the approach is named High Variability Phonetic Training (HVPT), as opposed to the studies dated back which involved the use of stimuli produced by a single speaker and with limited phonetic environments. Studies adopting HVPT approach will be introduced in Section 2.5.3.

2.4.1.3 Assessment Design

To assess the effect of perceptual training on participants' perception performance, researchers usually adopt a pretest-training-posttest design to compare the results before and after the training by measuring their perception accuracy or response time. If the study also investigates the transfer of perception effect to production, measuring the production performance before and after the training by using acoustic measurements or native speakers' judgment is also a common assessment method adopted in training studies. Nonetheless, most of the studies tested the transfer of perceptual learning to production but not vice versa. Pursuing this research direction, so as to draw more valid conclusion that the change of subject performance is attributed only to the training alone but not the test-wise effect, a control group without participating in the training and has other confounding variables avoided is always setup. If the link between perception and production is to be investigated by investigating the training effect of one mode to the other, a more controlled setting is necessary.

Besides the importance of training effect, one of the key issues addressed in perception training studies is the external validity, which concerns whether the training effect can be transferred to the sounds in new phonetic contexts. Thus, perception generalization tests are usually adopted. However, the generalization effect on the production of sounds – whether learning can extend to sentence level production or even connected speech – is usually ignored. I administered a generalization test for production called Test of Contextualization in which participants would record a passage with target sounds simulating more natural speech in a vowel perception and production training study (Wong, 2010, 2012, 2013a, 2013b). Results showed that the training effect in general can be transferred to contextualized situations. Added to this is the investigation of long-term retention of the training effects. This is important to both linguists and teachers, for the purpose of training is to offer solid and long-term learning to the subjects. Thus, some researchers give a delayed-posttest to participants after three or six months of the posttest to judge the effectiveness of the training in the long-term development of perception and production strategies and skills.

2.4.1.4 A Forgotten Variable: Training Intensity

Training intensity has usually been preset as a constant in many previous studies (e.g. Bradlow et al., 1997; Bradlow et al., 1999; Hazan et al., 2005; Lambacher et al., 2005; Lively et al., 1993; Lively et al., 1994; Logan et al., 1991; Iverson & Evans, 2007a, 2009; Nishi & Kewley-Port, 2007b, 2008), hence has been overlooked at how intensive a treatment should be so as to achieve optimal results, or whether different intensity levels will generate diverse training effects. Proactive as this factor is, training intensity is one intervention design which determines the subjects' rate of progress, although the tangible effect on performance depends more on the subjects' motivation and level of interest due to different configurations of intensity of training tasks (Warren, Fey & Yoder, 2007).

A ubiquitous consensus has been drawn in previous studies that an intensive perceptual training treatment per se can contribute to a significant perceptual and production gain (Lengeris, 2009; Wang, 2002). Nevertheless, it is also argued and shown in some previous studies (Bambara & Warren, 1993; Riches, Tomasello, Conti-Ramsden, 2005; Warren, Fey & Yoder, 2007) that efficacious learning is achieved when the same number of training episodes is spread over several sessions than massed into one or a few. Despite this controversy, no second language acquisition or training studies on speech perception and production have systematically investigated the effect of differential training intensity levels on subjects' performance, thus ignoring the possible link to any reachable optimality effect.

Only in recent decades has training intensity begun to receive some attention in research (e.g. Fey, Warren, Brady, Finestack, Bredin-Oja, Fairchild, Sokol & Yoder 2006; Gray, 2003, 2005; Yoder and Warren, 2001, 2002) on the language development of young children with developmental delays and disabilities. Although there are no standard definitions of training intensity in the field, it is in itself an intriguing parameter that is viewed quite differently in the literature, varying from the researcher-subject ratio, the level of subject participation or both the quality and quantity of intervention given in a specified period, to simply the number of hours of treatment in a given period of time. All these results are assets that enrich researchers' understanding towards the effect of this time factor on subject performance, further implying the value of studying the effects of differential training intensities.

Warren, Fey and Yoder (2007) proposed a set of terminologies for measuring treatment intensity and they are presented in the following table:

Table 2.1. Terminologies of Training Intensity Based on Warren, Fey and Yoder (2007)

Terminologies	**Definition**	**Example**
Dose	The number of properly administrated training stimuli presentations during a single training session, including the average rate of training stimuli presentation per unit time, the length of the training session and the distribution over the session (this term will	A dose of 60 training stimuli per one hour training session

	be further explicated in the sixth point)	
Dose form	The physical form in which the active ingredient (i.e. the training stimuli) is dispensed	Imitative prompts: subjects imitate the researcher's prompts
Dose frequency	The number of times a dose of training is provided per day and per week	One week one hour dose at one stimuli per minute
Total Intervention Duration	The time period over which a specified intervention is presented	10 weeks
Cumulative intervention intensity	The product of *Dose* × *Dose Frequency* × *Total Intervention Duration*	60 × 1 × 10 = 600
Training episode distribution	How the number of teaching trials is, i.e. whether they are massed into just a few of sessions or distributed over a number of them	Distributed

The understanding of training intensity of the present study will be built upon this framework. In spite of the fact that the terminologies are based on medical models and are devised for research studying particularly human language disorders, they still offer the present study a solid ground for how training intensity can be primarily understood and defined. With a view to better suiting the present study on second language perceptual training, the definitions of the terminologies are fine-tuned.

It is suggested that cumulative intervention intensity can be held constant while only other aspects should vary (Warren et. al, 2007). This can help preliminarily give some useful information on the optimal configuration of the training effectiveness by investigating other parameters. Pursuing the definition of training intensity set under these terminologies, the present study will follow the same line of reasoning and investigate the differences between two training intensity levels (they differ in the dose frequency, training episode distribution and hence total intervention duration while the cumulative intervention intensity is kept the same). The detailed administration of this variable in the current research will be illustrated in Chapter 3 under Section 3.1.

2.4.2 L2 Perceptual Training Research

2.4.2.1 High Variability Phonetic Training (HVPT)

The failure of transferring perceptual learning to natural tokens in Strange and Dittmann's (1984) study is speculated as related to the use of stimuli produced by only a single speaker and in a single context, which is the Low Variability Phonetic Training approach (LVPT), in contrast to HVPT technique. Although this study used discrimination tasks in lieu of identification tasks with low-variability stimuli, one later study (McCandliss, Fiez, Protopapas, Conway & McClellang, 2002) using LVPT and identification training still showed that low-variability stimuli, particularly those synthetic ones, could not induce effective training effects on novel and natural tokens. These results however motivated a battery of follow-up research studies in response to the suggestion pointed out by Strange and Dittmann (1984) and Jamieson and Morosan (1986, 1989) that a wider range of stimuli should be covered in the training paradigm, and they are the studies adopting the High Variability Phonetic Training (HVPT) approach.

HVPT emphasizes the use of multiple speakers and various phonetic contexts to increase the stimulus variability in the presentation of the natural training minimal pairs. The subjects under training can then be exposed to the natural tokens produced by different speakers. Results utilizing HVPT approach have displayed encouraging results in improving the subjects’ perceptual performance of confusing non-native contrasts.

HVPT was firstly adopted in a series of research studies in researching the effectiveness of stimulus variability on the perceptual training of /ɹ/ and /l/ contrast of Japanese speakers of English (Bradlow et al., 1997; Bradlow et al., 1999; Lively et al., 1993; Lively et al., 1994; Logan et al., 1991). *High variability* in the approach designated the use of a wider variety of speakers, and hence the stimuli of the target language to train the perception of the subjects. This approach has particularly been popular in training Japanese speakers’ perceptual learning of English /ɹ/ and /l/ contrast in the last two decades and yielded very salient improvements among the subjects.

Logan et al. (1991), as the first of the series, trained six Japanese speakers of English with different length of residence in the United States to identify the /ɹ/ and /l/ contrast under different phonetic environments. The subjects had to do a pretest and a posttest, in which the testing tokens were the same 16 minimal pairs used in Strange and Dittmann’s (1984) research. Fifteen training sessions using a two-alternative forced-choice identification task with immediate feedback were provided to the subjects. The subjects were exposed to 68 minimal pairs with /ɹ/ and /l/ contrast in different environments in the training. The minimal pairs were produced by five different native speakers and were presented as a block of tokens under the speaker, meaning that the subjects would hear a set of tokens produced by one speaker first before another. Besides doing the posttest, the subjects also had to do the first Test of Generalization (new words produced by a new speaker) and the second Test of Generalization (new words produced by a familiar speaker). The results showed that the identification improved from the pretest to the posttest in general and transfer to the two tests of generalization was observed. However, the effect of phonetic environments and the effect of speakers were found as tokens produced by some speakers were more accurately identified due to the grouping of tokens under a speaker, implying that the subjects had paid attention to the talker-specific information rather than the acoustic differences of the two consonants. This study was also criticized by Pruitt (1993) that there was no control group for comparison of experiment effects and hence lacked evidence to claim the efficacy of using HVPT over LVPT and the usefulness of stimulus variability.

The results of this first study were however significant in the way that it provided a solid threshold for further research. On the grounds of Pruitt’s (1993) critique on the methodologies and research design of the study of Logan et al. (1991), two more follow-up studies, one investigating the role of phonetic environments on perception (Lively et al., 1993) and the other testing the possibility of long-term retention of HVPT (Lively et al., 1994), were conducted. Lively et al. (1993) modified the methodology used in Logan et al. (1991) by doing two experiments to investigate the effects of talker variability and phonetic environments separately on the subjects’ perceptual learning. The first experiment trained native Japanese speakers’ perception by providing the subjects with natural tokens with English /ɹ/ and /l/ contrasts in three different environments (initial singleton, intervocalic positions and initial consonant clusters) produced by five native speakers. The second experiment used minimal pairs produced in five different environments by only one single speaker. Results showed that subjects in general improved from the pretest to the posttest in both experiments, but the tokens with higher speaker variability (experiment 1) were more efficient in training the subjects than using tokens with higher variability in phonetic environments (experiment 2), especially in generalization to new words produced by new speakers. This study provided more evidence that higher speaker variability was essential in cross-language perceptual learning, especially in generalization.

The third study of the series by Lively et al. (1994) tested whether the training effects can last long among a group of native Japanese speakers with no experience in residing in any English-speaking countries. The study was closely based on the methodology used in Logan et al. (1991), using highly variable training stimuli produced by five native speakers contrasting /ɹ/ and /l/ in four different phonetic environments. The posttest results showed that the subjects improved significantly and more generalization effects were found in new words produced by familiar speakers than by new speakers. The primary goal of the experiment was to gauge the effect of long-term retention of the perceptual learning by administering two delayed posttests three and six months after the training, the posttest and generalization tests were all completed. Significant improvement was still found in the three-month posttest and was partially retained in the six-month posttest.

The other two training studies in the series of HVPT approach in training Japanese speakers of English the perception of /ɹ/ and /l/ contrast (Bradlow et al., 1997; Bradlow et al., 1999) will be discussed in detail in Section 2.5.5 as these studies have expanded the research design and probed into the transfer of perceptual learning to production, which is a valuable area in the present study that is worth devoting a separate section for discussion. Yet, since HVPT approach was found to be so useful in improving the notoriously difficult object of perception, the English /ɹ/ and /l/ contrast, among native Japanese speakers, this training paradigm was also adopted in other cross-language studies aiming to improve the perception of other segmental and suprasegmental L2 contrasts. These include training native Mandarin speakers on the perception of French voiced and voiceless stops which were distinguished by VOT values (Rochet, 1995) or their perception of English word-final /t/ and /d/ (Flege, 1995b); training native English and Japanese speakers to perceive Hindi dental and retroflex stops (Pruitt, 1995; Pruitt, Jenkins & Strange, 2006); training native English speakers the perception of Korean stop voicing contrast (Kim & Hazan, 2010); training native English speakers the perception of Japanese vowel length contrast (Hirata, 2004; Hirata, Whitehurst & Cullings, 2007; Tajima, Kato, Rothwell, Akahane-Yamada & Munhall, 2008; Yamada, Yamada & Strange, 1996); training native English speakers the perception of Mandarin lexical tones (Wang, Jongman & Sereno, 2003; Wang, Spence, Jongman & Sereno, 1999) as well as training Catalan or Spanish speakers the perception of English word-initial /p/-/b/ and /t/-/d/ and the vowel contrasts /iː/-/ɪ/ and /æ/-/ʌ/ (Aliaga-García & Mora, 2009), just to mention a few.

2.4.2.2 Vowel Training Studies

Research on cross-language vowel training is however relatively scarce vis-à-vis research on L2 consonant training. Although there is an increasing number of investigations on cross-language vowel perception such as the perception of English vowels by Greek speakers (Lengeris, 2009; Podlipsky, 2005) or by German and Spanish participants (Iverson & Evans, 2007, 2009), most of the studies have tended to focus on investigating native Japanese speakers' perception (and production) of American English vowels (e.g. Lambacher et al., 2005; Nishi & Kewley-Port, 2005, 2007b; Sperbeck, Strange & Ito, 2005). Several cross-language vowel studies, with English as the subjects' L2, will be introduced in this section as they are insightful for the present research. They may not have adopted HVPT approach, but the research methodology was modeled or based on the design of HVPT paradigm.

Seventeen adult native speakers of Japanese were recruited in Nishi and Kewley-Port's (2007b) study, which aimed to compare the efficacy of two sets of stimuli in perceptual training: one was a full set with nine American English monophthongs and the other was only a subset with three most difficult American English monophthongs. Six of the subjects were assigned to the training group with the full set stimuli; the other six were put in the training group with only the subset; and the remaining five subjects were the control subjects receiving no training. The experiment followed the usual protocol of HVPT: the pretest, training, the posttest, and a three-month delayed posttest. In both the pretest and the posttest, the subjects were given nonsense words with consonant-vowel-consonant and schwa (CVCə) structure within six different consonantal contexts. Thirty-six real CVC words were used in the generalization test. The training, including nine sessions, adopted identification tasks with immediate feedback and the stimuli were those nonsense words used in the pretest and the posttest. The subjects were given at most 10 chances to re-listen to the correct answers if their identification was incorrect. The results indicated that both training groups improved their perception of the vowels from the pretest to the posttest, and their learning could also be generalized to new words produced by new speakers. However, the subjects trained under the subset were not able to improve the perception of vowels other than the three they had received training. This suggested that the use of the full set stimuli was more advantageous and beneficial than the subset. The authors explained that because the subjects trained under the full set stimuli could be exposed to a larger set of vowels, they could hence experience a wider range of spectral and temporal variability, which allowed them to be more successful learners than those with limited stimulus exposure.

Nishi and Kewley-Port (2008) conducted a follow-up study based on Nishi and Kewley-Port (2007b). They trained three groups of five Korean speakers for nine days on nine American English monophthongs. Three training protocols were compared: a full set of training stimuli used in Nishi and Kewley-Port (2007b), first three days on subset and six days on full set (3V-9V protocol), and lastly, first six days on full set and three days on subset (9V-3V protocol). The procedures in the training were the same as in Nishi and Kewley-Port’s (2007b) study. The performance of the subjects was assessed by a pretest, a posttest and a mid-training test and the results showed that all of the training protocols were effective in improving the subjects’ perceptual learning of the English vowels, but the two protocols involving the subsets were not found to have any advantage over the full set protocol. The results of 9V-3V and 3V-9V protocols were however very different. It was found that training first on the smaller set (3V-9V protocol) exerted strong and unexpected negative effects on the learning of the vowel set. As shown in the results, this protocol failed to train the subjects on vowel /ʊ/ in particular. The authors elucidated that because the subjects were guided initially to make label-exemplar associations for three particular vowels and when they were exposed to the whole set of vowels afterwards, they had to accommodate additional vowels while their newly learnt subset was still unstable and in a state of complication. This study had corroborated the general positive effects of HVPT and it has also shed light on how the order of training protocols and size of training stimuli set affect the effects of perceptual learning.

The last cross-language vowel training study that will be introduced in this section is the one conducted by Iverson and Evans (2009) who trained 17 Spanish and 16 German native speakers on their perception of 14 British English monophthongs and diphthongs under HVPT paradigm. The training was partly tailor-made to train the subjects on vowels that they had problems with. The performance of the subjects was evaluated by comparing their results in the pretest and the posttest in three ways: a) a vowel identification test, where the subjects chose the word they heard from four choices; b) an L1 vowel assimilation test, where the subjects chose the L1 vowel that sounded the closest to the stimulus they heard; and c) a vowel-space mapping test, where the subjects had to rate on a continuous scale how far away a stimulus they heard was from being a good exemplar of the words given. The results indicated that both groups improved their perceptual performance, while the German group improved more than the Spanish group. It was only after 10 more sessions of training that the Spanish speakers' performance could reach the level that the German group had achieved. The performance of both groups was also retained after four to five months. The authors concluded that a larger L1 vowel category inventory (German has 18 vowels whereas Spanish has only five vowels) seemed to facilitate new learning. Also, the results suggested that HVPT promoted perceptual learning by allowing the subjects to apply their knowledge of existing phonetic categories to identify new L2 vowels more accurately and efficiently.

All these vowel training studies have attested that HVPT approach is an effective approach in training the perception of L2 vowels. Modifications and the application of the approach to subjects with different L2 backgrounds were also attested and the results showed that the effectiveness was generally high, but it varies in different degree. Practically speaking, the eventual goal of learning the L2 sound system is to achieve the success in both the perception and production of the sounds in the target language, meaning that perceptual learning should be evaluated as how its effects can be transferred also to the production domain which can also help understand the link between perception and production as well as the second language acquisition processes. The studies which examined the effect of perceptual training on production are worth mentioning in the next section for understanding more about how these two domains can be assessed.

2.4.2.3 The Effects of Training Paradigm Design on Perceptual Learning

Studies mentioned in the above two sections mainly gauged the effectiveness of a training paradigm on non-native segmental contrasts. One study that aims to compare the effects of different training paradigms, including HVPT, on the learning of pitch is worth mentioning here to conclude major studies on perceptual learning before moving on to the studies that evaluate also the learning in the production domain. This study provides more empirical evidence of, on one hand, the possibility of training contrasts other than in the segmental aspect, and on the other, the importance of comparing training paradigms to maximize the training effects and design, which is of high relevance to the present research.

This study was conducted by Perrachione, Lee, Ha and Wong (2011) and had two parts. In the first experiment, the authors recruited 64 native speakers of American English to learn a phonological contrast, i.e. a lexical tone, to distinguish 18 simulated words. The primary goal of the experiment was to assess whether the amount of stimulus variability in the training paradigm would influence the learning results of the subjects with highly variable pretraining pitch-perception abilities. After assessing the perceptual abilities of the subjects, they were divided in to two groups: High-Aptitude Learners (HAL) and Low- Aptitude Learners (LAL). Half the subjects in each group were assigned to either HVPT (the authors called it the HV condition, with stimuli produced by four training speakers, intermixed within and between training sets) or LVPT (the authors called it the LV condition, with stimuli produced by only one of the four training speakers). All the test and training stimuli were pseudo words, yet naturally produced with synthetically manipulated pitch contours. Eight training sessions in which all the subjects were given 72 training trials per session were given to them, followed by a daily Word Identification Test (WIT) which aimed to track the learning progress. Their performance was measured by a test called Test of Learning Achievement (TLA) which involved the same tokens produced by four generalization speakers whose voice had not appeared in the training. With respect to the learning progress assessed by WIT, the results showed that the HAL group demonstrated faster learning than the LAL group when trained under both HVPT and LVPT, and faster learning in both groups was observed in the LVPT conditions. In TLA, however, the LAL group had a significant impairment in learning under HVPT, whereas the HAL group still outperformed the LAL group in both training types. By comparing the results in TLA with the WIT on the last day of training, the generalization effect was also measured. No differences were observed between the HAL and LAL groups in terms of performance, but both groups who received HVPT indicated better generalization to new talkers than those who received LVPT.

The impaired outcomes within the LAL group trained under HVPT, of which the authors dubbed as an observation of “individual-instructional interaction” (p.466), attracted particular attention. The authors justified that the observation was due to the difficulty and extraneous cognitive load that the low-aptitude learners encountered while listening to stimuli, trial-by-trial, with different voices. In means that whether a training paradigm is beneficial to the learners depends on the learners’ perceptual abilities. Meanwhile, the amount of time that subjects were only exposed to individual stimulus token produced by a specific talker in HVPT were obviously less than those in LVPT. It is claimed that the LAL subjects may require more extensive exposure of individual stimulus before they could achieve successful learning. These two factors were said to interact and made HVPT detrimental to LAL groups. A battery of follow-up experiments was carried out by the authors to further investigate variations in HVPT paradigm.

Sixty-one new participants were invited for a series of experiments which compared three variations manipulated on an ordinary HVPT paradigm: 1) Block high variability (HV-B); 2) Repeated high variability (HV-R); and 3) Blocked and repeated high variability (HV-BR). All the training stimuli (18 pseudoword stimuli) and procedures were the same as those in the experiment just mentioned. In HV-B, the stimuli were no longer randomized by talker; instead, they were blocked by talker, i.e. they heard 18 token produced by Talker A first, then 18 by Talker B, and so on. Subjects hence heard only one voice in one training block with different stimulus tokens. In HV-R, subjects were exposed to the same amount of tokens produced by the same voice as in LVPT condition (each word was repeated by a single talker four times), by having them hear four times tokens than before, i.e. 18 stimuli x 4 times = 72 tokens per talker. Yet, the stimuli were still intermixed to retain the trial-by-trial variability. The last HV-BR training condition combined the use of HV-B and HV-R by blocking all the tokens under the same talker in one training block and all the tokens were repeated four times.

Results demonstrated more rapid learning in the HAL group than the LAL group across conditions. Blocking and repeating the stimuli both resulted in more rapid mastery of the test items in both groups. However, the combining effect of both manipulations did not generate faster learning. The TLA scores again showed that the HAL group outperformed the LAL group across training conditions. Also, blocking the stimuli resulted in greater learning than not blocking the stimuli or repeating the stimuli, particularly for the LAL group as they could reduce the cognitive load across trials. The HAL group did not benefit from the blocking manipulation. Repeating the stimuli, however, was not at all useful to the learning achievement for both groups. These three HV paradigms were found equally effective for both groups when the generalization effect was concerned. The authors concluded that a manipulated version of HVPT, i.e. HV-B by removing the trail-by-trial variability that imposes excessive processing demands, can benefit the learners with lower perceptual aptitude. HVPT which adopts stimuli produced by different talkers was in general beneficial to learners with different cognitive and perceptual abilities and learning could also be generalized to novel stimuli.

2.4.2.4 The Effects of Perceptual Training on Production

Although it has been extensively researched, the link between perception and production has only been investigated mainly in terms of cross-sectional studies on the subjects' performance of consonants (e.g. Flege & Schmidt, 1995; Schmidt & Flege, 1995) and vowels (e.g. Flege, Bohn & Jang, 1997; McAllister et al., 2002) in both domains which overlooked the fact that improvement in production as a function of perceptual learning can also offer insights to the correlation between the two modes. Studies that directly assessed how changes in the perceptual domain affected the production domain under laboratory training have emerged in recent years and some investigating English as the L2 of the subjects are reported as follows.

The series of studies that focused on the perceptual learning of English /ɹ/-/l/ contrast by Japanese speakers was extended by Bradlow et al. (1997) to the transfer of perceptual learning to the production performance of the subjects. Under HVPT protocol, eleven Japanese subjects had to go through the pretest-treatment-posttest procedure and were pretested and posttested on both their perception and production performance. A group of control subjects was also present. The perception tests and training procedures followed the usual identification tasks (with immediate feedback in training sessions) while the production tests required the subjects to record a list of minimal pairs contrasting /ɹ/ and /l/. The production performance was assessed by a group of native English speakers in two ways: one by seeing whether the native speakers preferred the production of a word produced by the subjects in the pretest or the posttest; the other one by testing whether the native speakers can identify the subjects' production tokens in the pretest and the posttest accordingly. The results demonstrated that not only had the perceptual performance improved, corroborating positive results in previous relevant studies, but it also showed that the production performance also improved. The authors claimed that the effect of perceptual learning had brought about gains in the production domain. Despite the general gain across subjects, a lot of individual differences were found among them and this implied that the effects of the training vary across individuals. Bradlow et al. (1999) extended the above research by testing whether the improvements in both domains could be retained three months after the training. This study replicated the findings in Bradlow et al. (1997) and also revealed that HVPT approach was beneficial in retaining long-term learning in both perception and production of non-native contrasts. The authors also speculated that the two domains were closely linked. A similar study done by Hazan et al. (2005) which measured the effects of audiovisual training on the Japanese speakers' perception and production of English /ɹ/ and /l/ contrast, also yielded similar positive results in the attainment in both domains.

Besides examining consonants, some studies investigated the effect of L2 vowel perception training on production. Wang (2002) investigated the perception of three English vowel pairs, /i/-/ɪ/, /e/-/æ/, and /ʊ/-/u/, among a group of Mandarin and Cantonese speakers and the effect of perceptual learning on production. The subjects' production performance was assessed by a group of native speakers and also acoustic analysis of the words produced by the subjects in the pretest and the posttest. Trained under both synthetic stimuli and HVPT paradigm, the subjects were found to have significantly improved the perceptual identification of the three vowel contrasts, but the production improvement was not significant. There were, however, a lot of individual differences which provided evidence that some particular subjects' production performance had improved. The author justified that the difficulty was due to the fact that production of vowels lacked a fixed placed of articulation. More studies are thus needed to investigate more on the link between perception and production.

In contrast, the study done by Lambacher et al. (2005) offered positive results to both the perception and production of English vowels. This study trained 34 Japanese speakers on five vowels in American English, /æ/, /ɑ/, /ʌ/, /ɔ/, and /ɝ/, with a group of 20 subjects as control. The training involved the use of HVPT approach where the subjects had to identify 75 stimuli produced by five native speakers once per week for a total of 6 weeks. The perceptual test employed a five-alternative forced-choice identification task, with the five target vowels as the choices. The production tests required the subjects to record a set of words with the five target vowels within varied CVC contexts. The production performance was evaluated by a group of native speakers through identifying the vowels and by an acoustic analysis on their productions. It was found that both the perception and production performance of the subjects improved significantly, replicating the positive results in some consonant training studies.

Nobre-Oliveira (2007) trained the perception of three English vowel contrasts /i/-/ɪ/, /e/-/æ/, and /ʊ/-/u/ by Brazilian learners. Thirty-six subjects were involved in the study. Twenty-nine subjects participated in the perceptual training program while the remaining seven of them were in the control group. The usual pretest-treatment-posttest paradigm was used, and two training protocols were compared. One was HVPT approach, where the stimuli (nine monophthongs) were natural and produced by seven native American English speakers, whereas the other one used synthesized stimuli generated by a Praat script with the first three formant frequency (*F1*, *F2* and *F3*) values based on previous literature on American English. Prior to the perceptual training, the subjects attended a theoretical session about some basic knowledge of the vowels, such as their representations in the vowel chart. Then they proceeded to the perceptual training session, which was an identification task with immediate feedback. The training lasted for three weeks, 50 minutes each day, and the subjects were trained only on the front vowels in the first week, the back vowels in the second week and all the vowels in the third week. Before and after the training sessions, the subjects were required to do a pretest and a posttest, in which they had to identify 72 CVC target words plus 36 distracters. The results of perceptual performance from the pretest to the posttest were compared and both training methods had positive effects on the perceptual learning of the L2 vowels. Noticeable improvement was also found in the production of the vowel pairs, but only the /i/-/ɪ/ gain was statistically significant. This study also found that synthesized stimuli appeared to be more effective in training the subjects than natural stimuli, because synthesized tokens were enhanced by manipulating some crucial acoustic cues.

The last study discussed in this section that gauged the learning in both the perception and production domains after training was the one conducted by Iverson, Pinet and Evans (2011). Thirty-six French ESL learners participated in the study. The major goal of the study was to compare the training effects of HVPT on the perception and production of 14 British English monophthongs and diphthongs by two groups of ELS learners: one group was called the experienced who lived in England and had daily exposure to native English whereas the inexperienced group lived in their home country and had little exposure to English. The training stimuli utilized were the same as Iverson and Evans (2009) which were English words produced by eight native British English speakers. Although the study adopted a pretest-treatment-posttest procedure, it was different from previous works in the way that the stimuli used in both the pretest and posttest were not used in the training corpus and were produced by 10 native English speakers whose voices were not heard in the training. Generalization of learning was hence directly measured in the study. The pretest and posttest included a vowel identification test, category discrimination task and a vowel production test. All the subjects received eight sessions of HVPT over one to two weeks. Results of the three tests demonstrated learning to similar degrees in both groups. The authors concluded that phonetic training can improve the efficiency of categorization in a “nonsuperficial manner” that generalization and retention of learning can be observed (Iverson et al., 2011, p. 15). The degrees of learning in the three tasks were different and were not correlated to each other after training, with the greatest and broadest improvement found in perceptual identification, and small though significant improvement in a few contrasts in category discrimination and production tests.

2.4.3 L2 Production Training Studies

The aforementioned studies mainly investigated the effect of perceptual training on perceptual improvement in non-native or L2 segmental or suprasegmental contrasts while some tested if perceptual training can transfer also to production. Some studies focused on the other way round: whether production training would affect either/both perception or/and production. Since not only linguists but also teachers are interested in how to improve the performance of learners in perceiving and producing L2 sounds, production training studies have been managed and scrutinized in different ways, including training under laboratory conditions with computer-assisted programs, in classroom settings or with different means of feedback. This section delineates some L2 production training studies to set a foundation for an understanding of the present study.

As early as 1970, Catford and Pisoni (1970) designed an experiment to compare the training efficacy of perception and production training. The production training they provided was an explicit articulatory instruction, in which one group of participants with English as their L1 received systematic training in the production of some exotic phonemes (Group A). The participants were given minimal auditory exposure to the sounds while taught only the articulatory postures and movements before they produced and practised the pronunciation. Another group just received purely perceptual discrimination training (Group B). The experiment’s result demonstrated not only that Group A performed more than twice of Group B in the production posttest, but that they also performed significantly better at discerning the sounds during a perception posttest than the group trained under solely auditory techniques. The sample size of this study was not large enough, although the insights in the preliminary findings shed light on the forthcoming studies to investigate the link between speech perception and production.

Two groups of adult Dutch speakers learning the lexical tone system of Standard Chinese were trained under a counterbalance design (Leather, 1997). One group was trained to perceive the tones and was tested on how they produced the tones whereas the other group was trained to produce the tones followed by a perception test. Both were given computer-assisted training. The perception training group was presented with digitalized tone tokens with inter-speaker variation and had to label them before receiving the immediate feedback. Given not only verbal feedback but also visual ones through a real-time display of their F0 contours, the production training group produced the target tone words and could compare every production with a model presented on the screen. After the respective training, both groups improved in the mode in which they received no training. Leather (1997) hence concluded that there existed a carry-on of learning in one modality to another, suggesting that it tends to be sufficient for learners to be trained in either perception or production to achieve the improvement in the other mode. Nevertheless, this experimental design ignored the importance to compare the existing result with some base-line data of how perception training improves the perception of tones and how production training helps the production. Simply testing the transfer of learning from one mode to another can reveal the link between perception and production, but it cannot offer any further implication about the difference of direct training effectiveness in the specific mode with Leather’s counterbalance experimental design.

Akin to the methodology utilized in Leather (1997), production training can be administered by using computers as an assistant. Hirata (2004) trained a group of native English speakers to learn Japanese pitch and duration contrasts by using a computer program which provided the participants with the F0 contours as visual feedback after they produced the words. Their perceptual and production ability of words in sentences and words in isolation improved robustly after the training and significantly better than the control group with no training. A more recent pilot study compared the use of visual and auditory feedback in different means during production training and their effect on Japanese's and Korean's learning of the pronunciation of English sounds /w/, /ɪ/ and /iː/ (Carruthers, 2007). Provided with either a hand mirror or a webcam as a mirror to monitor their production of sounds, the participants improved their ability in the production of the sounds after the production training and performed better than the control group without training and the training group without the visual feedback. Although this pilot study needs expansion and improvement in subject size as there was only one subject in each group, it still gave some evidence of the potential benefits of using some tools to monitor the subjects' articulation of L2 sounds.

Nevertheless, not all production training studies demonstrated positive results in both the perception and production of the target sounds. Hattori and Iverson (2008) showed that ten sessions of production training provided to Japanese learners of English could only benefit their production of English /ɹ/-/l/ but not the perception of the pair. The training paradigm encompasses: 1) audio-visual examples and pronunciation instructions for English /ɹ/-/l/; 2) subjects produced the target sounds in words or minimal pairs with online feedback of the acoustic properties given; and, 3) subjects compared their production recordings with an enhanced version, i.e. an improved version with F3 and duration manipulated. Hattori (2009), who investigated also the perception of production of English /ɹ/-/l/ by Japanese speakers, found similar results as well. These two studies lead to my speculation that auditory training may affect only a specific level of processing, and perception and production may be two different domains that have separate underlying representations. Whether solely production training can benefit the gain of both domains and hence the how the two domains are linked together require more in-depth investigation.

These production training studies offered mixed evidence that production training sometimes can benefit learners' production of sounds, or in some cases, the perception, while some showed detrimental effects from production training. The present study will supplement the current literature of the effects of production training.

2.4.4 L2 Perceptual and Production Training Research

Besides training the subjects in only one domain, some studies began to investigate the effects of training in both modalities. Some compared perception or production training alone to training of perception with production training so as to address the question of the genuine efficacy of the training paradigms. Still, there were very few studies in this area and the ways of administering the training protocols varied in terms of the order and design of the training, whereas the results have also been mixed. This section will be devoting to introducing some which will shed some light to the present investigation, particularly for its methodological design.

Aliaga-García and Mora (2009) aimed to test which training paradigm can best promote the learning of native-like acoustic cues. The two training paradigms compared were modified versions of HVPT: identification (ID) or articulatory (ART) audiovisual HVPT. A total of 64 bilingual Catalan/Spanish leaners of English, divided into several groups, received the ID or ART training on 11 English RP vowels. This study adopted the pretest-treatment-posttest design. In the ID training, the subjects heard and watch a set of stimuli aided by video-clips produced by 10 different native English speakers and then chose the response from four choices. Immediate feedback was given. The ART training is a production-based perceptual training in which leaners were presented with the same set of words used in the ID training audiovisually. The subjects responded to each stimulus by imitating the native speaker's production, which were recorded for the replaying and comparing with native models at the end of each session. Results showed that the subjects who were trained under either the ID (perception only) or the ART (perception and production) could achieve more accurate perception and production of the target segments, although the size of the effect of training depended on the phonetic dimension of the sound contrasts. The results also revealed that after training, the subjects had more reliable use of cue weighting and less on duration.

Tsushima & Hamada (2005) recruited 45 Japanese speakers of English to receive perception and production training on six American English contrasts during a 13-week period. One group received perception training first before production, while another group received the training in a reversed order. Results showed that subjects who were trained with perception training first before production training demonstrated relatively weaker learning in production; whereas those who were trained under production training first before perception training showed significant and strong learning in perception. However, the order of training still was not significantly impacting the degree of improvement in both modalities. The results of this study indicated that the order of training might have some important implications for successful learning.

To examine the relationship between speech perception and production, Baese-Berk (2010) trained 40 subjects in the perception domain, in which resynthesized natural stimuli with manipulated VOT and formant transitions were provided. Another group with 41 subjects were trained in both perception and production domains. The training was administered by having the subjects, instead of responding through choosing a choice from the computer after listening to a stimulus, the subjects had to imitate the stimulus of which the repetition would be recorded. By comparing the pretest and posttest results which included a series of discrimination, categorization and repetition tasks, the author discovered that the subjects who were trained in perception only showed robust learning in the perception modality, but only small though significant improvement in production. Those who were trained under both domains demonstrated no changes in perception after training, but significant learning in the production modality was observed. The finding that receiving training in both perception and production domains led to disruption in perceptual learning was justified by the author in terms of the high cognitive demands required by in the training task. This experiment has shown that learning in one domain does not always transfer to the other modality; sometimes even disruptions occurred, meaning that a complicated relationship between speech perception and production exists.

2.4.5 Summary of Previous Research

Although investigations on vowels were still relatively rare, all the above training studies have offered substantial proof that phonetic training approaches can benefit the subjects' perceptual and/or production learning even on some difficult non-native contrasts at the segmental level. Native-like performance after the training can never be guaranteed, but significant improvement could still be found among many subjects with different L1s (e.g. Flege, 1995b; Hirata, 2004; Hirata, Whitehurst & Cullings, 2007; Kim & Hazan, 2010; Pruitt, 1995; Pruitt et al., 2006; Rochet, 1995; Tajima et al., 2003; Wang et al., 1999; Yamada & Munhall, 2008; Yamada et al., 1996). Added to this, using highly variabile training stimuli also promotes, rather than impedes, the perceptual learning of the subjects (e.g. Bradlow et al., 1997; Bradlow et al., 1999; Lively et al., 1993; Lively et al., 1994; Logan et al., 1991). Generalization effects to new words and new speakers were also found, especially when the subjects were trained with a wider range of stimulus variability (e.g. Bradlow et al., 1997; Wang, 2002; Wong, 2010). Several extended studies (e.g. Bradlow et al., 1997; Bradlow et al., 1999; Lively et al., 1994) have also shown that the training effects can be retained in the long run. Perceptual learning was also found to be capable of transferring to the production domain in general, although a lot of individual differences were observed across studies (e.g. Bradlow et al., 1999; Hazan et al., 2005; Lambacher et al., 2005). Researchers had also modified, expanded and combined some training techniques and paradigms together to maximize the effectiveness of the training approaches, among which HVPT paradigm was usually utilized as the skeleton of the modified training approach (e.g. Iverson & Evans, 2007, 2009; Lambacher et al., 2005; Nishi & Kewley-Port, 2005, 2007b; Sperbeck, Strange & Ito, 2005). Meanwhile, the mixed results shown from a few production training studies (e.g. Carruthers, 2007; Hirata 2004; Leather, 1997) suggest that production training such as explicit articulatory instructions or computer-assisted production learning may enhance or impede both the perceptual and production ability of the participants, which is an area worthy of research.

Based on all the above training studies, the present study also compares the efficacy of different combinations and modifications of HVPT and LVPT approaches and production training. Prior to the introduction of the methodology adopted in the present study, the background of the research will first be presented.

2.5 Aspects of the Present Study

The present research focuses on the perceptual learning of the English /ɪ/-/iː/, /e/-/æ/, and /ʊ/-/uː/ vowel contrasts among Hong Kong Cantonese ESL learners. The training effects on the production of the three pairs of vowels were also examined in this study. These vowel pairs are chosen for training because a number of studies (e.g. Chan, 2010, 2012; Chan & Li, 2000; Chang, 1975; Hensman, 1969; Hung, 2000; Leung & Brice, 2012; Meng, Zee & Lee, 2007, Sewell, 2009) on the perception and production performance of English vowels have indicated that the three pairs of vowels have posed identifiable problems among Hong Kong Cantonese speakers, particularly in their production. The L2 realizations of the vowel pairs, particularly /ɪ/-/iː/ and /e/-/æ/, were found to have caused intelligibility problems by native speakers as well (Brown, 1991; Jenkins, 2000; Sewell, 2009). Hung (2000) particularly ascribed the problem to the differences between the L1 (Hong Kong Cantonese) and L2 (English) phonological systems. L2 learners may find it difficult to acquire the second language, especially the sound system, as the learner's preexisting L1 system may have interfered with the perception of novel L2 sounds. It is worth scrutinizing and comparing the vowel systems of Hong Kong Cantonese and English first in the next section with a view to, on the one hand, understanding better the situation of the perception and production of English /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ contrasts among Hong Kong Cantonese speakers and, on the other, accounting for why these three particular vowel pairs were chosen as the major subject in the current study.

2.5.1 A Brief Overview of Cantonese Vowel System vs. English Vowel System and Cantonese Learners' Difficulties

English, a Germanic language, and Cantonese, a widely spoken Chinese dialect in Hong Kong within the Sino-Tibetan language family, have considerable typological differences which may create the difficulties for Cantonese learners to master the pronunciation of the English language (Chan & Li, 2000; Meng et al., 2007). To have a more focused understanding of the present research, this section will compare only the vowel systems of English and Cantonese. Learning difficulties attributed to the typological difference between the two language systems and in particular the three vowel pairs will be introduced.

English, if the Received Pronunciation (RP) accent is concerned, has 24 vowels in total, including eleven monophthongs, eight diphthongs and five triphthongs (Handbook of the IPA, 1999); whereas there are 18 vowels in Hong Kong Cantonese, with eight monophthongs and ten diphthongs (Barrie, 2003; Matthews & Yip, 1994; To, Cheung &

McLeod, 2012). The two vowel charts here display the basic differences between the vowel systems (only monophthongs are shown) of the two languages:

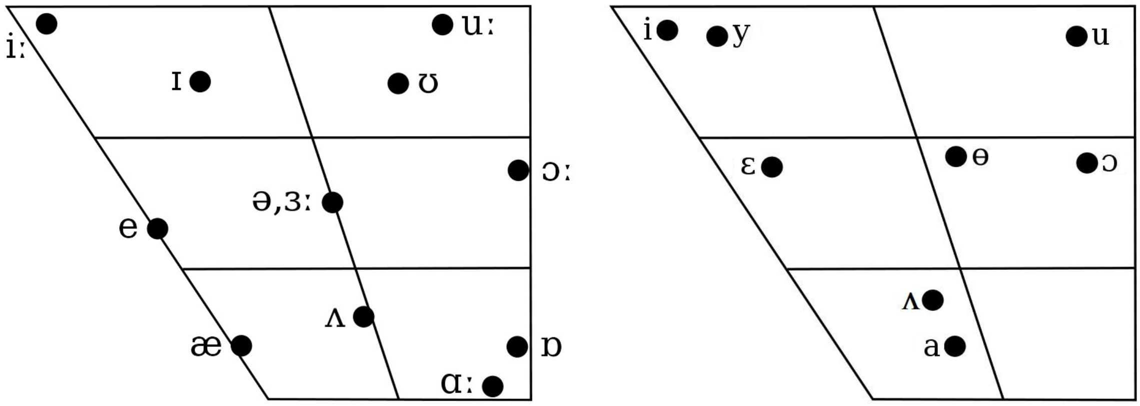


Figure 2.3. Left: RP vowel (phoneme) chart; Right: Hong Kong Cantonese vowel (phoneme) chart

2.5.1.1 Comparison of the Formant Frequencies of English /ɪ/-/iː/, /e/-/æ / & /ʊ/-/uː/ and the Expected Cantonese Assimilation Targets

Descriptions of the relative articulatory movements involved in vowel production offer only a scant understanding of the structure of the vowels. The acoustic measurements of the vowels can give more objective information concerning the properties of the sounds. They are the essential components in the intelligibility of sounds. The spectral peaks of the harmonic spectrum or the overtone structure of vowel[3], which is called the *formants*, is always the subject of acoustic analysis.

Formants arise due to the vibration of air in the vocal tract. Different vowels have different formant frequencies as they are produced with different shapes of the vocal tract, which are formed by the movement of the articulators. To distinguish one vowel from another, one can find out the first three formants of a vowel: F1, F2 and F3. The

[3] An overtone is defined as any frequency which is higher than the fundamental frequency. A vowel is characterized by different pitches simultaneously and various overtone pitches will give the vowel its particular quality that distinguishes itself from another.

lowest formant, symbolized as F1, is corresponds to the vowel openness: the more open a vowel is, the higher F1 values are. The second formant frequency, F2, represents vowel frontness/backness. Front vowels have higher F2 frequencies than back vowels. The third formant, F3, provides information on the quality distinctions and is related more to the position of the tip of the tongue.

Different speakers also have their distinctive vocal tract shapes and sizes; as a result, the formant values of one vowel differ from one person to another. Gender and age also make a difference in the pitch and hence the vowel quality. The various formant values obtained from different speakers about one single vowel are thus informative as they are usually being plotted in an F1 × F2 acoustic space (F1 and F2 to visualize the space where the vowels can lie in. The space is always represented as a circle or ellipse to delineate the possible range of a vowel produced by the native speakers of the language. The following example from Deterding (1997) displays a vowel acoustic space of Standard Southern British English obtained from approximately 10 occurrences per vowel from five male speakers:

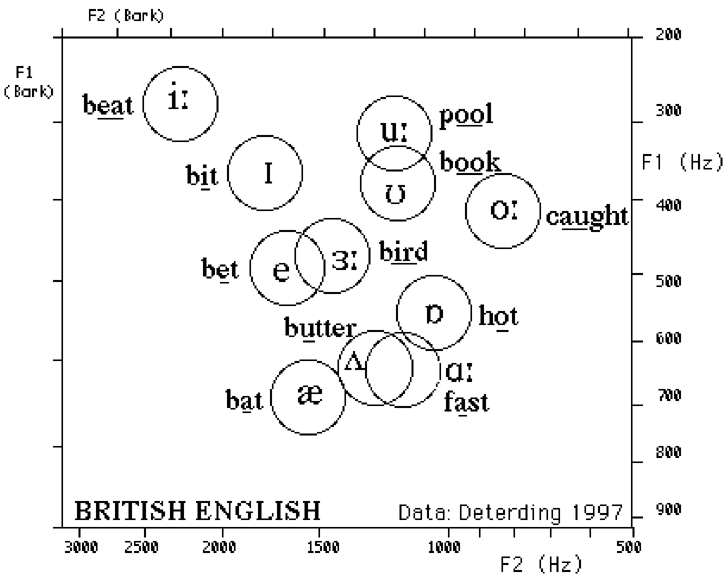


Figure 2.4. First (F1) and second (F2) formant frequencies of 11 English vowels as in Deterding (1997)

In the present study, the following three vowel pairs, /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/, are the focus. As the first two formant frequencies can already provide adequate information to disambiguate one English vowel (except the rhotic vowel /ɝ/ in General American accent, but it is beyond the focus of the present research) from another, the third formant frequency (F3) will not be discussed thereafter (Ladefoged, 2005). Instead, the temporal measurements of the vowel will be regarded as another subject of comparison as they can provide information on the vowel length.

The formant frequencies of the three English vowel pairs across several studies investigating different accents of English including General American (Carmell,1997; Hillenbrand, Getty, Clark, & Wheeler, 1995; Lambacher et al., 2005), Australian (Bernard & Mannell, 1986), South Australian (Butcher, 2006), Southern British English (Deterding, 1997) and Received Pronunciation (Hawkins & Midgley, 2005) are presented in Table 2.2 on the next page as reference. Since the English vowel pairs adopted in the present study cover all these accents, a comparison of their formant frequencies can shed light on the forthcoming discussion and provide a general picture of the vowel space of the three target vowel pairs.

Take the formant frequency and vowel duration values of female speakers in Hillenbrand et al. (2005) with General American accent as an illustration. It is given in Table 2.3.

For the first vowel pair, /ɪ/-/iː/, the vowel /ɪ/ has a higher F1 value and lower F2 value than that of /iː/, meaning that /ɪ/ is a higher and fronter vowel than /iː/. The vowel pair /e/-/æ/ are relatively lower than /ɪ/-/iː/, with the vowel /e/ being a lower

Table 2.2

Formant Frequencies and Vowel Durations for English Vowels /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ *produced by Native Speakers of English with Different Accents and Gender in several Studies*

	F1	*F2*	*F3*	Duration	Gender	Accent	Source
/ɪ/	370	2210	2740	N/A	Male	Australian	Bernard & Mannell (1986)
	410	2623	N/A	N/A	Female	South Australian	Butcher (2006)
	400	1900	2550	N/A	N/A	General American	Carmell (1997)
	350	1800	N/A	N/A	Male	Southern British	Deterding (1997)
	384	2145	N/A	N/A	Male	Received Pronunciation	Hawkins & Midgley (2005)
	483	2365	3053	237	Female	General American	Hillenbrand et al. (1995)
	390	1990	2550	N/A	Male	General American	Peterson & Barney (1952)
/iː/	300	2280	2800	N/A	Male	Australian	Bernard & Mannell (1986)
	396	2737	N/A	N/A	Female	South Australian	Butcher (2006)
	280	2250	2900	N/A	N/A	General American	Carmell (1997)
	260	2250	N/A	N/A	Male	Southern British	Deterding (1997)
	273	2325	N/A	N/A	Male	Received Pronunciation	Hawkins & Midgley (2005)
	437	2761	3372	306	Female	General American	Hillenbrand et al. (1995)
	270	2290	3010	N/A	Male	General American	Peterson & Barney (1952)
/e/	460	2040	2650	N/A	Male	Australian	Bernard & Mannell (1986)
	558	2352	N/A	N/A	Female	South Australian	Butcher (2006)
	550	1770	2490	N/A	N/A	General American	Carmell (1997)
	480	1620	N/A	N/A	Male	Southern British	Deterding (1997)
	556	1901	N/A	N/A	Male	Received Pronunciation	Hawkins & Midgley (2005)
	731	2349	2979	254	Female	General American	Hillenbrand et al. (1995)
	530	1840	2480	N/A	Male	General American	Peterson & Barney (1952)
/æ/	640	1870	2600	N/A	Male	Australian	Bernard & Mannell (1986)
	871	1853	N/A	N/A	Female	South Australian	Butcher (2006)
	690	1660	2490	N/A	N/A	General American	Carmell (1997)
	680	1550	N/A	N/A	Male	Southern British	Deterding (1997)
	807	1524	N/A	N/A	Male	Received Pronunciation	Hawkins & Midgley (2005)
	669	2058	2972	332	Female	General American	Hillenbrand et al. (1995)
	665	1726	2423	157	N/A	General American	Lambacher et al. (2005)
	660	1720	2410	N/A	Male	General American	Peterson & Barney (1952)
/ʊ/	400	910	2360	N/A	Male	Australian	Bernard & Mannell (1986)
	430	1038	N/A	N/A	Female	South Australian	Butcher (2006)
	450	1030	2380	N/A	N/A	General American	Carmell (1997)
	390	1200	N/A	N/A	Male	Southern British	Deterding (1997)
	397	1135	N/A	N/A	Male	Received Pronunciation	Hawkins & Midgley (2005)
	519	1225	2827	249	Female	General American	Hillenbrand et al. (1995)
	440	1020	2240	N/A	Male	General American	Peterson & Barney (1952)
/uː/	350	1600	2350	N/A	Male	Australian	Bernard & Mannell (1986)
	417	1960	N/A	N/A	Female	South Australian	Butcher (2006)
	310	870	2250	N/A	N/A	General American	Carmell (1997)
	310	1250	N/A	N/A	Male	Southern British	Deterding (1997)
	289	1476	N/A	N/A	Male	Received Pronunciation	Hawkins & Midgley (2005)
	459	1105	2735	303	Female	General American	Hillenbrand et al. (1995)
	300	870	2240	N/A	Male	General American	Peterson & Barney (1952)

Note: F1, *F2* and *F3* values are measured in Hertz while duration in millisecond. Unless otherwise specified, the values are averaged between gender.

Table 2.3

F1, F2 and Vowel Durations for English Vowels /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ *of a group of Female Speakers with General American accent in Hillenbrand et al. (2005)*

	F1	*F2*	Duration
/ɪ/	483	2365	237
/iː/	437	2761	306
/e/	731	2349	254
/æ/	669	2058	332
/ʊ/	519	1225	249
/uː/	459	1105	303

Note: F1 and *F2* values are measured in Hertz while duration in millisecond.

vowel than /æ/. The F2 value of /æ/ is smaller than /e/, as it is not as front as /e/ is. The vowel /æ/ is also longer than /e/. The two back vowels /ʊ/ and /uː/ are similar to /ɪ/ and /iː/ as they are also high vowels with relatively smaller F1 values, with /ʊ/ being a higher vowel than /uː/. Their relatively smaller F2 values indicate that they are back vowels. The vowel /uː/ also has a longer duration than /ʊ/.

Since the present study also recruited some native speakers of English to record the training stimuli for the experiment (details to be covered in Chapter 3), their productions of the three target vowel pairs are worth being taken as reference to better the understanding of the English vowel system. Figure 2.5 on p. 101 showing two F1 × F2 acoustic spaces were plotted by using the acoustic measurements produced by three male and three female native English subjects to visualize the results. By reading and comparing the F1, F2 and vowel duration values of the vowels, one can determine the vowel quality of any production. The acoustic information produced by native speakers of English is then an important benchmark for comparing the vowel productions made before and after training by the L2 subjects in this study.

It is worth nothing that in Cantonese, vowel length difference characterizes syllable structures as short vowels only occur in diphthongs (e.g. as [e] in [sej] "number four", underlying form: /ɛ/) or closed syllable (e.g. as [ʊ] in [hʊk] "cry"], underlying form: /u/). Since the short vowels [ɪ], [e] and [ʊ] are the surface forms of the long vowels /i/, /ɛ/ and /u/, they demonstrate some overlapping in their acoustic properties, as shown in Figure 2.6 on p.102. At this point, the acoustic properties of these Canontonese vowels, which Cantonese ESL learners always assimilate the three target pairs of English vowel to, are also provided for comparison:

Table 2.4

Averaged Formant Frequencies and Vowel Durations for Cantonese Vowels /i/, [ɪ], /ɛ/, [e], /a/, /u/ and [ʊ] *produced by 11 Males and 10 Females Native Speakers of Cantonese*

	Male			Female		
	F1	*F2*	Duration	*F1*	*F2*	Duration
/i/	393	1849	151	449	2319	130
[ɪ]	499	1790	133	571	2138	114
/ɛ/	573	1863	203	747	2234	316
[e]	412	2047	174	516	2309	164
/a/	736	1248	190	973	1569	201
/u/	464	1420	189	505	1220	159
[ʊ]	550	1368	145	634	1220	133

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

However, as noted earlier, the three English vowel pairs, /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ have been documented in the literature as posing both perception and production difficulties among Cantonese ESL learners in Hong Kong who usually assimilate their production to the most similar Cantonese counterpart, resulting in audience's comprehension difficulties: /ɪ/-/iː/ to /i/; /e/-/æ/ to /ɛ/ and /ʊ/-/uː/ to /u/ (cf. Chan 2010, 2012; Chan & Li, 2000; Hung, 2001; Meng et al., 2007; details of these research

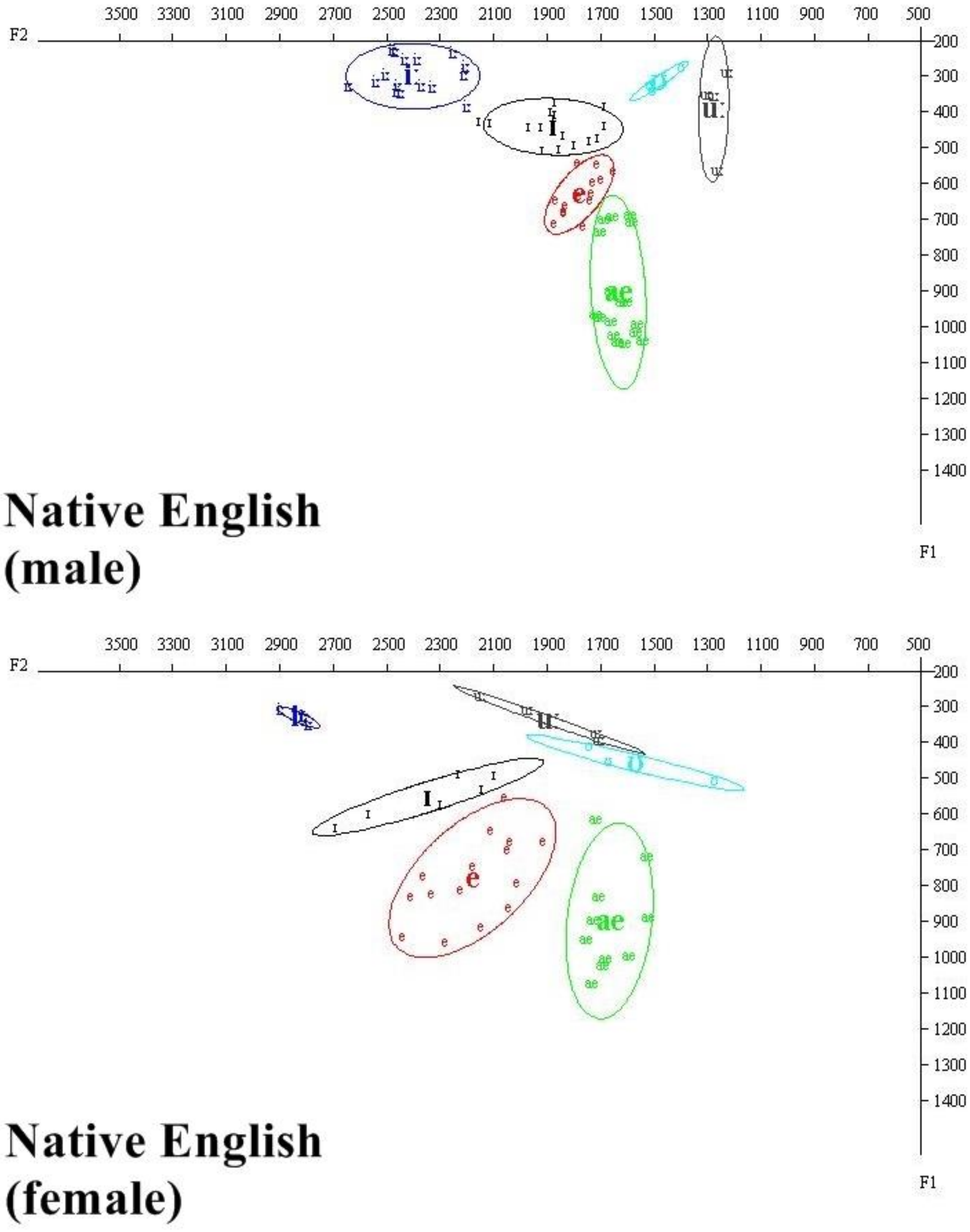


Figure 2.5. First (F1) and second (F2) formant frequencies (in Hertz) of the three English vowel pairs produced by three male (above) and three female (below) speakers in the present study.

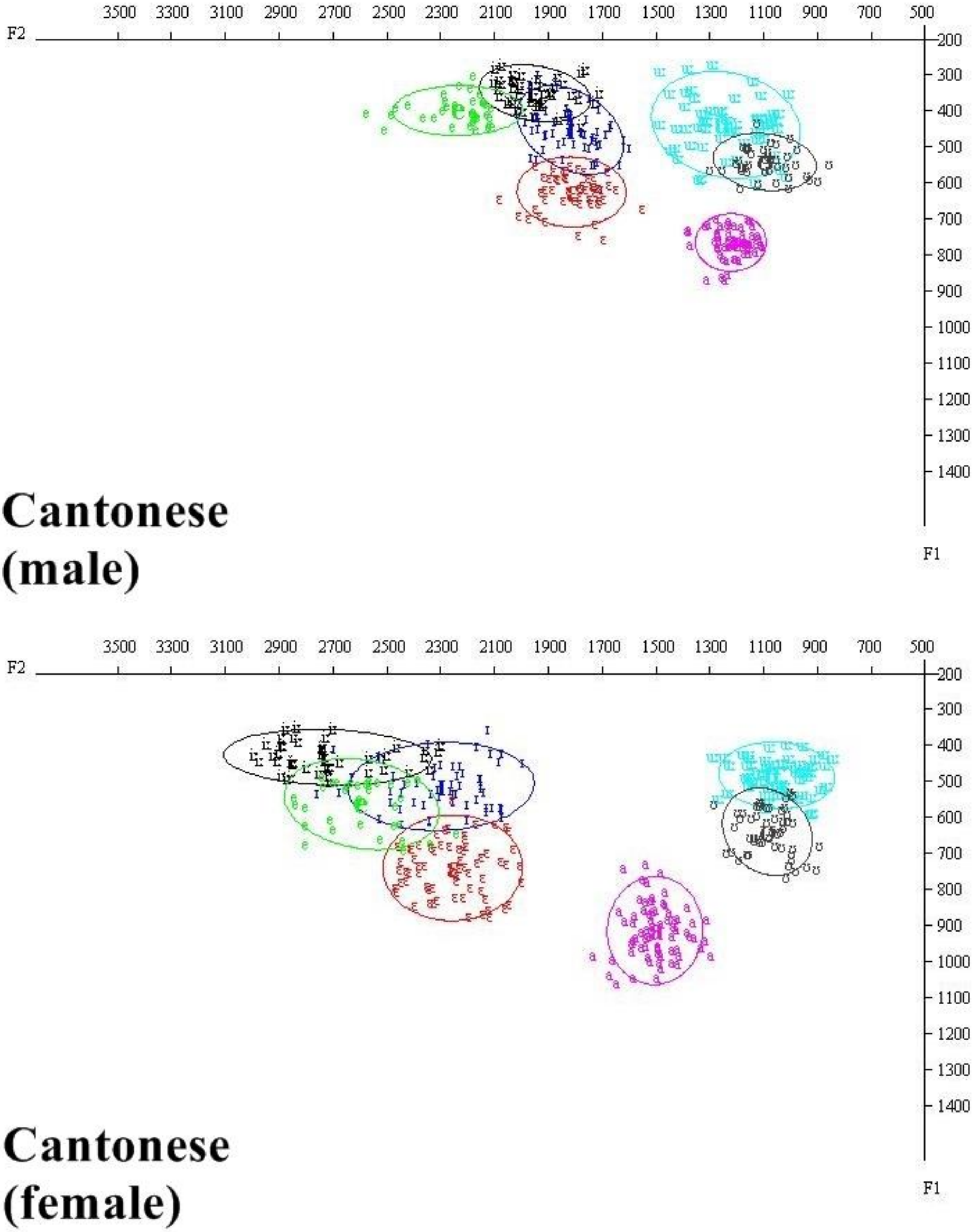


Figure 2.6. First (F1) and second (F2) formant frequencies (in Hertz) of the seven Cantonese vowel produced by 11 male speakers (above) and 10 female speakers (below)

findings will be discussed in Section 2.6.2 below). The learners generally have a high degree of homogeneity in the assimilation and they had no tense-lax distinction both in perception and production of the three vowel pairs.

2.5.2 CANTONESE LEARNERS' DIFFICULTIES

Contrasting the differences between Cantonese and English phonemes as a prelude to understanding and accounting for the problems and difficulties in the phonological acquisition of Cantonese learners has been done preliminarily by Chan and Li (2000). This present research can be benefited from Chan and Li's study through their prediction of Hong Kong learners' developmental problems in acquiring the English sound system, one of which is the confusion in producing the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/. As stated in Chan (2010, 2012) and Chan and Li (2000), most of the Cantonese speakers of English do not distinguish between /e/-/æ/ especially in the mouth openness, as /e/ is mid-low vowel while /æ/ is a low vowel; or tense or lax vowels (i.e. /ɪ/-/iː/ and /ʊ/-/uː/) as Cantonese speakers under-differentiate the distinction of the vowel length. Cantonese speakers of English tend to substitute /e/ for /æ/, or the other way round, and also use the short vowel for long, the long vowel for short, or the vowel with length somewhere in between the long and short ones. This suggests that Cantonese speakers of English generally have problems in producing the three pairs of vowels which are not present, or have different realizations in their language, thus sometimes cause communication problems. The investigation by Meng et al. (2007) also stated the same over-generalization problem of the three vowel pairs prevalent among Cantonese speakers of English in Hong Kong.

Additionally, Hung (2000) also scrutinized the possibility of an existing variety called *Hong Kong English*[4], and hence Hong Kong English phonology. As with the English /e/-/æ/ pair, Cantonese speakers generally use /ɛ/, which exists in Cantonese as a mid vowel, to substitute for both /e/ and /æ/, meaning that words with the two English vowels will be produced as the same in the speech of a Cantonese speaker of English. The following table shows the results obtained in Hung (2000):

Table 2.5

Average Formant Frequencies of the Three Vowel Pairs produced by 15 Subjects in Hung's (2000) Study

Vowel	Items	F1	F2	Duration
/ɪ/	Hid	368	2601	138
	Hit	372	2557	117
/iː/	Heed	346	2572	188
	Heat	350	2604	117
/e/	Head	668	1961	160
	Bet	741	1919	169
/æ/	Had	701	1941	144
	Bat	689	1959	179
/ʊ/	Hood	402	926	147
	Hook	485	973	109
/uː/	Who'd	372	796	224
	Hoot	404	992	121

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

The formant frequencies, F1 and F2, of the words with /e/ and /æ/ produced by the subjects in the study were very close, meaning that the two vowels produced were highly similar. The production of Cantonese /i/ for both /ɪ/ and /iː/ and /u/ for both /ʊ/ and /uː/ is of a similar situation, with very close F1 and F2 values, and highly similar

[4] Whether *Hong Kong English* in terms of it grammar and usage exist still remains controversial because the government and most sectors in Hong Kong do not accept this variety of English and this is beyond the discussion in this dissertation. However, it is evident that most Cantonese speakers of English produce a different vowel set with the native speakers of English in general, hence the phonology of *Hong Kong English*. Most of the Cantonese ESL learners speak with a *Hong Kong English* accent which is characterized by a flat intonation, a syllable-timed rhythm, simplified consonant clusters, etc.

vowel duration in particular for words ending with a voiceless obstruent, suggesting that the Cantonese subjects generally lack tense and lax distinction during their production. The findings obtained by Deterding, Kirkpatrick, & Wong (2008), who investigated the production of *Hong Kong English* produced by 15 female teacher trainees, also converge with Hung's (2000) results that both /ɪ/-/iː/ and /e/-/æ/ pairs tend to be merged and paired.

Hung (2000) also investigated the perceptual aspect of the subjects. It is found that if the two vowels are produced rather distinctively by even a native RP speaker, Cantonese speakers may still find that they are the same (in a perception test done by Hung, with 15 native Cantonese speakers, the subjects can only distinguish 47% of /e/ while only 60% of /æ/produced by native English speakers). Whereas for both /ɪ/-/iː/ and /ʊ/-/uː/ pairs, the problem lies in the inability in distinguishing the difference in the vowel quality (tense vs. lax) and the length (long vs. short) of the two pairs of vowels. The low perception accuracy in identifying the two vowel pairs (particularly low for /ɪ/ (33%) and /uː/ (53%); and 67% for /iː/ and 73% for /ʊ/) was reported by Hung (2000) in the same study. From these accounts, one can know that Cantonese speakers of English suggest also the difficulties in both the perception and production of the three vowel pairs. It is acknowledged that these modifications can be dealt with by using specific training approaches, such as HVPT, which focuses on raising the perceptual ability as a necessary step for better production, to raise both the perception and production of specific sounds.

Productions of the three English vowel pairs by 15 male and 15 female Cantonese

ESL learners without any training given in the present study were also collected for comparison prior to the experiment. Table 2.6 display the averaged F1 and F2 values and vowel durations produced by Cantonese ESL learners. The findings replicate the results of Hung (2000), as the F1, F2 and durations values of the vowel pairs respectively, were very close, signifying that the subjects had difficulty in discriminating the English vowel pairs.

Table 2.6

Average Formant Frequencies of the Three English Vowel Pairs produced by 15 Males and 15 Females in the Present Study

	Male			Female		
	F1	F2	Duration	F1	F2	Duration
/ɪ/	610	2184	123	507	2167	122
/iː/	602	2269	177	442	1991	166
/e/	704	1872	155	847	1789	180
/æ/	739	1773	168	949	2122	175
/ʊ/	546	1555	174	514	1576	172
/uː/	460	1581	193	496	1590	196

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

From these accounts, one can know that Cantonese speakers of English may also have difficulties in both the perception and production of the three vowel pairs. It is acknowledged that these modifications can be dealt with by using specific training approaches, such as HVPT, which focuses on raising the perceptual ability as a necessary step for better production, to raise both the perception and production of specific sounds.

2.5.3 THE PRESENT RESEARCH

Regarding the Hong Kong context, no research has been done on the effectiveness of HVPT training method, emphasizing only on the training in perception which leads to an improvement in production. The teaching and training of phonology has mainly focused on the fine articulatory movements and production as the final product, hence neglecting and deterring the development of more conducive and all-rounded pedagogical approaches and training techniques.

The studies listed above mainly focus on describing the errors made by adult learners. Even for those that have investigated the effectiveness of HVPT, many studies take training intensity as a constant. They have not considered a possible extraneous factor, the intensity of the perceptual training approach, which may influence the effectiveness of the training methods. The aforementioned investigations also have not tested high-form secondary schools students who are supposed to have got over the Critical Period and have acquired the language. This research there may shed some lights on a novel way of training Hong Kong students' English vowel perception, which focuses on improving the perceptual ability of learners as the prerequisite of improving the production, as well as providing a more lucid account of the interaction between the perception and production of speech sounds.

Differed from the Low Variability Phonetic Training (LVPT) which only adopts natural stimuli produced by one talker (e.g. Strange & Dittmann, 1984), HVPT is believed to have forced the participants to focus on phonetic cues that underlie categorical distinctions and to enhance long-term recall. HVPT has been proved effective in training non-native contrasts in some studies (e.g. Bradlow et al., 1997; Lively et al., 1993; Lively et al., 1994; Logan et al., 1991). However, most of these studies only focused on consonants and how the perceptual training affected the production in one-way (e.g. Lambacher et al., 2005; Wang, 2002). Additionally, the production performance of the participants was checked only by inviting native speakers as judges, but not through also spectrogram analysis to enhance the reliability and validity.

At present, except my M.Phil. research (Wong, 2010) and related studies (Wong, 2012, 2013a, 2013b) on comparing HVPT and LVPT on improving the perception and production of English vowel contrasts (e.g. /ɪ/-/iː/ and /e/-/æ/) produced by Hong Kong Cantonese speakers of English, which are long-standing difficulties among them, no research has been conducted to investigate and compare useful training approaches for improving the perception and production of L2 segments (Chan 2010, 2012; Chan & Li, 2002; Hung, 2000). One study by Wang (2002) was insightful in its investigation in training Mandarin and Cantonese speakers' learning of /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ by HVPT, but the subject size was rather small and production training has not been investigated for exploring the link between perception and production. So as to understand more about the test design in a Hong Kong context, my previous research of a similar nature (Wong, 2010) will be briefly presented here.

2.6 WONG, 2010

This study probed into the effectiveness of two phonetic training approaches, High-Variability Phonetic Training (HVPT) approach and Low-Variability Phonetic Training (LVPT) approach, on the modification of the perception and production of

English vowels /e/ (as in *bed*, *said*, and *pet*) /æ/ (as in *bad*, *sad* and *pat*) among a group of Cantonese-speaking secondary school students.

This study adopted a pretest-treatment-posttest design with supplementary generalization tests. 17 subjects were trained under HVPT approach while 24 were trained under LVPT approach for 10 training sessions. Another 23 subjects served as the control group and did not take part in any training. Under each training paradigm, they are subdivided into two groups according to their general English proficiency levels (high and low, determined by the listening and oral exam results in a recognized public exam in Hong Kong) as most of the previous studies tested highly-advanced ESL learners only.

In the pretest phase, all the subjects produced a list of words with the two target vowels. They then identified one of the counterparts of the /e/-/æ/ minimal pair in the perception test. Both statistical and acoustic analyses showed that the subjects generally confused the two vowels in both perception and production. All the trained subjects only received perceptual training, which was a two-alternative forced-choice identification test with immediate feedback. The subjects had to identify the vowel (either /e/ or /æ/) from a list of words after listening to each token. The difference between the training approaches of the two groups were that a set of perceptual training stimuli produced by six native speakers of English were utilized in HVPT while the same set of words produced by only a single native speaker was adopted in LVPT. Both training approaches were effective in improving the subjects' perception of the two vowels, while those trained under HVPT approach demonstrated even greater improvement. The effects of proficiency was however an insignificant effect and in other interactions. Perceptual learning could also be generalized to new words and new speakers for the two experimental groups.

As for the transfer of perceptual learning to the production, the trained subjects improved significantly from the pretest to the posttest and HVPT group also outperformed LVPT group. Nevertheless, no significant difference was discovered for subjects with different proficiency levels. The formant frequencies and durations of the vowels produced by the trained groups after the training were also found to be closer to native-like productions. However, although the number of target productions in the generalization test of production at the sentence-level (a passage reading task was given) was high, no significant difference was found across groups.

The findings in this study imply that training in perception alone appeared to be useful for improving both the perception and production of the non-native contrast among the subjects, let alone their proficiency level. As the first English vowel training study in a Hong Kong context, this pioneering research draws our attention to the link between speech perception and production. It also offers preliminary empirical support to the theoretical issues related to cross-language vowel training and second language acquisition. Yet, investigating the relationship between the two domains requires further research such as investigating whether production training, comparing different test designs, etc. can also affect perception or the effectiveness of the training approaches to other vowel pairs.

2.7 RESEARCH QUESTIONS

The present study will fill the research gaps identified in previous studies, by comparing HVPT and LVPT, increasing the subject size, modifying the methodology such as the test design and investigating the factors of training intensity and a combination of test types on the effectiveness of the training on the three target vowel pairs. All seven research questions are as follows:

RQ1. Are the two perceptually-based phonetic training approaches, HVPT and LVPT, effective in improving

a. the perception and
b. the production of the English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/?

RQ2. Which training approach, HVPT or LVPT, is more effective in improving

a. the perception and
b. the production of vowels /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/, and how different are they?

RQ3. What are the training effects of different perceptual training intensities in

a. the perception and
b. the production of the English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/?

RQ4. Is the production training effective in improving

a. the perception and
b. the production of the English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/?

RQ5. Are there are any differences in both

a. the perception and
b. the production performances of the English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ when the participants are
    i. trained only under either HVPT or LVPT;
    ii. trained only under the explicit production training or
    iii. trained under both the perception approach (either HVPT or LVPT) and the production training, and how different are they?

RQ6. Can any learning effect be generalized to

a. the perception of new words produced by both familiar and new speakers, and
b. to the production in a more naturalistic environment?

RQ7. Are there any differences in the ease of

a. perceptual identification or
b. production of the three English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/?

CHAPTER 3

RESEARCH METHODOLOGY

3.0 INTRODUCTION

In the previous chapter, it has been established that the current study mainly aims to compare the effectiveness of the HVPT and LVPT approaches on the modification of the perceptual and production performance of non-native contrasts by Cantonese ESL learners as well as how much perceptual training intensity and diversity will influence the training results. This chapter opens with Section 3.1, which introduces the background and grouping details of the research participants, followed by Section 3.2 detailing the research design and procedures. Section 3.3 will be devoted to illustrating the stimuli and training materials and lastly Section 3.4 will explain how evaluations and transcription reliabilities are achieved in the present study.

3.1 RESEARCH PARTICIPANTS

Eighty-five Secondary 5 local Hong Kong students, aged between 16 and 17 (an average age of 16.3) at the time of the study, were recruited as the participants. They all spoke Cantonese as their L1 and English as their L2. They were English learners at the upper-intermediate level with an age of learning English as an L2 in Hong Kong at 3.62 of age on average i.e. for an average of 12.70 years of learning. The school used English as the medium of instruction (MOI)[5], meaning that except Chinese language classes, all other subjects were taught in English. The teachers in the school were all advanced Cantonese ESL learners with only two native-speaking ones. Their English classes included no explicit pronunciation instruction during the time of the study and only a small amount of production learning was received before they were enrolled into the study. All of them were Band 1[6] students, i.e. the top one-third students in Hong Kong when they were firstly admitted to the school, whose English proficiency levels and learning aptitudes were closely similar to each other.

According to self-report, they seldom interacted with native English speakers and had very little exposure to English outside of school. None of them had resided in any English-speaking countries for more than a month. No one reported hearing or speaking impediments. All volunteered to join the experiment after the researcher's introduction

---

[5] Chinese (written form; Cantonese as the spoken form) and English are the official languages of Hong Kong. While most Hong Kong people spoke Cantonese as their first language, English is usually the second language, or even foreign language to some of the locals, depending mostly on a wide range of social factors (Bolton, 2003). Being an international metropolis and a once British colony, Hong Kong has placed English as an important instructional language at different levels of institutions since the colonial period. In 1997, the "Mother Tongue Policy" has been enforced in secondary education. Since then, only 114 schools in Hong Kong were allowed to use English as the MOI while others had to use Cantonese as the MOI (Bolton, 2002). Almost all schools in Hong Kong used British and/or American English (the englishes in the *Inner Circle*, according to Kachru, 1992) as the models. In recent years, the MOI arrangement has been loosened and the school can choose their own MOI during lessons. The present participating school has used English as its MOI since its establishment in 1985, with both British and American English as their teaching models. The teachers are with diversified background and speak with British, General American or Hong Kong English accent. All the participants are mostly exposed to these three varieties of English in the daily life.

[6] *Information Leaflet on the Secondary School Places Allocation System* [Leaflet]. (2012). Hong Kong: Education Bureau.

and teachers' promotion at the school. In between training breaks, they were given some refreshments.

The subjects were randomly assigned to one of the 10 groups, 9 of which received perception and/or production training while one group receiving not intervention, i.e. the control group[7]. So as to shed light on the research questions, the subjects were grouped under the three facets:

1) according to the perceptual training method utilized
(the HVPT, the LVPT, or no perceptual training);
2) according to the training intensity of perceptual training approaches
(either 2 days – intensive, or 10 days – standard) and,
3) according to whether or not they received the production training.

Training intensity, as one aspect of the grouping trichotomy, is one major dependent variable that the present study investigates. How it is implemented and defined play an influential role in the experiment. As explicated in Chapter 2, training intensity has various definitions based on the emphasis of different research. The present study has adopted some vital elements of definition of training intensity that suit

[7] Originally, this all-inclusive design was planned to have a total of 250 students in the whole Secondary 5 of the school to see the interaction between each factor. Yet only 198 volunteered to joined the experiement, and with a large percentage of drop-outs after completing the pretest phase or the mid-way of the training phase, only 95 students were left and they were put into the ten groups randomly with the first 5 groups having 10 participants and the last 5 groups with 9 participants. Due to the students' difficulties in time management and compromising on their schedule, three participants in the training groups requested to become the control subjects or would just receive the explicit production training because they found that their schedule did not fit the whole experiment agenda. Another eight students dropped out at the last phase of the experiment; two suffered from long-term illnesses and could not participate anymore. As a result, the ideal case of having a large number and an equal number of participants in each group could not be achieved. However, the number of participants in each group is still very close and is shown on the next page.

the experimental goal. This study controls for dose, dose form and cumulative intervention intensity, allowing only dose frequency, total intervention duration and training episode distribution to vary. Table 3.1 below elaborates on how training intensity is defined in the present study and elucidates the differences between the two types of training groups based on the terminologies used in Warren et al. (2007):

Table 3.1

Details of the Training Intensity in the Perceptual Training Approaches

	Intensive	**Standard**
Dose	120 training stimuli received per one training session (self-paced, but within 15 minutes)	
Dose form	Listening to a perception program (exposed to either HVPT/LVPT stimuli)	
Dose frequency	10 sessions per day	2 sessions per day
Total Intervention Duration	2 days (on alternative days)	10 days (on alternative days)
Cumulative intervention intensity	(Dose × Dose Frequency × Total Intervention Duration) 120 × 10 × 2 = 2400 tokens	
Training episode distribution	Massed	Distributed

Based on the above definition of training intensity in the current study (see Chapter 2, Section 2.4.1.4), the researcher randomly assigned the subjects into different groups:

1. *HSP group (**H**VPT + **S**tandard + **P**roduction training)*
   9 subjects trained under the HVPT (standard) with production training
2. *HIP group (**H**VPT + **I**ntensive + **P**roduction training)*
   9 subjects trained under the HVPT (intensive) with production training
3. *HSN group (**H**VPT + **S**tandard + **N**o production training)*
   9 subjects trained under the HVPT (standard) with no production training
4. *HIN group (**H**VPT + **I**ntensive + **N**o production training)*
   8 subjects trained under the HVPT (intensive) with no production training
5. *LSP group (**L**VPT + **S**tandard + **P**roduction training)*
   7 subjects trained under the LVPT (standard) with production training

6. *LIP group (**L**VPT + **I**ntensive + **P**roduction training)*
   7 subjects trained under the LVPT (intensive) with production training
7. *LSN group (**L**VPT + **S**tandard + **N**o production training)*
   7 subjects trained under the LVPT (standard) with no production training
8. *LIN group (**L**VPT + **I**ntensive + **N**o production training)*
   8 subjects trained under the LVPT (intensive) with no production training
9. *COP group (**CO**ntrol + **P**roduction training)*
   11 subjects with no perceptual training but with production training
10. *CON group (**CO**ntrol + **N**o production training)*
   10 subjects with no perceptual training and no production training

A summary of the subjects' training conditions, language backgrounds and other personal details are also shown in the table:

Table 3.2
Summary of the Training Conditions, Language Backgrounds and Personal Details of the Subjects in the 10 Groups

Group Name	Perceptual Training Approach	Perceptual Training Intensity	Production Training	Gender	Mean		
					Age	Age of Onset	Years of Learning
HSP	HVPT	Standard	Yes	3M6F	16.44	3.22	13.22
HIP	HVPT	Intensive	Yes	4M5F	16.11	3.33	12.78
HSN	HVPT	Standard	No	3M6F	16.33	3.67	12.89
HIN	HVPT	Intensive	No	2M6F	16.38	3.75	12.50
LSP	LVPT	Standard	Yes	3M4F	16.14	4.00	12.14
LIP	LVPT	Intensive	Yes	2M5F	16.29	3.71	12.57
LSN	LVPT	Standard	No	2M5F	16.43	3.29	13.00
LIN	LVPT	Intensive	No	2M6F	16.50	3.50	13.00
COP	Control	N/A	Yes	4M7F	16.22	3.22	13.00
CON	Control	N/A	No	4M6F	16.60	5.40	11.10

3.2 Research Design and Procedures

3.2.1 The Research Setting

The study was conducted from March to August in 2012. The implementation of the entire experiment took 1 to 4 weeks on weekdays[8], depending on their group nature. So as to let the subjects be relaxed while participating in the experiment, all tests and the training sessions were conducted in a language laboratory located at the subjects' school. The language laboratory was a 45-seated multi-media learning center (MMLC) equipped with good sound insulation system and adequate facilities such as computers with audio-playing and recording software, headphones and microphones. These facilities were utilized in both the training and testing sessions. Two separate sound-proof booths were used when the subjects recorded the test tokens. The researcher and two IT technicians were present to supervise the proper use of all the equipment and software. They could also answer subjects' enquires during the training and testing sessions.

3.2.2 THE DESIGN IN DETAIL

With a three-phased "pretest-treatment-posttest" structure, the present study employed an experimental design with eight groups of target subjects taking part in either the High Variability Phonetic Training (HVPT) sessions or the Low Variability Phonetic Training (LVPT) sessions in one to four weeks[9], one group receiving just the

---

[8] A very strict schedule could not be maintained as the student volunteers only participated in the experiment in their spare time during lunch break or after school on normal school days. The experiment hence straddled weekends.

[9] Since it would be very difficult to request the subjects to participate in the experiment on Saturdays or Sundays when they had no school, all tests and training sessions were conducted only during weekdays for the convenience of the subjects. Thus for the case when 10 sessions of perceptual training plus 4 sessions of production training were given, the participation period would be spanned over 4 weeks as the subjects took each type of training on every other day.

production training, and also a control group receiving no treatment. All subjects, except the control group, participated in all three phases:

PHASE 1. Pretest phase, which included 1 production pretest and 1 perception pretest;

PHASE 2. Treatment phase:
a) Perceptual Treatment Phase: all groups except CP and CN participated in this training. This phase comprised a total of 20 training sessions (either the HVPT or LVPT; except for two control groups) in which they attended 2 (standard) or 10 (intensive) sessions in a day;
b) Production Treatment Phase: 5 of the groups received explicit production training for the three vowel pairs for 4 30-minute sessions on 4 separate days, and finally,

PHASE 3. Posttest phase with 1 production posttest, 1 perception post-test, 1 production Test of Contextualization and 3 perception Tests of Generalization.

A flowchart on the next page summarizes and describes all the phases in which the subjects had participated; another table on the page after next also illustrates the participation schedule for each group. Detailed information of the research procedures and materials utilized will be introduced in detail later.

3.2.2.1 Before the Experiment: the Preparatory Stage

One month before the experiment commenced, all subjects filled in a consent form to confirm that the researcher had the access right to collect the experimental data for research purposes. They also completed a language background and personal information, e.g. their English learning experience, their willingness to take part in the experiment and their schedule, etc., before choosing potential candidates for this study.

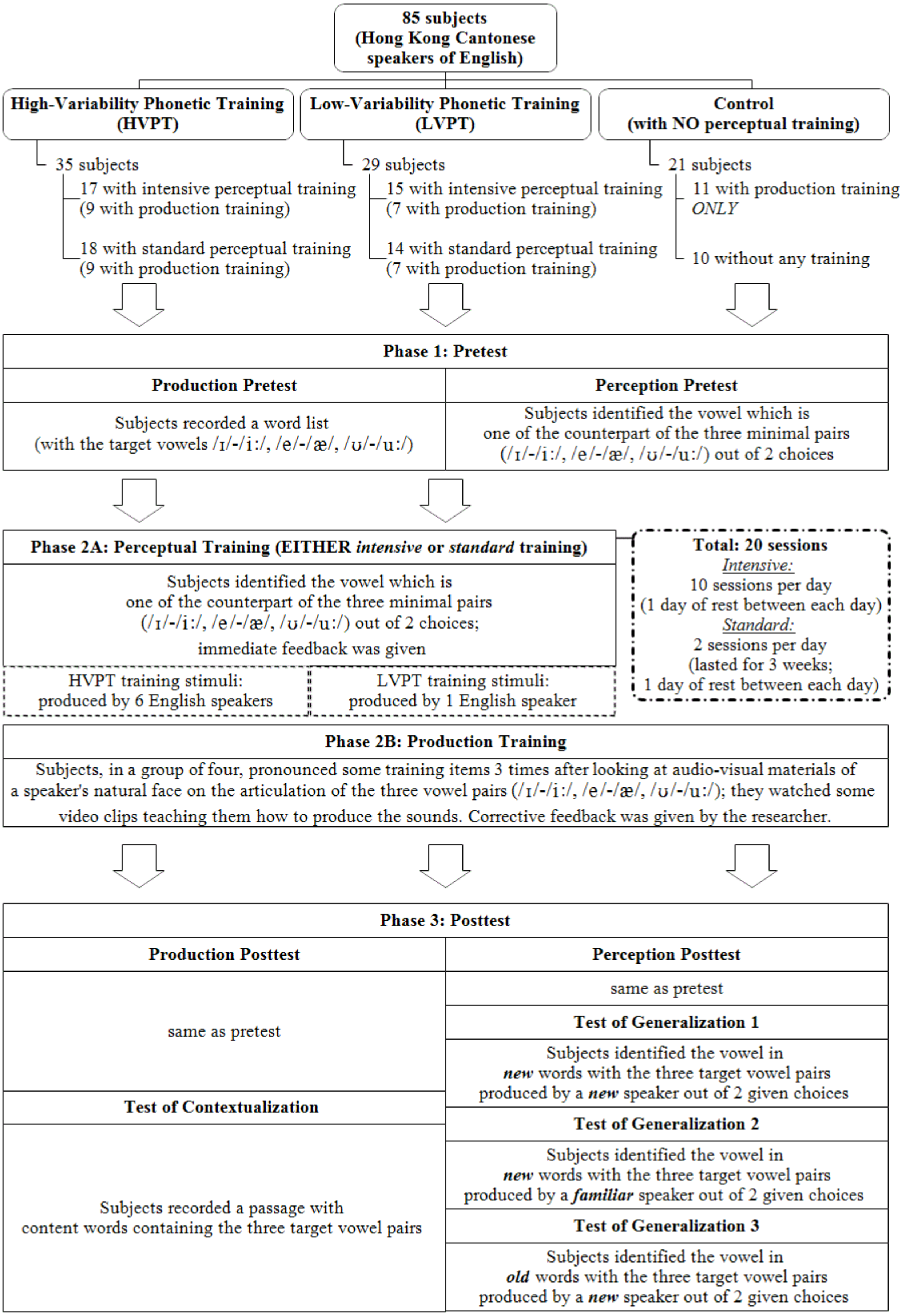


Figure 3.1 Flowchart showing the research procedures

		Week 1							Week 2							Week 3							Week 4						
Group	Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
HSP	Test	Pre																							Post				
	Perc. Train	H		H		H	No School			H		H		No School		H		H		H	No School		H		H			No School	
	Prod. Train		✓		✓				✓		✓																		
HIP	Test	Pre									Post																		
	Perc. Train	H		H			No School							No School							No School							No School	
	Prod. Train		✓		✓				✓		✓																		
HSN	Test	Pre																					Post						
	Perc. Train	H		H		H	No School		H		H		H	No School		H		H		H	No School		H					No School	
	Prod. Train																												
HIN	Test	Pre		Post																									
	Perc. Train	H		H			No School							No School							No School							No School	
	Prod. Train																												
LSP	Test	Pre																							Post				
	Perc. Train	L		L		L	No School			L		L		No School		L		L		L	No School		L		L			No School	
	Prod. Train		✓		✓				✓		✓																		
LIP	Test	Pre									Post																		
	Perc. Train	L		L			No School							No School							No School							No School	
	Prod. Train		✓		✓				✓		✓																		
LSN	Test	Pre																					Post						
	Perc. Train	L		L		L	No School		L		L		L	No School		L		L		L	No School		L					No School	
	Prod. Train																												
LIN	Test	Pre		Post																									
	Perc. Train	L		L			No School							No School							No School							No School	
	Prod. Train																												
COP	Test	Pre							Post																				
	Perc. Train						No School							No School							No School							No School	
	Prod. Train	✓		✓		✓			✓																				
CON	Test	Pre							Post																				
	Perc. Train						No School							No School							No School							No School	
	Prod. Train																												

Figure 3.2 An illustration of participation schedule for each group

The consent form and questionnaire are provided in Appendix A.

Before the pretest phase, the researcher ensured that the subjects knew the meanings and pronunciations of all the minimal pairs contrasting /ɪ/-/iː/, /e/-/æ/, and /ʊ/-/uː/ on a word list which was to be used in both the production and perception pretests. This word list is in Appendix B. When the subjects reported that they had not encountered the word, the researcher explained the meanings to them and played the pronunciation clip of the words produced by a female native speaker (with General American accent) who had not involved in any tests or training sessions of the study. This ensured that the pronunciation of the words can reflect an authentic performance of acquisition of the three pairs of vowels. The subjects also did not have to guess the pronunciation of the segments apart from the vowels. Since the subjects had difficulties in distinguishing the vowel pairs in perception and production, the possibility that the subjects were imitating directly from the audio clips was low; meanwhile, the subjects could accurately produce the segments apart from the target vowel.

So as to provide constructive and fair results, the researcher ascertained that all the subjects were not overqualified for this research. They must have lower than 50% of perception or production accuracy in at least one pair of vowels before joining the experiment. This was achieved by immediately examining all the pretest results. All subjects were found to have both perception and production confusion in at least one pair of vowels.

3.2.2.2 Phase 1: Pretest Phase

The pretest phase involved two tests: 1) the production pretest, which was a word list reading task, and 2) the perception pretest. All the subjects including the control group participated in both tests. Each time at most 10 subjects were present in the language laboratory to complete the test. They were required to sit far away from each other, particularly when they did the production tests and training so that they would not disturb their neighbors. The production pretest was administered first to avoid subjects' cueing or exposure to the items which would appear later in the perception pretest and training.

3.2.2.2.1 PHASE 1A: PRODUCTION PRETEST – RECORDING ISOLATED WORDS

This test was conducted in the language laboratory, with the use of a microphone and a recording software named *Adobe Audition 1.5* installed in a computer (OS: 64-bit Windows 7). Each time only at most two subjects were recording (as only two booths were available) and the remaining ones had to wait silently.

The subjects were given a word list with 50 words with target vowels (10 /ɪ/, 10 /iː/, 10 /e/, 10 /æ/, 5 /ʊ/, and 5 /uː/) plus 10 distractors, all arranged in random order so that they were not juxtaposed to remind the subjects of the contrast between the any pairs of vowels. The codas of the words were either plosives or fricatives. The subjects were presented with one word at a time shown on the computer screen to avoid list intonation since all the words were recorded in isolation. The word list is shown in Appendix C.

The target words were monosyllabic words with different CVC structures and would appear in the perception pretest and the training. The instructions of this production pretest were offered to the subjects by recording 10 practice trials. They were reminded to record the tokens with natural loudness and speaking rate. No audio prompts or instructions were provided once the recording began; yet, they could pause and resume recording at their own pace. The test took less than 5 minutes to complete.

3.2.2.2.2 Phase 1b: Perception Pretest – Perceptual Identification Test

Following the production pretest, the subjects then completed the perception pretest – in the form of multiple choices with four answers – in a computer program which was designed by the researcher and coded by a professional programmer. All the computers in the language laboratory were equipped with headphones and were connected to the school's public drive installed with the computer program designed particularly for this experiment which allowed the subjects to complete the perceptual identification test with convenience and the data could be saved systematically and retained accurately. Their answers were saved into a password-protected *Microsoft Access* database immediately when the subjects submitted their answer. This holds the most updated reference of the subjects' scores for further analysis.

At most 10 subjects sat in the laboratory to complete the test. The researcher offered brief instructions to the subjects before the commencement of the test. Prior to the experiment, the subjects were given a username and password to login to their own account so that they would not mix up their results with the others. They entered the main menu page, in which they could start the experiment by finishing the 10 practice trial questions in "Trial" session to get familiarized with the format of test and training. They were also allowed to adjust the amplitude of the stimuli to a comfortable level. After this, the subjects were presented with the "Pretest" stage consisting of 120

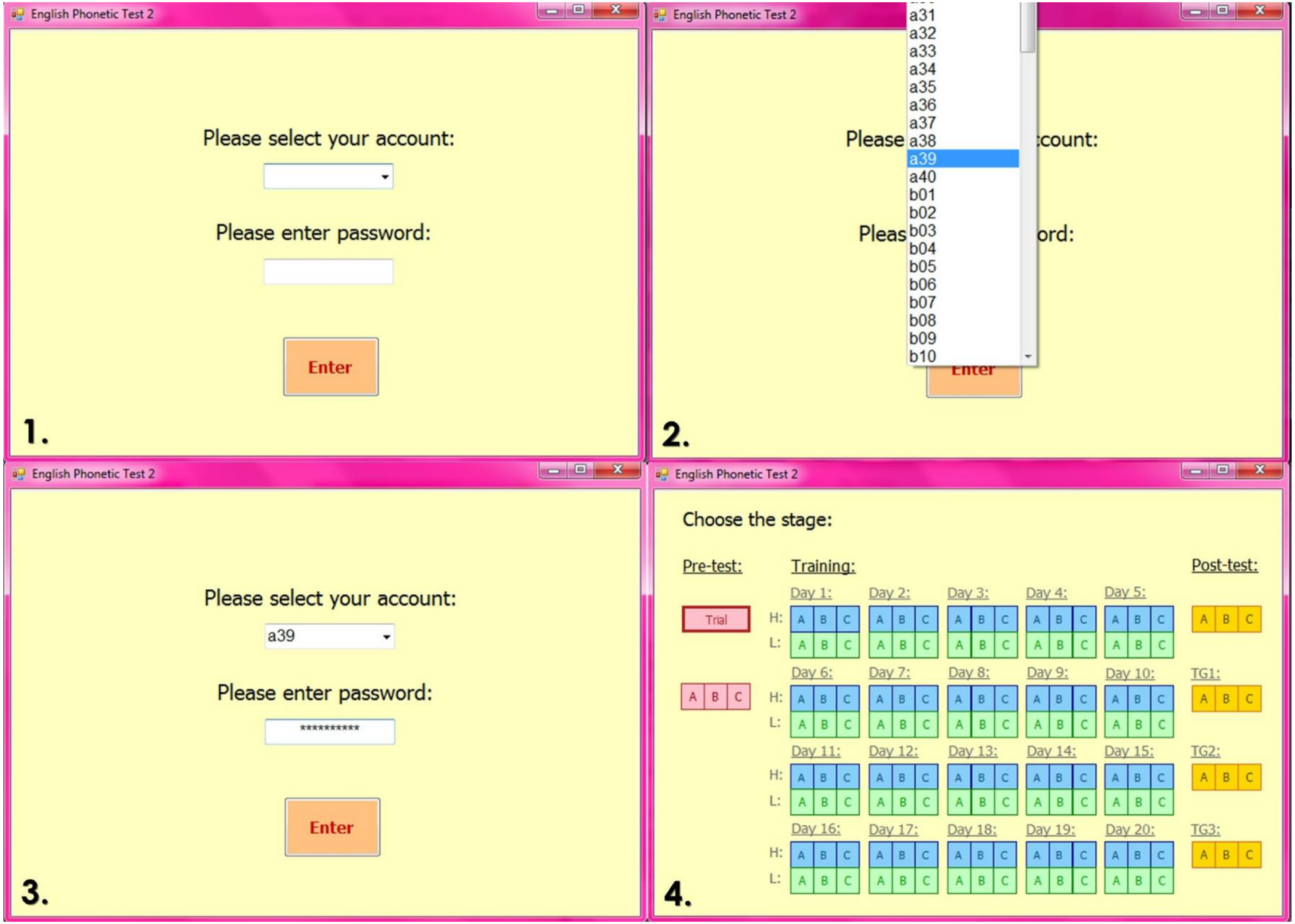
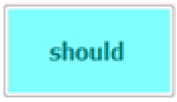


Figure 3.3 Screenshots of the computer program used in the current study

questions in total, spread over three blocks, with 40 /ʊ/-/uː/ stimuli in Block A, 40 /e/-/æ/ stimuli in Block B, and 40 /ɪ/-/iː/ stimuli in Block C (details of the stimuli will be discussed in Section 3.3.1). All the test stimuli were prepared by a male native English speaker with General American accent. The blocks in the "Training" session thereafter also followed the same arrangement. Screenshots of the program are shown on the previous page in Figure 3.3. Four procedures of how a subject could enter the program to finish the experiment were shown there.

When the subjects clicked into one of the blocks of the pretest, they would find an "audio ()" button, three buttons with conventional English orthography (the target minimal word pair and a distracter, for example, 

, 

, and 

) and a blank () in which they could type their own answer. They were reminded that the frequency of occurrence of the correct answer appeared in the four serial positions, i.e. word 1, word 2, word 3, free answer, were equal, so that it allowed the chance level to be correctly and fairly inferred at 25%, while any bias towards or away from the free answer can be eliminated. A screenshot of the test session was shown in Figure 3.4 on the next page.

This specific four-alternative answering test design has not been adopted in previous studies (e.g. Lambacher et al., 2005; Logan et al., 1991; except in my previous work, see Wong, 2010; Wong, 2012) using the training approaches, HVPT and LVPT, and this design was derived to avoid the flaws existing in offering only two-alternative forced choices for the subjects when it is used to evaluate their performance, which has been briefly analyzed in the Literature Review section. This design can thus reveal the

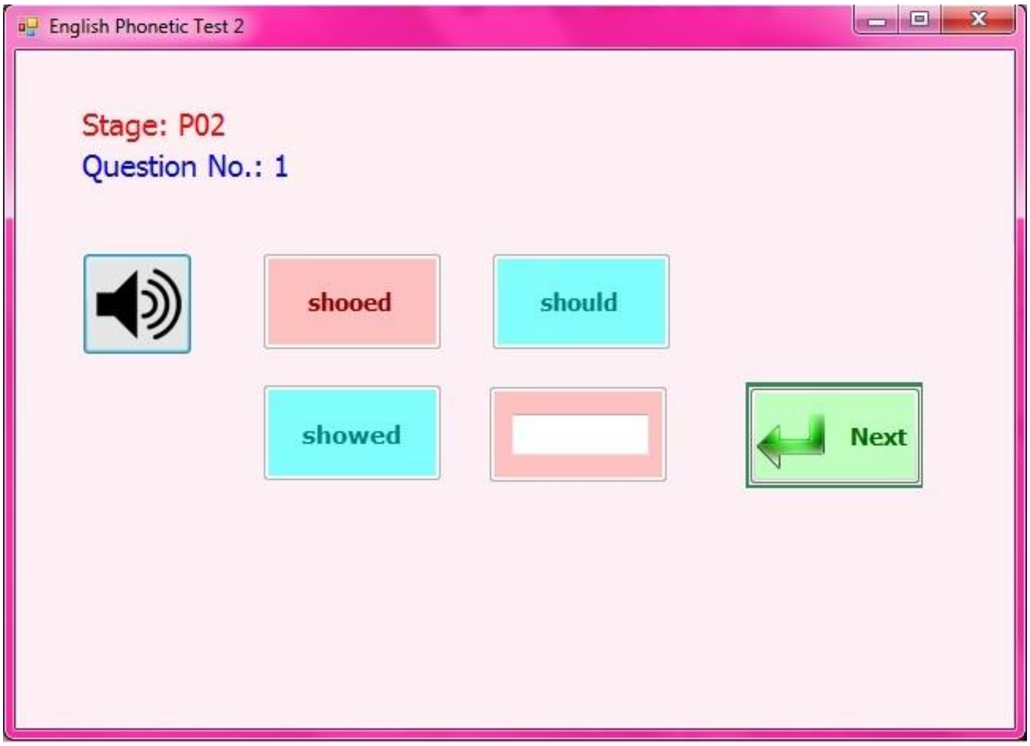


Figure 3.4 Screenshot of the pretest layout in the computer program (test session)

genuine performance of the subjects as it can avoid offering them 50% probability to guess the correct answer even if they could not perceive the sound or that they had different perceptions. A wider variety of choices can hence raise the reliability and the fairness of the data collected. Unlike the tests, the subjects were deliberately provided with only two-alternative forced choices (2AFC) in all training sessions. They would be able to pay more attention to and focus on identifying only the two vowels during training as there were no without interference from other sounds, so that the intensiveness of the training effect can be raised.

Before confirming the answer from the four choices, the subjects could press the "audio ()" button as many times as they needed to listen to the stimulus. No practice effect was expected to deduce due to repeated playing of the stimulus for the perceptual performance of the subjects. The subjects were reminded that they must make the decision before clicking the answer as the program would lock the answer button once they clicked one of the four choices. Then they could click the "Next (

)" button to move on to the next question. At the instance when they clicked the button, the answer was sent to the database and was saved. This real-time update of the scores allowed the researcher to monitor the subjects' progress and avoided data loss if the program was closed unexpectedly. All questions required an answer and the subjects were required to finish all the questions in one block before proceeding to the next. Unlike the training sessions which will be mentioned later, no immediate feedback was given in the test sessions.

3.2.2.3 Phase 2: Training Phase

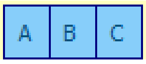
All the subjects (except the control group) received both/either type(s) of training: perceptual training and/or production (explicit articulation) training based on the group they were in. The following sections delineate the details of the training paradigm.

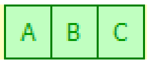
3.2.2.3.1 Perceptual Training

Two perceptual training approaches were adopted in different groups: four groups were trained under the High Variability Phonetic Training (HVPT) approach whereas four other under the Low Variability Phonetic Training (LVPT) approach. Under each paradigm, two groups received the training organized in an intensive level (10 sessions a day, in a total of two days) while two groups with standard training (two sessions a day, in a total of 10 days).

3.2.2.3.1.1 The HVPT Training Procedures

The High Variability Phonetic Training was administered in the form of a two-alternative-forced-choice (2AFC) identification task, in which the subjects received either an intensive or standard form of the training. All subjects received a total of 20 sessions. Each session lasted under 15 minutes. In between each training session, they were given at least 5 minutes for rest. The group receiving intensive perceptual training took around 2.5 hours to complete 10 sessions whereas the normal intensity group spent around 35 minutes to complete 2 sessions in one day.

They had to complete the pre-test before they could proceed to the training since the computer program would inhibit them from joining any training session. On the main menu page of the program, the subjects were instructed to choose either the HVPT (the blue blocks: 

) or the LVPT (the green blocks: 

) according to their assigned group and had to complete the whole series prior to the post-test. Once a color block was chosen, they would not be allowed to change the training paradigm, hence ensuring them to stay with one training approach throughout the experiment. Referring back to Figure 3.2 gives a clearer picture of the program's layout.

A total of 35 subjects (intensive: n = 17; standard n = 18) were trained under the HVPT approach. The same as the perceptual pretest, they were presented a total of 120 stimuli spreading over 3 blocks: 20 /ʊ/ and 20 /uː/ stimuli in Block A, 20 /e/ and 20 /æ/ stimuli in Block B, as well as 20 /ɪ/ and 20 /iː/ stimuli in Block C, all produced by six different native speakers to enhance the stimulus variability (a more detailed account and explanation of how stimulus variability was implemented will be discussed in

Section 3.3.1.2).

Unlike the perceptual pretest, the training session was presented as a 2AFC identification task which was intended to raise the training effect by allowing the subjects to focus on only the pair of vowels. The subjects firstly heard the stimulus played by the program automatically when they were directed to a question. They could choose the answer at that moment; or else they could press the “audio ()” button again to listen to the token before they pressed the answer. They hence had the control over the pace of each training session. Allowing sufficient time for the subjects between successive items let them selectively attend to and elaborate on the acoustic cues without time pressure (Martin et al., 1989). At the instant when they submitted their choice, an immediate feedback (shown on the screen next to the answer, represented as a tick or cross) was given to indicate whether they had identified the target word correctly. Apart from the immediate feedback given after each trial, the subject could be notified of their progress and performance from time to time at the bottom right corner where the cumulative score was shown. Added to this, the subjects were also shown their total scores at the end of each training session. All these measures allowed both the researcher and the subjects to keep track of the training and learning progress. Figure 3.5 on the next page showed the screenshots of the program used during training. A complete list of words chosen in the perceptual training will be provided in Appendix D and the details about how the stimuli were arranged and organized will be introduced in Section 3.3.1.

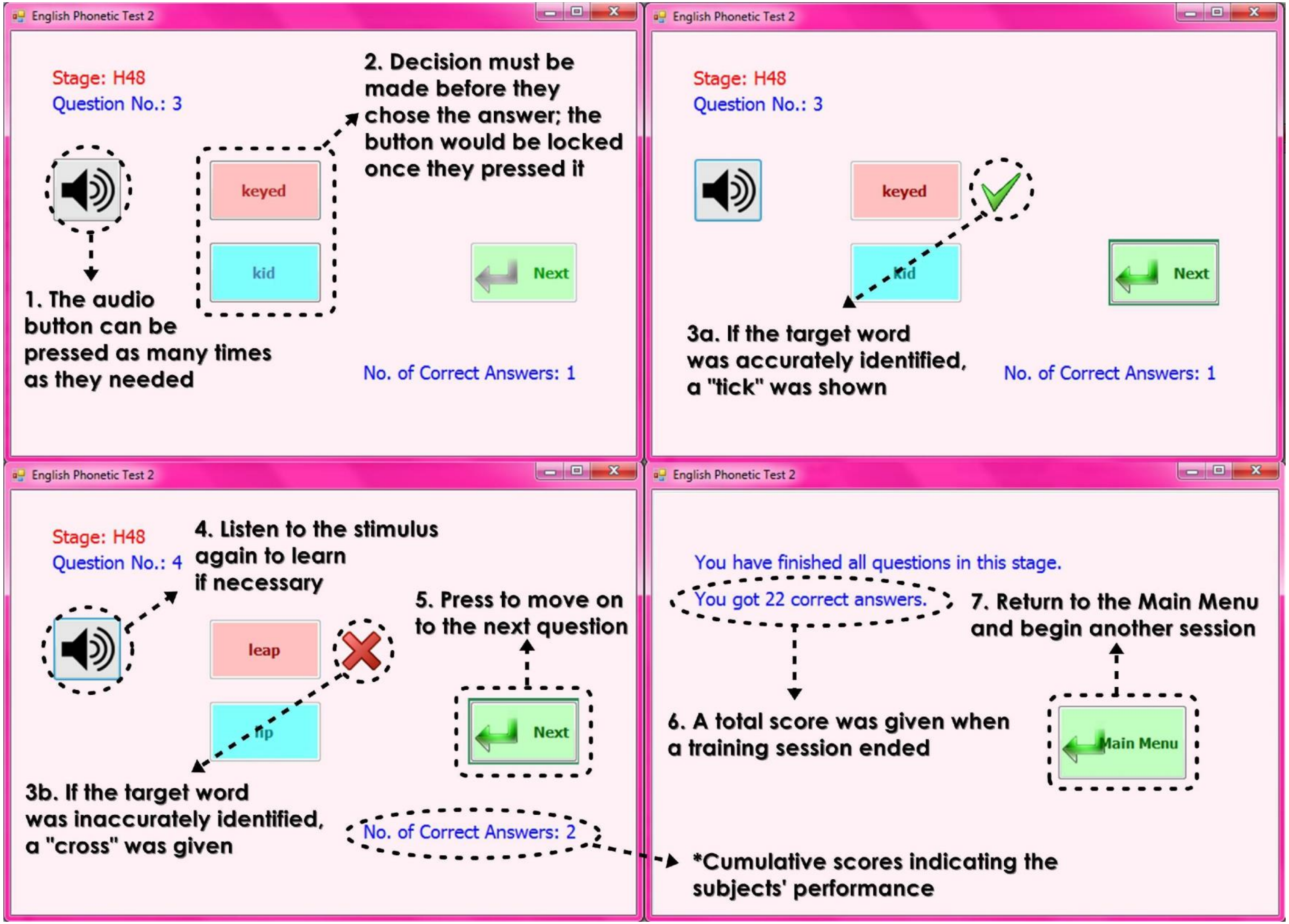


Figure 3.5 Procedures of the perceptual training

3.2.2.3.1.2 The LVPT Training Procedures

Twenty-nine subjects (intensive: n = 15; standard: n = 14) participated in the LVPT training paradigm. The subjects used the same computer program to complete the test, except that all training stimuli were produced by only one male native speaker of English. The training procedures followed those adopted in the HVPT training: the subjects were trained to perceive of the three vowel pairs by using an identification task with immediate feedback.

3.2.2.3.2 Production Training

The production training was administered in the form of explicit articulation training. A total of 43 subjects participated in the training. The subjects, in a group of four, were provided with 4 video clips adapted online and modified by the researcher in which a native speaker demonstrated the articulation of the vowels. The subjects had to watch the videos first and imitate the minimal pair words after the video host. Three word lists (each with at least 10 target minimal pair words) and short passages targeting the vowel contrasts were given to the subjects and they pronounced after the researcher at least three times and corrective feedback was given. Articulation information of the vowel pairs, i.e. the tongue height/vowel openness, tongue frontness/backness as well as the length of the vowels, was also emphasized in the each session explicitly with pictures as illustrations and mirrors to help the subjects better articulate.

Each session lasted for 30 minutes and each subject had participated in four sessions of production training. It would be ideal to give the same number of production training sessions as the perceptual training sessions, i.e. 10 sessions; but training in the production aspect through learning to accurately articulate, receiving corrective feedback from the researcher and practiced, as well as describing how the vowels should be produced, took more than 10 minutes to complete, which was the averaged time of a perception training session. Thus, four 30-minute training sessions instead of ten 12 or 15-minute sessions were given to optimize the learning environment. Details concerning the videos and word lists are shown in Appendix E.

3.2.2.4 Phase 3: Posttest

Immediately after the training, all trainees participated in the posttest phase which involved two parts: the production posttest and the perception posttest. The control group also took the posttests at about the same timeframe of the trainees. Besides directly comparing the pretest-posttest results to reflect the training efficacy, a series of generalization tests was also administered to reveal the external validity of the training; hence, a Test of Contextualization and 3 Tests of Generalizations were added to the testing regime. Same as the pretest, the production posttests were all completed before the perception posttests to avoid cueing effects.

3.2.2.4.1 Phase 3a: Production Posttest

3.2.2.4.1.1 Production Posttest: Recording Isolated Words

All 85 subjects recorded the same list of isolated words used in the pretest. The procedures were also the same as those adopted in the pretest.

3.2.2.4.1.2 Test of Contextualization (TC): Passage Reading

Immediately after recording the posttest wordlist, the subjects participated in Test of Contextualization (TC) in which they were asked to record a passage named "April Fools" with around 500 words. The passage was adapted from a teachers' resources website online (Hashway, n.d.). It had simple content and no difficult vocabulary items for the subjects so that they were not distracted or disrupted by the unfamiliar words and meanings during the recording. With the researcher's adaptations to suit the research purpose, the passage included content words at the elementary or intermediate level with the target vowels: 15 /ɪ/, 15 /iː/, 15 /e/, 15 /æ/, 10 /ʊ/ and 10 /uː/. The length of the passage was also shortened so as to sustain the subjects' interest and attention. The words were either monosyllabic words or multisyllabic in various phonological environments, with all the target vowels placed in the strong syllable with primary stress. Below shows an excerpt of the passage, "April Fools" (full version in Appendix F):

April Fools

If you've ever been pranked on April Fools' Day, you may wonder how this tradition started. Well, you're not alone. No one knows for sure how April Fools' Day began, but the most likely explanation has to do with the calendar.

No, that's not an April Fools' Day joke. People used to celebrate New Year's Day on April 1st. Just like today, people would choose to have big parties and cook some good food to celebrate. Over time, the calendar changed and so did the date for New Years. In the Fifteen Hundreds, the new calendar put New Year's Day as January 1st. But because there was no Internet or other means to spread the word, the news traveled slowly by word of mouth. It took a while for everyone to speak to others about the change, but some people are not happy with it. They looked at it as a bad push. Due to this reason, they carried on celebrating New Years on April 1st. These people were given the cool nickname, "April fools."

Figure 3.6 An excerpt of the passage, April Fools, used in the Test of Contextualization

The subjects were given around 10 minutes to get familiarized with the passage. A sample recording read by a female native speaker who recorded also the wordlist sample were given to the subjects as well. They could listen to the sample recording prior to the test. During recording, they were reminded to read naturally, at their own pace and loudness.

3.2.2.4.2 Phase 3B: Perception Posttest

3.2.2.4.2.1 Perception Posttest: Identification Test

This test was the same as the perception pretest aforementioned. All the procedures followed those adopted in the perception pretest. All the subjects had to participate in the perception posttest and the three Tests of Generalization.

3.2.2.4.2.2 Test of Generalization 1 (TG1)

In TG1, the subjects listened to three blocks of new words (40 for each vowel pair) spoken by a new speaker whose voice was not heard in any of the training or tests. The procedures were the same as those administered in the perception pretest and posttest (i.e. using the same software) in which the subjects were given four choices to choose from.

3.2.2.4.2.3 Test of Generalization 2 (TG2)

Different from TG1, in TG2, the subjects had to listen to three blocks of new words (40 for each vowel pair) spoken by a familiar speaker, who had been one of the speakers in the training stimuli (i.e. the one who produced all the tokens in the LVPT).

Procedures were the same as those in TG1.

3.2.2.4.2.4 Test of Generalization 3 (TG3)

In the last perception generalization test, The subjects revisited some familiar words which they had come across in the perception training sessions before; but the words were produced by a new speaker (40 for each vowel pair). Again, the procedures were the same as those in TG1 and TG2.

3.3 Materials

3.3.1 Perceptual Training/Test Stimuli

3.3.1.1 Preparation of Stimuli

Table 3.3 summarizes the background information of the native speakers who contributed to the perceptual training stimuli:

Table 3.3

Gender, Age and Accents of Speakers A-F, X and Z

Speaker	Gender	Age (range)	Accent
A	Female	25-30	North American
B	Female	30-35	General American
C	Female	20-25	British English
D	Male	30-35	General American
E	Male	25-30	Canadian English
F	Male	40-45	General American
G	Female	20-25	British English
X	Female	20-25	British English
Z	Female	20-25	General American

Since stimulus variability is stressed in the HVPT paradigm, the stimuli in the training were produced by 1) increasing the *speaker variability* through using natural tokens produced by 6 different native English speakers (three male and three female); 2) increasing the *accent variability* by adopting speakers with different British or American accents; and 3) increasing the *phonetic variability* by using tokens in different phonetic environments. A total of nine native speakers of English with various backgrounds participated as the speakers for the test or training stimuli.

The accents were particularly chosen since they were more comprehensible and prevalent among Hong Kong students as the education system widely promotes the use of these two accents, American and British, as *standards* or *models*. Although previous findings (Escudero & Chládková, 2010) have reported that the L2 perception initial stage will be different if the variety that the learners are exposed to is different, it is speculated that the present study will not show heave bias on either accent, due to the fact the participants all have wide exposure to the two varieties in their daily life.

Speakers A-F contributed to both the perceptual pre/posttest tokens and the perceptual training stimuli. All of them had produced 20 minimal pairs with both /ɪ/-/iː/ and /e/-/æ / vowels and 5 minimal pairs with /ʊ/-/uː/vowels (the amount of minimal pairs for /ʊ/-/uː/was extremely scarce, hence the words were used repeatedly throughout the experiment). As the tokens were also used in the LVPT, Speaker D recorded all the minimal word pairs. All words are monosyllabic in various consonant-vowel-consonant (CVC) contexts with different onsets and codas. To provide a base set for the training and avoid co-articulatory effects on the vowels, the researcher discarded using the minimal pair tokens with approximants or nasals as codas; hence all codas of the training instances were plosives, with a few fricatives as exceptions. To balance the effects of voiced and voiceless codas on the length of the preceding vowel, a balanced number of voiced and voiceless codas were adopted.

Speaker G also recorded the same set of minimal pairs, but the productions were utilized in TG3 (familiar words by a new speaker). Speaker D, i.e. a familiar speaker to the subjects, recorded another word list for TG2 (new words by a familiar speaker). The list included words with 20 /ɪ/-/iː/, 20 /e/-/æ/ and 5 /ʊ/-/uː/ minimal pairs which had not appeared in the training sessions. Speaker X who had not recorded anything for the training stimuli or the pre/posttests, was therefore, a new speaker to all the subjects, recorded another new list with 20 /ɪ/-/iː/, 20 /e/-/æ/ and 5 /ʊ/-/uː/ minimal pairs for TG1 (new words by a new speaker). All the minimal pairs for TG1, TG2 and TG3 were with various consonant-vowel-consonant (CVC) contexts and syllable structures (mono-, di- and poly-syllabic) with a view to testing the transfer of learning under various conditions (c.f. pre/posttest or training, tokens were only monosyllabic). Speaker Z recorded the audio prompts in production pre/posttest. The word lists in TG1, TG2 and TG3 can be found in Appendix G.

All words were recorded in isolation. To avoid list intonation, each speaker was given one word on a computer screen at a time to record. They also read the tokens at least three times so as to avoid intra-speaker variability in vowel productions, as reported in previous study (Munro, Wang & Li, 2000). It avoided the use of a single token per speaker for the stimuli as well. The three tokens for each word were adopted in the training program randomly while all the words in each training session were also arranged randomly. It could avoid any speaker effect that might arise when the subjects were only exposed to the stimuli produced by only one speaker at a time, as some studies (e.g. Logan et al., 1991; Lively et al., 1993; Lively et al., 1994) had reported on the influence of the speaker effect when the stimuli produced were arranged according to the speaker. The speaker effect would cause the subjects to obtain speaker-specific information rather than acoustic cues of the vowels, which distorted the original purpose of the training program.

The speakers were reminded to read at normal speed and loudness. All the stimuli were made by the native speakers reading into a headphone-mounted microphone connected to a computer (OS: 64-bit Windows 7) with *Adobe Audition 1.5* for digitalization. The words were saved separately, digitalized at a sampling rate of 44.1 kHz and 16-bit resolution, normalized for peak amplitude before incorporating them into the computer software. All minimal pairs were evaluated by three other native English speakers to confirm the quality and intelligibility of the words. Minimal word pairs with /e/-/æ/ among which no difference was found in the production of some American accents were excluded.

3.3.1.2 Training Stimuli Organization

In both the HVPT and LVPT, the training stimuli were grouped into three blocks: Block A with /ʊ/-/uː/ stimuli, Block B with /e/-/æ/ stimuli and Block C with /ɪ/-/iː/ stimuli. Particularly for the HVPT which involves tokens produced by multiple speakers, the tokens appeared randomly in different sessions, but not grouped by the speakers. Previous studies showed that the identification of isolated words was affected by speakers' voices and whether the words are produced by only one speaker or more (Mullennix, Pisoni & Martin, 1998; Pisoni & Lively, 1995). The present study thus grouped the tokens under vowels so that the subjects could attend to six voices in a single session and adapt to the underlying properties of the vowels. Moreover, so as to allow a wider exposure to token variability, each stimulus was read at least by two different speakers and they appeared in the training program with equal frequency. The vowel pairs were blocked so as to increase subjects' attention when learning to identify the difference between a pair, but not among all target vowels.

Each block contains 40 tokens. In Block B (/e/-/æ/) and C (/ɪ/-/iː/), the 40 tokens were the 20 minimal word pairs (20 pairs × 2 × 1 repetition = 40 stimuli), i.e. each word appears once in a session; in Block A (/ʊ/-/uː/), due to limited pairs existing in English, the 40 tokens are made up of only 4 minimal word pairs, each pair appears 5 times (4 pairs × 2 × 5 repetitions = 40 stimuli).

3.3.2 SURVEY FORMS

All the subjects were given a survey form prior to the experiment. They filled in their personal information, language background, English learning experience and daily English usage and habit, etc. so that the researcher can ensure that their language backgrounds have no significant difference for fair analysis and comparison of data. Together with a consent form, the researcher obtained the rights to collect their personal information and data for research purposes. The complete survey form together with the consent form is shown in Appendix A and a summary of the subjects' language background information was in Appendix H.

3.4 Data Analysis

3.4.1 Vowel Identification Data

Prior to statistical analysis, all identification data in tests and training sessions were transferred to SPSS and variables such as group types, test, training intensity and vowels were identified and set. Several statistical tests were performed to answer all the research questions.

3.4.2 Vowel Production Data

With a view to obtaining accurate and objective evaluation of the subjects' production performance and investigating their individual differences, the present study employed certain procedures including phonetic transcription, acoustic analysis as well as native speaker's judgment and rating tests to achieve this aim.

3.4.2.1 Phonetic Transcription

The production data of the pretests and posttests of all the subjects were transcribed in narrow phonetic transcription by the researcher with Cantonese a L1 and English as L2 twice. Two other phonetically-trained researchers were also invited to check the inter-rater reliability by each doing 23.5% of the transcriptions, i.e. 20

subjects. They also had Cantonese as L1 and English as L2 and reported no hearing or speaking deficits. The transcribers were of similar phonetic training backgrounds so that a consensual validity of the transcription can be yielded (Perry, 2005; Shriberg, Kwiatkowski & Hoffmann, 1984). Productions which conformed to the target vowels were deemed as accurate. The intra-rater reliability was 93.27% while the inter-rater reliability reached 91.74% and 92.62% respectively.

3.4.2.1.1 Evaluation of Transcription Reliability

3.4.2.1.1.1 Aim

Monitoring the consistency of the transcribed data is always crucial to the reliability and validity of the presentation and analysis of the subjects' production performance as even phonetically transcribed data "may not necessarily be an accurate reflection of a subject's actual speech output" (Van Borsel, 1989). Hence, in the present study, Cronbach's alpha was used to examine the reliability of transcriptions: the higher the alpha coefficient, the more consistent the data. It is central to setting the threshold of accurate production performance analysis.

3.4.2.1.1.2 Intra-rater Reliability

The productions of all the subjects were transcribed twice by the researcher with Cantonese as L1 and English as L2. The interval between each transcription was about one to one and a half months. The original transcription was not referred to when the second transcription was done. The intra-rater reliability was calculated by using the total number of target productions produced by all the subjects in the second transcription divided by the first trial of transcription. The reliability was 93.27%. Also, to further determine the transcription consistency, Cronbach's alpha of the two transcriptions was computed and the alpha coefficient was .898. It implies a very high consistency in the intra-rater judgment.

3.4.2.1.1.3 INTER-RATER RELIABILITY

About 23.5% of the subjects (20 subjects) were randomly selected for each inter-rater, making a total of 20 subjects × 2 = 40 subjects. Their production recordings were transcribed for an inter-rater reliability check. Subjects' productions, in MPEG-1 Audio Layer 3 format, were given to two phonetically-trained researchers in Applied English Linguistics with Cantonese as L1 and English as L2 to transcribe the data phonetically. The reliability check was completed without referring to any completed transcriptions and it was 91.74% and 92.62%. The Cronbach's alpha was also obtained at $\alpha = .807$ and $\alpha = .821$ respectively. These figures were high and the disagreement in transcriptions was reached between the raters upon discussions on the transcriptions.

Although both the intra-and inter-rater reliabilities were high and consensus were reached upon discrepancies of transcriptions, the transcribers were still non-native speakers of English. So as to uplift the validity of the transcriptions, acoustic analysis and selective native speaker judgment tests were conducted to check whether both transcriptions align with the other measures. The details of the acoustic analysis and native judgment tests will be introduced in the following sections.

3.4.2.2 Native Speaker's Judgment and Rating Test

What is of interest to the present study with regard to the effectiveness of the training regimes is whether the productions of minimal words pairs were improved to become closer to native speakers' productions and more intelligible perceptually by native speakers. Besides using phonetic transcriptions, production results of the present research were also assessed by a native speaker's judgment. One monolingual native English speaker who was raised in an American family in Hong Kong was the judge of this task. She was familiar with the International Phonetic Alphabet and reported no hearing or speaking deficits. This judgment task was aimed to supplement the transcriptions by confirming their accuracies and showing additional information on the intelligibility of the productions.

Three tasks were given to the judge. Among the pretest and posttest production recordings of 20 subjects (two subjects randomly picked from each group, i.e. 2 subjects × 10 groups = 20 subjects), the judge had to 1) transcribe orthographically the token presented; 2) transcribe phonetically; and 3) rate on how similar the production was to the vowel. The 9-point scales on the next page were utilized in the judgment task.

The rating was quantified and the two extremes, 1 and 9, represented the target and exact productions. The native speaker, after listening to each production made by the subjects, could rate on how close she deemed the production was to a particular vowel according the scale. The productions would be discarded and excluded from the rating only when the vowel produced was not one of the six targets (i.e. a misreading of the vowel), meaning that the judge would stop at the second task, i.e. phonetic transcription,

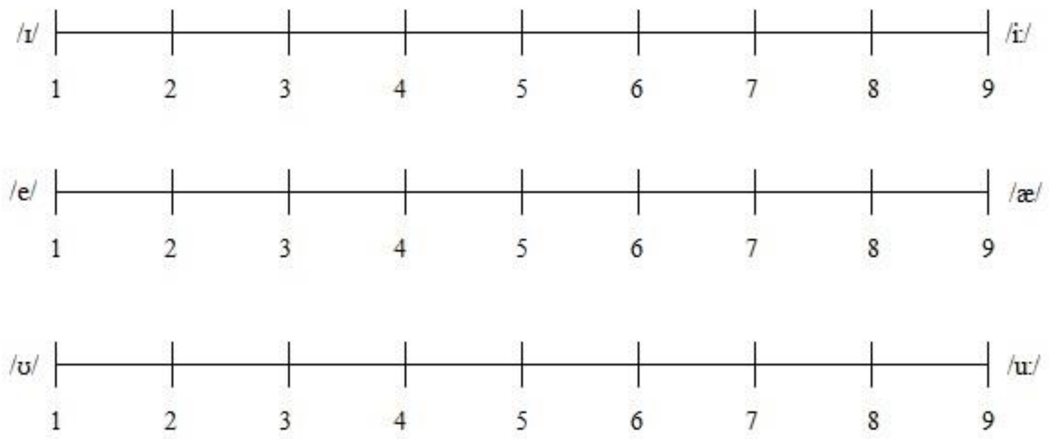


Figure 3.7. Three 9-point scales adopted in the native speaker rating task

without giving the rating. Upon comparison with the transcription data and discussion with the native judge, the researcher confirmed that phonetic transcriptions highly aligned with the native speaker's judgment. Table 3.4 on the next page summarizes the rating results. For vowels /ɪ/, /e/ and /ʊ/, the smaller the number, the more accurate the production; for vowels /iː/, /æ / and /uː/, the larger the number, the more accurate the production. Hence, a negative difference between pretest and posttest means improvement for vowels /ɪ/, /e/ and /ʊ/, whereas a positive difference denotes an improvement for vowels /iː/, /æ / and /uː/.

3.4.2.3 ACOUSTIC ANALYSIS

So as to avoid any subjective perceptions from the researchers' transcriptions and to reveal the productions' actual acoustic properties, the data was finally analyzed acoustically by using the *Praat* speech analysis software (Boersma & Weenink, 2005), which gives spectrometric representations of the sounds. By using this software, the

Table 3.4

Summary of Rating of the 6 Target Vowels in Native Speaker's Rating Task (positive improvement denoted in boldface)

Groups	/ɪ/		Difference
	Pretest	Posttest	
HSP	6.70	1.00	**-5.70**
HIP	6.20	2.60	**-3.60**
HSN	4.30	2.00	**-2.30**
HIN	2.70	1.20	**-1.50**
LSP	4.00	3.70	**-0.30**
LIP	5.00	3.30	**-1.70**
LSN	1.10	1.00	**-0.10**
LIN	3.80	1.00	**-2.80**
COP	7.80	1.70	**-6.10**
CON	6.20	5.80	**-0.40**

Groups	/e/		Difference
	Pretest	Posttest	
HSP	7.50	7.30	**-0.20**
HIP	6.60	5.90	**-0.70**
HSN	7.11	3.00	**-4.11**
HIN	6.80	4.40	**-2.40**
LSP	6.50	4.80	**-1.70**
LIP	6.78	3.11	**-3.67**
LSN	6.60	4.90	**-1.70**
LIN	8.20	4.20	**-4.00**
COP	7.50	4.10	**-3.40**
CON	5.10	6.30	1.20

Groups	/ʊ/		Difference
	Pretest	Posttest	
HSP	5.80	4.20	**-1.60**
HIP	4.40	3.00	**-1.40**
HSN	4.00	2.80	**-1.20**
HIN	3.60	2.80	**-0.80**
LSP	4.40	3.00	**-1.40**
LIP	4.80	4.40	**-0.40**
LSN	3.40	4.20	0.80
LIN	5.40	6.40	1.00
COP	6.20	2.60	**-3.60**
CON	3.00	3.25	0.25

Groups	/i:/		Difference
	Pretest	Posttest	
HSP	5.90	9.00	**3.10**
HIP	8.20	9.00	**0.80**
HSN	7.44	6.30	-1.14
HIN	8.60	8.70	**0.10**
LSP	8.00	7.90	-0.10
LIP	7.90	8.10	**0.20**
LSN	9.00	8.90	-0.10
LIN	5.40	7.90	**2.50**
COP	8.90	9.00	**0.10**
CON	7.80	7.50	-0.30

Groups	/æ/		Difference
	Pretest	Posttest	
HSP	6.10	8.40	**2.3**
HIP	6.10	9.00	**2.9**
HSN	6.90	8.90	**2**
HIN	6.50	9.00	**2.5**
LSP	6.44	9.00	**2.56**
LIP	6.50	8.20	**1.7**
LSN	6.40	8.60	**2.2**
LIN	5.20	8.90	**3.7**
COP	5.40	8.60	**3.2**
CON	4.78	5.11	**0.33**

Groups	/u:/		Difference
	Pretest	Posttest	
HSP	8.50	9.00	**0.50**
HIP	4.20	9.00	**4.80**
HSN	6.80	8.20	**1.40**
HIN	8.20	9.00	**0.80**
LSP	8.60	7.40	-1.20
LIP	7.80	7.60	-0.20
LSN	7.60	8.80	**1.20**
LIN	6.60	7.40	**0.80**
COP	8.80	8.80	**0.00**
CON	6.00	6.00	**0.00**

first two formant frequencies (F1, F2) and the temporal measurements of the vowels were gauged at the midpoint to evaluate how similar or different the vowel productions were after training. Productions by the subjects were also used to compare to those produced by the native English speakers invited for the recording of training stimuli. This could allow a contrastive analysis and suggest whether the posttest productions had become more native-like after participating in the training program. Moreover, by analyzing the data acoustically, the production performance of the subjects could be measured as a follow-up to the transcription reliability and become some of the evidence for or against the effectiveness of the training approaches. Detailed acoustic illustrations and analyses will be discussed in Chapter 4.

CHAPTER 4

RESULTS AND FINDINGS

4.0 INTRODUCTION

The current study aims at testing and comparing the effectiveness of the High Variability Phonetic Training (HVPT) approach, the Low Variability Phonetic Training (LVPT) approach and explicit articulation training approach on training the perception and production of English vowel contrasts among Cantonese ESL learners. Perceptual training intensity and the effects of different combinations of the above training paradigms are also of central concern in the present investigation. With reference to the research insights and methodology outlined in previous chapters, this chapter reports the subjects' perceptual and production performances obtained in the experiment. This prepares the ground work to answer all six research questions.

The first two research questions – whether the perceptual training approaches are effective (RQ1) and how effective they are (RQ2) – will be answered by comparing and contrasting the pretest and the posttest results of the trained groups and the control group. How the results were varied when an addition of explicit production training was provided (RQ3) will also be answered, hence shedding light on the effects of a combination of a perception and a production training program (RQ4). In addition to these, the effect of training intensity of the perceptual paradigm (RQ5) as well as the generalization effects of the training approaches (RQ6) will be investigated. Last but not least, whether there is any difference in the training of the three pairs of vowels will also be examined (RQ7). There are two major sections in this chapter: Section 4.1 will report the results of the perceptual performance whereas 4.2 will report the results of the production performance. Section 4.3 will provide a conclusion of the findings.

4.1 Perceptual Performance

This section presents the results of the perceptual performances of the subjects. Results of statistical analysis on different factors will also be reported. Then an overview of the general performance will be shown in 4.1.1 and 4.1.2. After presenting a justification of the choice of statistical analysis in 4.1.3, I will start with delineating the statistical results of the effects of the perceptual training approaches in 4.1.4, of which the various factors in investigation will be discussed under different sub-sections. Section 4.1.5 will focus on the effects of different factors on the generalizability of the training approaches. This section presenting the results of perceptual learning will come to a close with 4.1.6 which summarizes the entire section.

4.1.1 Preview of the Overall Perceptual Performance

Table 4.1 on the next page gives an overview of the results of the statistical analyses that will be presented in detail in the coming sections. All the factors that are involved in the present investigation were fed into different statistical procedures to verify the effect and power of the data. To let readers have a basic understanding of the data gathered, I have displayed clearly whether a specific effect or interaction was significant before actual statistics was presented.

TABLE 4.1

Overview of the Statistical Results of the Effects of Different Factors on Perceptual Learning

Factors in investigation	**Results**	**Statistically significant?**
A. Training effectiveness (Pretest vs. Posttest)		
1. Perceptual Training (HVPT vs. LVPT vs. Control)	a. HVPT ~ LVPT b. HVPT > Control c. LVPT > Control	✗ ✓ ✓
2. Perceptual Training Intensity (Standard vs. Intensive)	a. HVPT: Standard ~ Intensive b. LVPT: Standard ~ Intensive	✗ ✗
3. Production Training (With vs. Without)	With > Without	✓
4. Training Diversity (Perceptual Training vs. Perceptual Training + Production Training vs. Production Training)	a. HVPT > Production only b. HVPT + Production > Production only c. LVPT > Production only d. LVPT + Production > Production only	✓ ✓ ✓ ✓
B. Generalization (Pretest vs. TG1/TG2/TG3)		
1. Perceptual Training (HVPT vs. LVPT vs. Control)	a. HVPT > LVPT b. HVPT > Control c. LVPT > Control	✓ ✓ ✓
2. Perceptual Training Intensity (Standard vs. Intensive)	HVPT: Standard ~ Intensive LVPT: Standard ~ Intensive	✗ ✗
3. Production Training (With vs. Without)	With > Without	✓
4. Training Diversity (Perceptual Training vs. Perceptual Training + Production Training vs. Production Training)	In TG1: a. HVPT > LVPT b. HVPT > LVPT + Production c. HVPT > Production only d. HVPT + Production > LVPT e. HVPT + Production > LVPT + Production f. HVPT + Production > Production only In TG2: a. HVPT > Production only b. HVPT + Production > Production only In TG3: a. HVPT > Production only b. HVPT + Production > Production only	 ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓

✓ = Yes

✗ = No

4.1.2 Overall Perceptual Performance of All 10 Groups

The overall results of perceptual identification performance of the experimental and control groups are shown in Tables 4.2 a, b and c on the next several pages (for a detailed version with individual subject's scores and percentages, see Appendix I). The tables display the averaged score and mean percentage of correct identification in all the five perceptual tests, the pretest, the posttest, the Test of Generalization 1 (TG1), the Test of Generalization 2 (TG2) and the Test of Generalization 3 (TG3) of the three vowel pairs across all 10 groups. In general, all the eight experimental groups receiving perceptual training improved their identification of all three pairs of vowels.

Figures 4.1 a, b and c further illustrate a comparison of the perceptual test results across groups. For front vowel pair /ɪ/-/iː/, all four groups trained under the HVPT[10] paradigm had an improvement of 11% over, with two groups having again over 20%; the four LVPT groups also improved after training, with three of them improving around 15% and one group around 22%; the group with production training had a slight gain of 3.41% while the control group had very similar performance from the pretest to the posttest.

The identification performance of the low vowel pair /e/-/æ/ improved over 20% among all the perceptually trained groups: the four HVPT groups improved from 25% to 36%; the four LVPT groups correctly identified over 22% of tokens more, with two of them improving over 30% and one doing exceptionally well after training for an

---

[10] Four groups trained under HVPT (with other modifications): HSP, HIP, HSN, HIN; four groups trained under LVPT (with other modifications): LSP, LIP, LSN, LIN; one group with production training only: COP; and the control group: CON. Details can be recalled by reading the first three columns in Table 4.2.

TABLE 4.2a

Mean Scores and Mean Percentage of the Identification of the Vowel Pair /ɪ/-/iː/ *in 5 Perceptual Tests (Pretest, Posttest, Test of Generalization 1, 2 & 3) across different Groups*

						Perceptual Identification of /ɪ/-/iː/									
	Group Name	Perceptual Training	Training Intensity	Production Training	Total Score:	Pretest 40	%	Posttest 40	%	TG1 40	%	TG2 40	%	TG3 40	%
Trained	HSP	HVPT	Standard	Yes	*Average:*	30.00	75.00%	34.67	86.67%	31.67	79.17%	31.89	79.72%	33.00	82.50%
					S.D.:	*2.00*	*5.00%*	*3.32*	*8.29%*	*3.67*	*9.19%*	*5.56*	*13.89%*	*3.04*	*7.60%*
	HIP	HVPT	Intensive	Yes	*Average:*	27.78	69.44%	36.78	91.94%	32.11	80.28%	31.89	79.72%	34.89	87.22%
					S.D.:	*3.23*	*8.08%*	*1.30*	*3.25%*	*2.52*	*6.31%*	*4.83*	*12.08%*	*4.81*	*12.02%*
	HSN	HVPT	Standard	No	*Average:*	27.78	69.44%	35.33	88.33%	33.89	84.72%	33.00	82.50%	33.11	82.78%
					S.D.:	*2.54*	*6.35%*	*3.43*	*8.57%*	*2.52*	*6.31%*	*4.47*	*11.18%*	*4.94*	*12.34%*
	HIN	HVPT	Intensive	No	*Average:*	28.00	70.00%	36.50	91.25%	34.88	87.19%	33.00	82.50%	34.38	85.94%
					S.D.:	*2.62*	*6.55%*	*1.60*	*4.01%*	*1.64*	*4.11%*	*3.02*	*7.56%*	*4.50*	*11.25%*
	LSP	LVPT	Standard	Yes	*Average:*	28.14	70.36%	34.71	86.79%	29.14	72.86%	31.29	78.21%	33.57	83.93%
					S.D.:	*2.12*	*5.29%*	*4.42*	*11.06%*	*3.44*	*8.59%*	*4.39*	*10.97%*	*5.00*	*12.49%*
	LIP	LVPT	Intensive	Yes	*Average:*	27.43	68.57%	33.71	84.29%	26.71	66.79%	29.00	72.50%	33.71	84.29%
					S.D.:	*3.91*	*9.77%*	*3.55*	*8.86%*	*3.04*	*7.60%*	*4.00*	*10.00%*	*3.95*	*9.87%*
	LSN	LVPT	Standard	No	*Average:*	29.86	74.64%	36.43	91.07%	28.86	72.14%	31.00	77.50%	32.71	81.79%
					S.D.:	*1.35*	*3.36%*	*3.31*	*8.27%*	*5.01*	*12.54%*	*3.83*	*9.57%*	*4.23*	*10.58%*
	LIN	LVPT	Intensive	No	*Average:*	27.63	69.06%	34.25	85.63%	27.38	68.44%	29.50	73.75%	31.38	78.44%
					S.D.:	*4.00*	*9.99%*	*1.98*	*4.96%*	*2.62*	*6.54%*	*2.73*	*6.81%*	*2.13*	*5.33%*
Control	COP	Control	N/A	Yes	*Average:*	27.27	68.18%	28.64	71.59%	24.45	61.14%	26.18	65.45%	28.27	70.68%
					S.D.:	*5.14*	*12.85%*	*4.70*	*11.74%*	*5.18*	*12.96%*	*5.51*	*13.78%*	*5.20*	*12.99%*
	CON	Control	N/A	No	*Average:*	27.40	68.05%	26.90	67.25%	22.40	56.00%	26.00	65.00%	26.20	65.50%
					S.D.:	*2.27*	*5.68%*	*1.91*	*4.78%*	*2.55*	*6.37%*	*3.09*	*7.73%*	*4.21*	*10.53%*

TABLE 4.2b

Mean Scores and Mean Percentage of the Identification of the Vowel Pair /e/-/æ/ *in 5 Perceptual Tests (Pretest, Posttest, Test of Generalization 1, 2 & 3) across different Groups*

	Group Name	Perceptual Training	Training Intensity	Production Training	Total Score:	Perceptual Identification of /e/-/æ/ Pretest 40	%	Posttest 40	%	TG1 40	%	TG2 40	%	TG3 40	%
Trained	HSP	HVPT	Standard	Yes	*Average:*	23.89	59.72%	36.78	91.94%	30.44	76.11%	31.00	77.50%	32.78	81.94%
					S.D.:	*3.86*	*9.64%*	*2.59*	*6.47%*	*2.19*	*5.46%*	*4.69*	*11.73%*	*5.74*	*14.35%*
	HIP	HVPT	Intensive	Yes	*Average:*	22.89	57.22%	37.44	93.61%	34.89	87.22%	33.33	83.33%	35.44	88.61%
					S.D.:	*5.21*	*13.02%*	*1.59*	*3.97%*	*2.26*	*5.65%*	*1.12*	*2.80%*	*2.83*	*7.08%*
	HSN	HVPT	Standard	No	*Average:*	25.44	63.61%	35.56	88.89%	28.00	70.00%	30.56	76.39%	31.44	78.61%
					S.D.:	*3.61*	*9.02%*	*3.75*	*9.36%*	*4.58*	*11.46%*	*4.53*	*11.33%*	*6.97*	*17.42%*
	HIN	HVPT	Intensive	No	*Average:*	23.88	59.69%	34.88	87.19%	33.13	82.81%	33.13	82.81%	34.00	85.00%
					S.D.:	*4.29*	*10.73%*	*3.52*	*8.81%*	*1.55*	*3.88%*	*2.30*	*5.74%*	*3.55*	*8.86%*
	LSP	LVPT	Standard	Yes	*Average:*	24.43	61.07%	38.00	95.00%	26.71	66.79%	29.43	73.57%	32.57	81.43%
					S.D.:	*3.26*	*8.15%*	*1.73*	*4.33%*	*2.36*	*5.90%*	*2.64*	*6.59%*	*4.61*	*11.53%*
	LIP	LVPT	Intensive	Yes	*Average:*	25.14	62.86%	36.86	92.14%	26.71	66.79%	29.71	74.29%	30.43	76.07%
					S.D.:	*2.12*	*5.29%*	*2.04*	*5.09%*	*2.87*	*7.18%*	*3.40*	*8.50%*	*4.65*	*11.62%*
	LSN	LVPT	Standard	No	*Average:*	23.86	59.64%	37.00	92.50%	26.57	66.43%	31.00	77.50%	32.14	80.36%
					S.D.:	*6.59*	*16.48%*	*2.71*	*6.77%*	*2.30*	*5.75%*	*3.00*	*7.50%*	*5.73*	*14.32%*
	LIN	LVPT	Intensive	No	*Average:*	23.13	57.81%	35.63	89.06%	25.88	64.69%	29.38	73.44%	29.25	73.13%
					S.D.:	*5.33*	*13.33%*	*3.66*	*9.16%*	*3.23*	*8.07%*	*6.28*	*15.69%*	*7.78*	*19.45%*
Control	COP	Control	N/A	Yes	*Average:*	22.45	56.14%	26.82	67.05%	24.73	61.82%	26.09	65.23%	28.73	71.82%
					S.D.:	*6.07*	*15.18%*	*4.12*	*10.30%*	*3.07*	*7.67%*	*7.85*	*19.64%*	*6.74*	*16.85%*
	CON	Control	N/A	No	*Average:*	22.10	55.25%	19.40	48.50%	23.30	58.25%	22.80	57.00%	23.90	59.75%
					S.D.:	*3.75*	*9.39%*	*4.70*	*11.74%*	*2.00*	*5.01%*	*4.61*	*11.53%*	*4.51*	*11.27%*

TABLE 4.2c

Mean Scores and Mean Percentage of the Identification of the Vowel Pair /ʊ/-/uː/ in 5 Perceptual Tests (Pretest, Posttest, Test of Generalization 1, 2 & 3) across different Groups

						Perceptual Identification of /ʊ/-/uː/									
	Group Name	Perceptual Training	Training Intensity	Production Training	Total Score:	Pretest 40	%	Posttest 40	%	TG1 40	%	TG2 40	%	TG3 40	%
Trained	HSP	HVPT	Standard	Yes	*Average:*	27.67	69.17%	32.22	80.56%	30.44	76.11%	30.33	75.83%	27.22	68.06%
					S.D.:	*4.66*	*11.66%*	*2.99*	*7.48%*	*2.74*	*6.86%*	*4.56*	*11.39%*	*2.54*	*6.35%*
	HIP	HVPT	Intensive	Yes	*Average:*	27.22	68.06%	36.00	90.00%	32.78	81.94%	31.78	79.44%	31.56	78.89%
					S.D.:	*3.90*	*9.75%*	*1.87*	*4.68%*	*2.77*	*6.93%*	*5.59*	*13.96%*	*4.33*	*10.83%*
	HSN	HVPT	Standard	No	*Average:*	27.22	68.06%	33.33	83.33%	33.44	83.61%	32.56	81.39%	32.11	80.28%
					S.D.:	*4.21*	*10.52%*	*3.28*	*8.20%*	*2.19*	*5.46%*	*5.32*	*13.29%*	*1.62*	*4.04%*
	HIN	HVPT	Intensive	No	*Average:*	26.50	66.25%	34.25	85.63%	32.50	81.25%	33.50	83.75%	32.13	80.31%
					S.D.:	*3.30*	*8.24%*	*3.69*	*9.23%*	*1.60*	*4.01%*	*5.83*	*14.58%*	*1.89*	*4.71%*
	LSP	LVPT	Standard	Yes	*Average:*	26.29	65.71%	37.71	94.29%	25.57	63.93%	31.57	78.93%	30.29	75.71%
					S.D.:	*2.14*	*5.35%*	*2.98*	*7.46%*	*4.43*	*11.07%*	*6.85*	*17.13%*	*4.31*	*10.77%*
	LIP	LVPT	Intensive	Yes	*Average:*	27.14	67.86%	33.71	84.29%	26.57	66.43%	27.00	67.50%	26.86	67.14%
					S.D.:	*4.41*	*11.03%*	*4.23*	*10.58%*	*2.37*	*5.93%*	*7.02*	*17.56%*	*3.13*	*7.83%*
	LSN	LVPT	Standard	No	*Average:*	27.43	68.57%	36.86	92.14%	30.29	75.71%	34.29	85.71%	29.14	72.86%
					S.D.:	*5.35*	*13.37%*	*4.91*	*12.28%*	*2.50*	*6.24%*	*3.64*	*9.10%*	*4.56*	*11.40%*
	LIN	LVPT	Intensive	No	*Average:*	25.50	63.75%	35.25	88.13%	26.88	67.19%	26.38	65.94%	29.25	73.13%
					S.D.:	*3.42*	*8.56%*	*4.59*	*11.48%*	*3.76*	*9.40%*	*5.42*	*13.56%*	*3.37*	*8.43%*
Control	COP	Control	N/A	Yes	*Average:*	25.55	63.86%	27.45	68.64%	21.82	54.55%	25.55	63.86%	24.64	61.59%
					S.D.:	*3.88*	*9.71%*	*2.21*	*5.52%*	*4.64*	*11.61%*	*5.45*	*13.62%*	*3.80*	*9.50%*
	CON	Control	N/A	No	*Average:*	26.90	67.25%	25.40	63.50%	22.80	57.00%	24.30	60.75%	24.50	61.25%
					S.D.:	*2.18*	*5.46%*	*1.84*	*4.59%*	*2.66*	*6.65%*	*5.98*	*14.96%*	*2.84*	*7.10%*

improvement of 45.71%; the subject trained under production training also increased around 10% of correct identification while the control group had a slight decrease of around 2% in performance. The back vowel pair /ʊ/-/uː/ had a similar improvement pattern in terms of the subjects' identification across groups: the increase in correct identification of the four HVPT groups varies from around 11% to 22%; the LVPT groups' improvement ranged from 16.5% to 31.5%; the group trained under the production paradigm only had a slight gain of around 5% whereas the control group regressed for 8%.

In general, all nine experimental groups demonstrated learning after training, while the control group had similar performance throughout all the tests, showing the overt-time fluctuation of performance of subjects without intervention. No definite pattern could be observed when comparing the groups receiving production training on top of the perceptual training, with the groups receiving only perceptual training. Yet, the perceptually trained groups with intensive HVPT training were consistently performing better than groups with standard HVPT training in all target vowel pairs; the LVPT groups did not have a similar performance pattern: LSP group in particular performed exceptionally well in contrast to other LVPT groups and was the group which gained the most among all nine experimental groups.

One of the objectives of the present study is to investigate whether learning can be promoted to new words and/or new speakers, i.e. if there is generalization effect of the learning. Intrigued as it may seem, the results of the three generalizations test – TG1 where stimuli were new and were produced by new speaker, TG2 where stimuli were

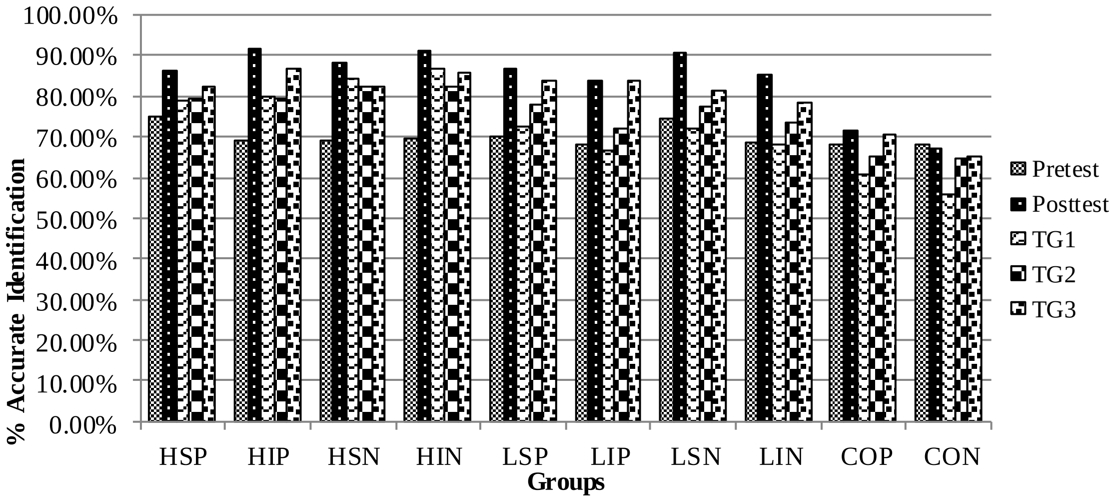


Figure 4.1a Mean percentages of correct perceptual identification performance of the vowel pair /ɪ/-/iː/ for the 9 training groups and control subjects in the pretest, the posttest and three Tests of Generalization (TG1, TG2 & TG3)

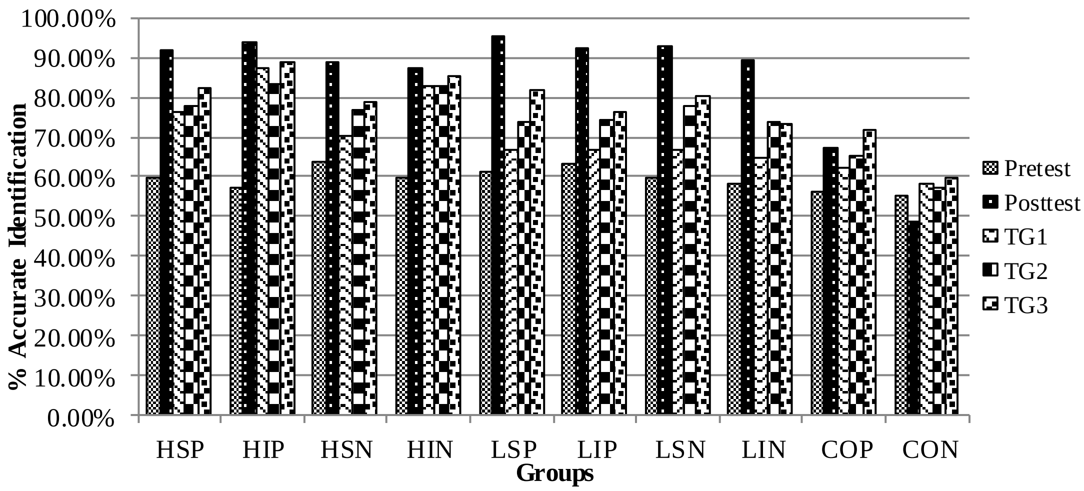


Figure 4.1b Mean percentages of correct perceptual identification performance of the vowel pair /e/-/æ/ for the 9 training groups and control subjects in the pretest, the posttest and three Tests of Generalization (TG1, TG2 & TG3)

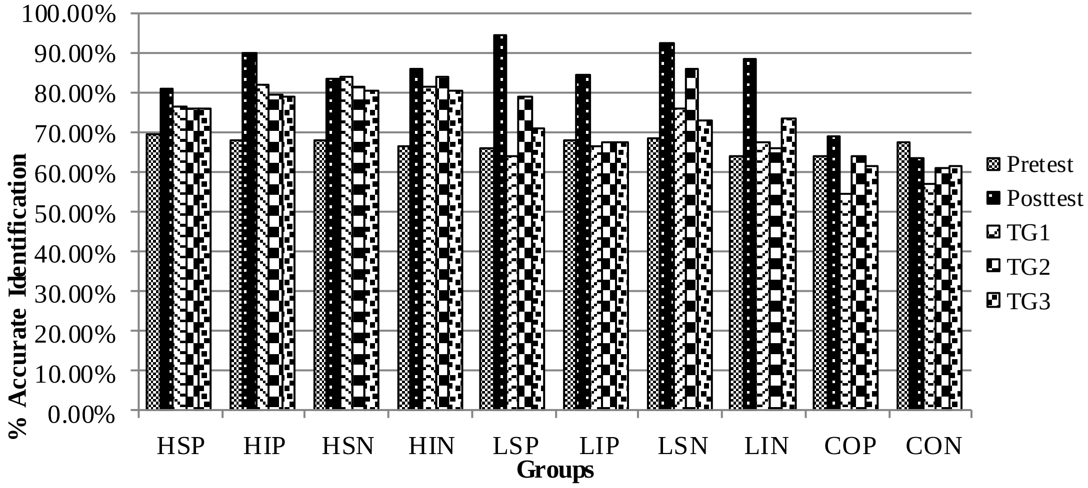


Figure 4.1c Mean percentages of correct perceptual identification performance of the vowel pair /ʊ/-/uː/ for the 9 training groups and control subjects in the pretest, the posttest and three Tests of Generalization (TG1, TG2 & TG3)

new yet were produced by a familiar speaker and TG3 where stimuli were the same tokens used in pre/posttests but were produced by a new speaker – demonstrated a learning pattern in different groups. Generally speaking, almost all experimental groups performed better, if not similar, in both TG2 and TG3 than in TG1 for all three pairs of vowels. Groups that received the HVPT approach consistently outperformed all groups that there trained under the LVPT paradigm, suggesting that the generalization effects could be observed only among some experimental groups. Since the group received only production training showed very slight gain in after treatment, generalization effects were barely expected. The results confirmed this speculation. The control group again displays similar results in all TGs and posttest.

To sum up, all the HVPT and the LVPT groups showed improvement after the perceptual training in terms of the posttest identification accuracies and the percentage of improvement of the three pairs of vowels; yet it appears that the HVPT groups had better performance in the generalization tests. The production training group showed slight improvement after training, but the learning was not generalizable to new conditions. These findings have offered a general picture of the subjects' performance, while statistical computing in the following sections allows a more thorough analysis and understanding of the data can be garnered through a series of statistical analyses.

To know whether the various intervention designs play a part in modifying the perceptual patterns of the subjects, I will start with comparing the results in perceptual pretest and posttest with different factors taken into account, namely perceptual training types, training intensity production training and training diversity. These provide a consistent measure of performance across group, time and vowel pairs under different conditions. With a view to better understanding the effects of training and design on the generalization of learning, there will be further analyses on the results of three Tests of Generalization so as to shed more light on the subject matter.

4.1.3 STATISTICAL ANALYSIS: 5-WAY ANOVA VS. 3-WAY ANOVA

The present study involves a total of five major factors with different levels: Test (5 levels: pretest, posttest, TG1, TG2, TG3), Vowel (front: /ɪ/-/iː/; low: /e/-/æ/; back: /ʊ/-/uː/), Perceptual Training Type (HVPT, LVPT, none), Production Training Intensity (Standard, Intensive) and Production Training (with, without). It will be ideal if all the factors are included in one analysis, i.e. a 5-way mixed analysis of variance (ANOVA) so as to see the interactions and the relation among them. However, due to the difficulty in presentation and the limitation of small subject size that will not result in enough power for the five factors, the 5-way ANOVA has been avoided as the major analysis of the current study. Also, given that a preliminary analysis of all 5 factors showed that they are similar to the results of several 3-way ANOVAs[11], the current study presents the 3-way ANOVA as the major exposition which will be more appropriate to interpret and understand the results. The same analysis applies to the production aspect in 4.2. All the significance levels in multiple comparisons were also adjusted using Bonferroni.

[11] The 5-way mixed ANOVA was conducted with Perceptual Training Type (HVPT, LVPT, none), Production Training Intensity (Standard, Intensive) and Production Training (with, without) as between-subject factors and Test (5 levels: pretest, posttest, TG1, TG2, TG3) and Vowel (front: /ɪ/-/iː/; mid: /e/-/æ/; back: /ʊ/-/uː/) as repeated measures is shown in Appendix J for readers' reference.

4.1.4 Results of Perceptual Learning (Pretest vs. Posttest)

4.1.4.1 Factor 1: Perceptual Training Approaches

To reveal the efficacy of the two perceptual training approaches - HVPT and LVPT - on the effect of perceptual learning, the mean identification accuracy scores of the three vowel pairs in both the perceptual pretest and posttest across the two perceptually trained groups as well as the control group were compared directly. Several ANOVAs were computed to gauge the statistical significance. Table 4.3 displays the averaged scores and the respective percentages in the pretest and posttest across the HVPT (n = 17), the LVPT (n = 15) and the control group (n = 10).

Table 4.3

Mean Test Scores, Mean Percentage, and Mean Difference of all the groups in Perceptual Pretest and Posttest

		Perceptual Performance					
		Pretest	%	Posttest	%	Difference	%
	Total Score:	40		40			
/ɪ/-/iː/	HVPT	28.40	71.00%	35.80	89.50%	7.40	18.50%
	S.D.:	2.69	6.73%	2.66	6.64%		
	LVPT	28.24	70.60%	34.76	86.90%	6.52	16.30%
	S.D.:	3.09	7.72%	3.36	8.39%		
	Control	27.33	68.33%	27.81	69.52%	0.48	1.19%
	S.D.:	3.94	9.85%	3.67	9.17%		
/e/-/æ/	HVPT	24.03	60.07%	36.20	90.50%	12.17	30.43%
	S.D.:	4.20	10.49%	3.01	7.52%		
	LVPT	24.10	60.26%	36.83	92.07%	12.72	31.81%
	S.D.:	4.50	11.25%	2.68	6.71%		
	Control	22.28	55.71%	23.28	58.21%	1.00	2.50%
	S.D.:	4.98	12.45%	5.73	14.32%		
/ʊ/-/uː/	HVPT	27.17	67.93%	33.94	84.86%	6.77	16.93%
	S.D.:	3.90	9.76%	3.21	8.02%		
	LVPT	26.55	66.38%	35.86	89.66%	9.31	23.28%
	S.D.:	3.85	9.63%	4.31	10.77%		
	Control	26.31	65.78%	26.48	66.19%	0.16	0.41%
	S.D.:	3.19	7.97%	2.25	5.62%		

A three-way ANOVA with Perceptual Training Type (HVPT, LVPT and Control) as between-subject factor, Test (pretest and posttest), and Vowel (front, low and back) as repeated-measures revealed a significant main effect of Perceptual Training Type [$F(2,39) = 56.00$, $p < .001$], Test [$F(1,39) = 393.84$, $p < .001$] and Vowel [$F(2,78) = 22.25$, $p < .001$], due to the overall identification accuracy of the two trained groups among all three pairs of vowels from the pretest to the posttest. There were also significant two-way interactions between Test × Perceptual Training Type [$F(2,39) = 70.78$, $p < .001$], Test × Vowel [$F(2,78) = 18.99$, $p < .001$] and Perceptual Training Type × Vowel [$F(4,78) = 4.59$, $p = .002$], indicating that the perception accuracy of different perceptual training types and vowels increased significantly after training. The three-way interaction between Perceptual Training Type, Test and Vowel [$F(4,78) = 4.20$, $p = .003$] was significant as well.

Planned comparisons with Bonferroni correction on the interaction between Test and Perceptual Training Type revealed a significant effect of Test within both the HVPT [$F(1,39) = 329.80$, $p < .001$] and the LVPT [$F(1,39) = 321.01$, $p < .001$] groups, but not in the control group ($p = .35$), indicating that the 21.95% and 23.79% of increase observed in the HVPT and LVPT groups respectively were robust. Significant differences in the posttest between HVPT > control for 23.64% [$F(1,54) = 184.0$, $p < .001$], LVPT > control for 24.90% [$F(1,48) = 167.14$, $p < .001$] were observed as well, but not in the pretest (HVPT vs. control: $p = .075$; LVPT vs. control: $p = .178$) . Nonetheless, the difference between the HVPT and the LVPT groups in the posttest was not significant ($p = .422$), meaning that no claim can be established for which approach is more effective at this stage. This will be discussed in detail in Chapter 5.

Post-hoc pairwise comparisons using the Bonferroni correction on the Test ×

Vowel interaction showed a significant difference between vowels in the pretest: front > low vowels (mean difference = 12.00%, $p < .001$), low < back vowels (mean difference = 7.91%, $p < .001$), front > back vowels (mean difference = 3.38%, $p < .001$). The same did not happen in the posttest: front vs. low vowels ($p = .311$), low vs. back vowels ($p = 1.00$), front vs. back vowels ($p = .342$), indicating that the starting points of accurate identification of the three vowels were different, but the accuracy scores of the three vowel pairs did not differ significantly after training. There were significant increases from the pretest to posttest in each vowel as well: front vowels [$F(1,39) = 131.10$, $p <.001$], low vowels [$F(1,39) = 252.62$, $p <.001$] and back vowels [$F(1,39) = 108.66$, $p <.001$] due to the improvement of the sum of accuracy scores of all trained groups.

For the Perceptual Training Type × Vowel interaction, post-hoc pairwise comparison using the Bonferroni correction shows that the subjects found the low vowel pair the most difficult to identify, although only the LVPT group and the control group showed some significant differences in the accuracy scores (LVPT: front > back for 3.86%, $p = .017$; front > low for 4.96%, $p = .002$; Control:; front > low for 11.96%, $p < .001$; back > low for 8.87%, $p < .001$).

A further analysis on the three-way interaction, Perceptual Training Type × Test × Vowel, gives more details on how the three groups of subjects performed before and after training in each pair of vowels. The interaction also provides understanding of the performance of each vowel across groups and time. Figure 4.2 on the next page shows three box-and-whisker plots[12] visualizing the distribution of the perceptual performance

[12] The upper and lower boundaries of the box represent the upper quartile and the lower quartile of the

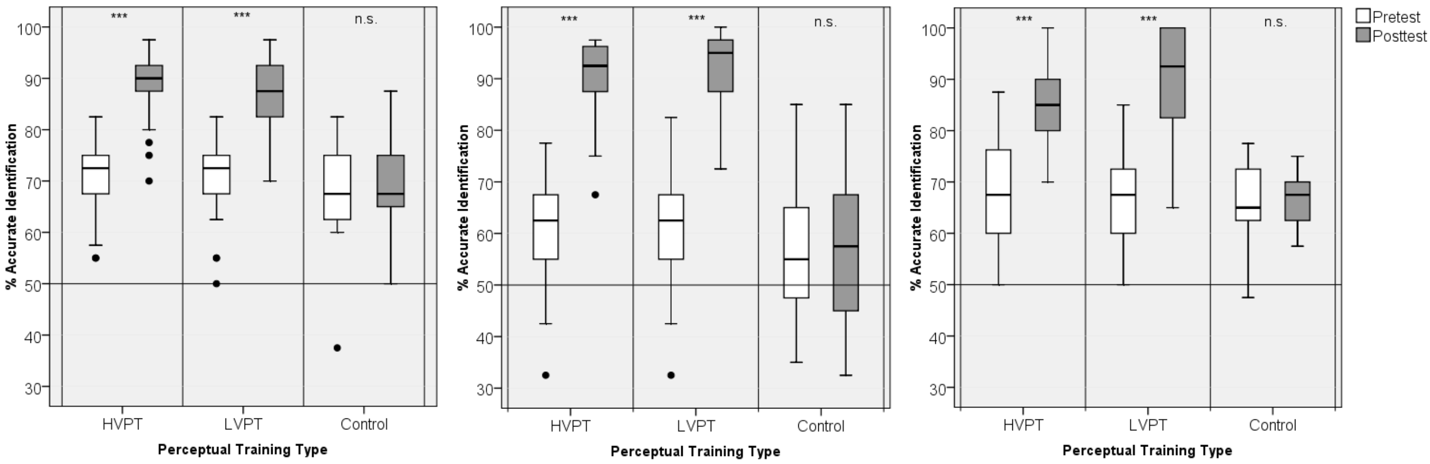


Figure 4.2 (From left to right) Mean percentages of accurate identification of the trained groups with significant difference between the pretest (white boxes) and the posttest (dark boxes) [*** = $p < .001$] and non-significant difference between the pretest and the posttest (n.s. = $p > .05$) in the control group for vowel pairs, /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/. The black dots represent outliners. The horizontal line indicates the chance level (50%) performance.

of all three groups of subjects across the three vowel pairs respectively in both the pretest and the posttest.

Planned comparisons with Bonferroni correction on the three-way interaction revealed that for the front vowel pair, both the HVPT and the LVPT improved significantly from the pretest to the posttest: HVPT for 18.5% [$F(1,39) = 134.23$, $p < .001$] and LVPT for 16.29% [$F(1,39) = 86.26$, $p < .001$], but the control group showed no significance ($p = .565$). Concerning the low vowel pair, the same pattern was observed: HVPT improved for 30.43% from the pretest to posttest [$F(1,39) = 216.18$, $p < .001$], LVPT increased the identification accuracy for 31.81% [$F(1,39) = 195.76$, $p < .001$] whereas the control group showed no significant difference ($p = .352$). Both

---

performance of that test in the groups respectively, while the line inside the box corresponds to the median. The two extended whiskers from the box indicate the maximum and minimum values of the perceptual scores of the groups.

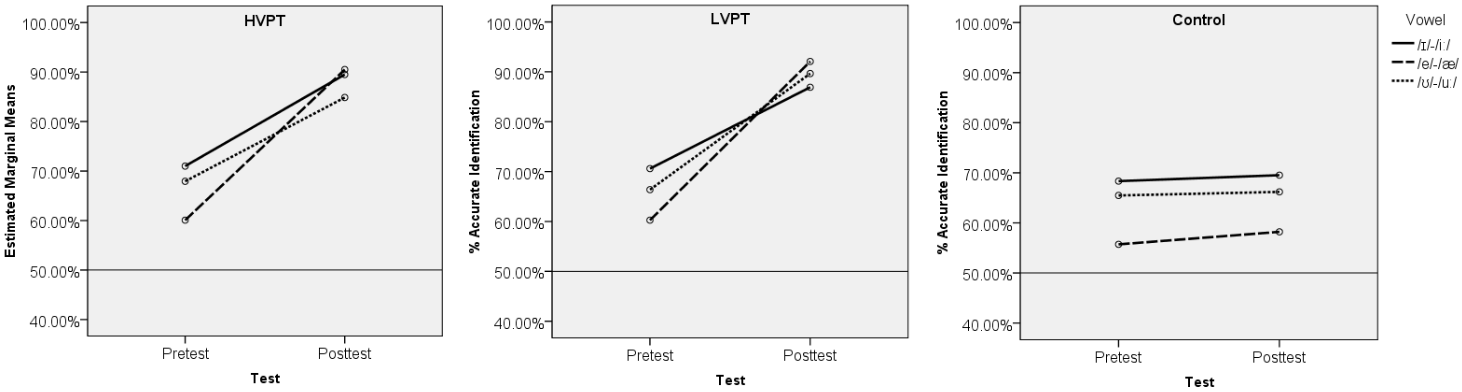


*Figure 4.3 (*From left to right) Mean percentages of accurate identification of the HVPT, the LVPT and the control groups in the pretest and the posttest. The three separate lines represent each target vowel pair. The horizontal line indicates the chance level (50%) performance.

training groups also improved the identification of the back vowel pair after training, with HVPT for an increase of 16.93% [$F(1,39) = 72.04$, $p < .001$] and LVPT for an increase of 23.28% [$F(1,39) = 112.85$, $p < .001$] but the control group had no significant improvement ($p = .782$).

The plots above illustrate the interaction between the performance of each group of each vowel from the pretest and the posttest. Each plot represents the performance of the HVPT, LVPT and control groups respectively. The solid and dotted lines are the accurate identification of each vowel pair from the pretest to the posttest. The slopes of the lines are crucial for the understanding of the performance of the subjects: as the steeper the positive slopes are, the better the subjects had performed. Both the HVPT and LVPT groups improved the identification of the three vowel pairs from the pretest to posttest whereas the control group had only very slight changes. Meanwhile, both training groups had the most improvement in the low vowel pair, as shown by the steepest slopes observed. All the above results indicate that both HVPT and LVPT paradigms facilitate the perceptual learning of the three vowels to different extents.

4.1.4.2 FACTOR 2: PERCEPTUAL TRAINING INTENSITY

Perceptual training intensity is one exploratory element that is central to the investigation of the present study. Figures 4.4a and b display the percentage of accurate identification obtained by the HVPT and LVPT groups with different intensity levels across the three target vowel pairs. It is very clearly shown in the figures that groups with different intensities have highly similar performance.

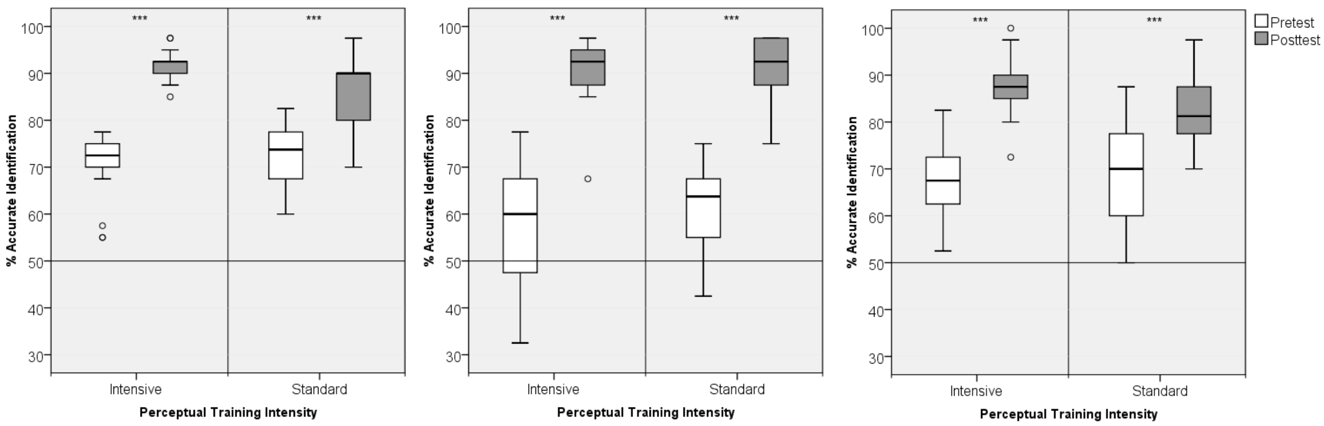


Figure 4.4a Mean percentages of accurate identification of /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right) by the HVPT with different intensity levels. All show significant differences between the pretest (white boxes) and posttest (dark boxes) performance (*** = $p < .001$). The circles represent outliners. The horizontal line indicates the chance level (50%) performance.

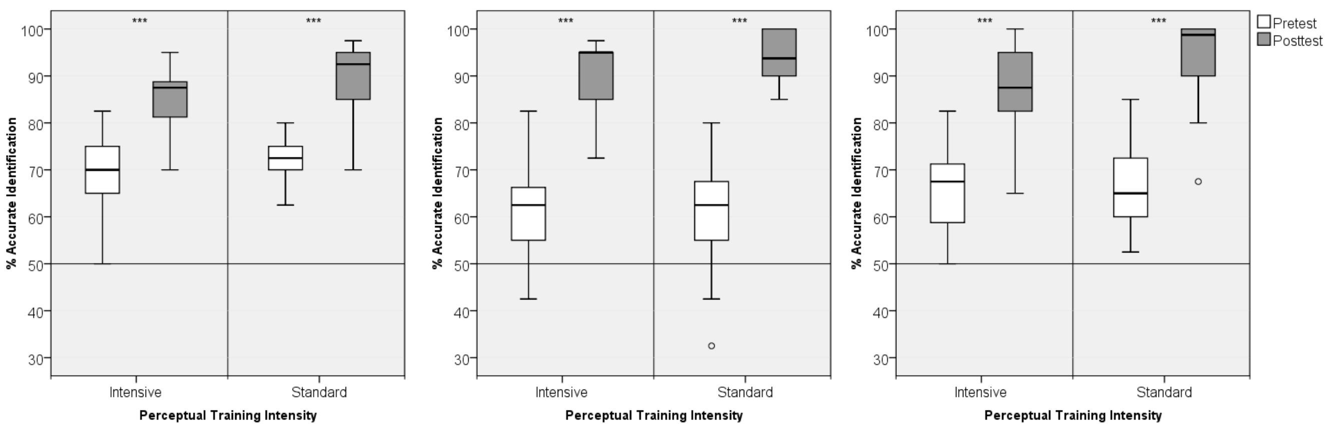


Figure 4.4b Mean percentages of accurate identification of /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right) by the LVPT with different intensity levels. All show significant differences between the pretest (white boxes) and posttest (dark boxes) performance (*** = $p < .001$). The circles represent outliners. The horizontal line indicates the chance level (50%) performance.

Separate three-way repeated-measures ANOVAs with Perceptual Training Intensity (standard and intensive) as between-subject factor, Test (pretest and posttest), and Vowel (front, low and back) as repeated-measures were carried out on the HVPT and LVPT groups' perceptual identification scores. For the HVPT groups (Intensive, n = 9; Standard, n = 8), the ANOVA yielded a significant main effect of Test [$F(1,15) = 156.20$, $p < .001$] and Vowel [$F(2,30) = 3.59$, $p = .040$] and Test × Vowel interaction [$F(2,30) = 3.45$, $p = .045$], as reported in the previous section that both groups improved the identification of vowels from the pretest to the posttest. However, the main effect of Intensity ($p = .903$), Test × Intensity ($p = .408$) Vowel × Intensity ($p = .498$) and Test × Vowel × Intensity ($p = .957$) were not significant. It suggests perceptual training intensity does not play a significant role in the perceptual learning among the two HVPT groups: the group with intensive training had an averaged increase of 22.71% whereas the group with standard training had an averaged improvement of 19.82%.

The same situation happens on the LVPT groups (Intensive, n = 8; Standard, n = 7). The ANOVA revealed only a significant main effect of Test [$F(1,13) = 119.10$, $p < .001$], and an interaction of Test × Vowel [$F(2,26) = 4.70$, $p = .018$], but the main effect of Vowel ($p = .180$), Intensity ($p = .084$), Test × Intensity ($p = .961$), Vowel × Intensity ($p = .877$) and Test × Vowel × Intensity ($p = .970$) were not robust. Again, perceptual training intensity does not significantly influence the modification of the perception of the three vowel pairs as the two LVPT groups with different intensity levels performed similarly in the pretest and the posttest: the group with intensive training had an improvement of 24.06% and the group with standard training had an increase of

24.29%.

From the above analyses, it appears that the training itself is already effective for the perception of the three vowel pairs of subjects. Intriguing as it may seem, the finding is not in line with a common view that learning the perception of L2 sounds should be spread over a period of time, meaning that the present study instead showed that intensified training can also benefit an L2 learner in the learning of a non-native contrast. Studies in the field of speech language pathology indicated that whether intensive intervention is more beneficial still remains unequivocal although a number of studies have provided positive evidence that intensified interventions are more favorable. Although no definite conclusion can be drawn at this stage, it is still evident that perceptual training intensity does not have much impact on the modification of the perception of the vowel pairs. Further explication and explanation will be discussed in the next chapter.

4.1.4.3 FACTOR 3: EXPLICIT PRODUCTION TRAINING

Besides the effects of perceptual training on the perception of non-native contrast, the training effects of the explicit production training is also a focus in the present investigation. Figure 4.5 on the next page shows the interquartile range of percent identification scores obtained by the subjects trained under the production training (n = 11) and the control group (n = 10) across the three target vowel pairs. As can be seen from the figures, the two groups showed similar performance in both the pretest and the posttest across the three vowel pairs. The identification scores were submitted to a

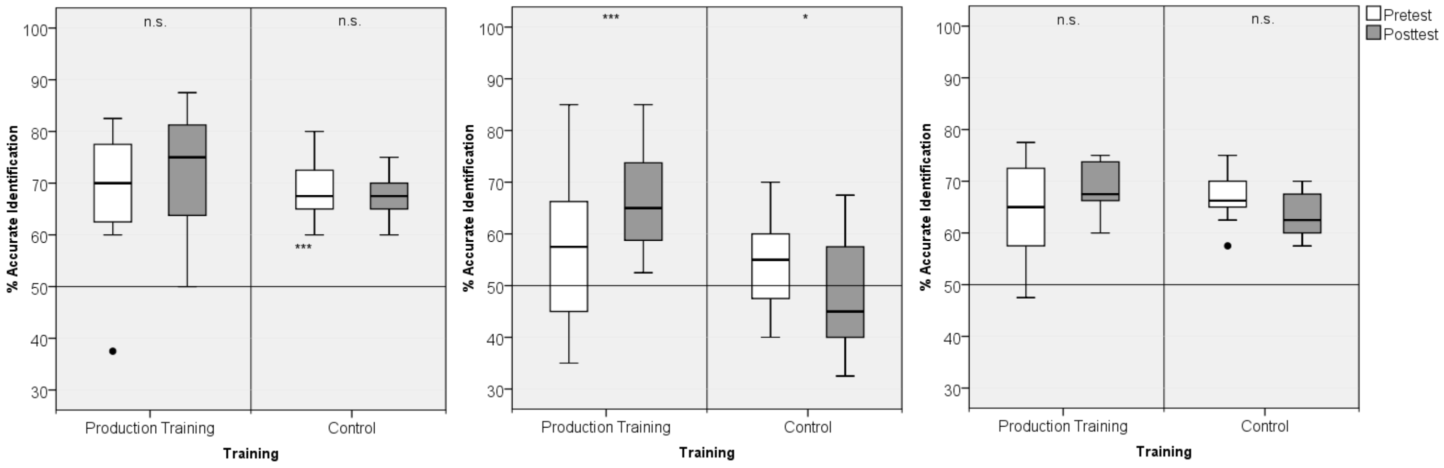


Figure 4.5 Mean percentages of accurate identification of /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right) by the group trained under the production training and the control group between the pretest (white boxes) and the posttest (dark boxes). For both /ɪ/-/iː/ and /ʊ/-/uː/ vowel pairs, there were non-significant differences between the pretest and the posttest (n.s. = $p > .05$); whereas for /e/-/æ/, there was significant differences between the pretest and posttest performance (*** = $p < .001$; * = $p < .05$). The black dots represent outliners. The horizontal line indicates the chance level (50%) performance.

three-way repeated-measures ANOVA with Production Training (with or without) as a between-subject factor and Test (pretest and posttest) and Vowel (front, low, back) as repeated measures. Interestingly, only the main effect of Vowel [$F(2,38) = 15.73$, $p < .001$] was robust because the low vowel pair was consistently more accurately identified than the front and back vowel. The interaction Test × Production Training [$F(1,19) = 22.79$, $p < .001$] was also significant, because the group with production training improved while the control group's performance deteriorated from the pretest to the posttest, constituting a crossover interaction effect. However, the main effects of Test ($p = .270$), Production Training ($p = .142$) and other interactions were not robust, due to the fact that significant improvement could only be observed among the group trained under the production training paradigm and the low vowel pair.

The following figure illustrates the crossover interaction of the two factors based

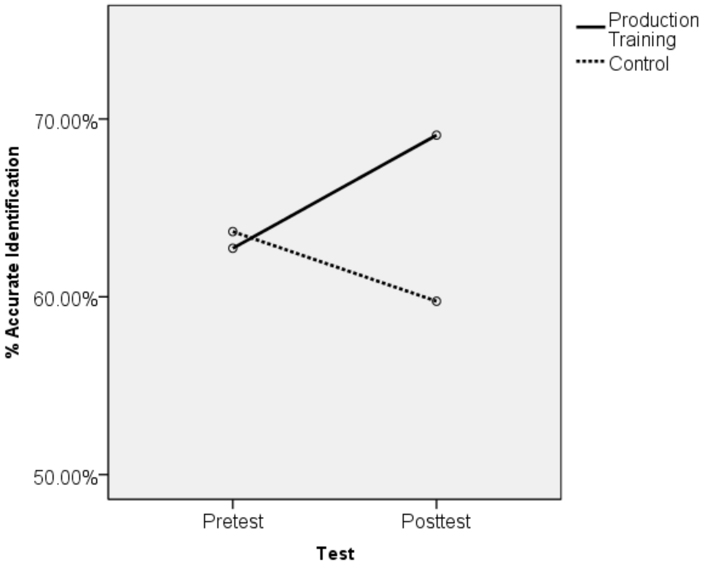


Figure 4.6 The interaction of Test × Production Training.

on the test of simple effects on the interaction Test × Production Training. It revealed that there were no significant differences between the performance of the group which received production training and the control group in the pretest (p = .790); while in the posttest, the group with production training performed significantly better than the control group for 9.34% in sum of all three vowel pairs [$F(1,19) = 16.70$, $p < .001$]. The group with production training also improved in the perception of the three vowel pairs for 6.36% [$F(1,10) = 10.90$, $p < .001$] from the pretest to posttest whereas the control group slipped back for 3.92% [$F(1,9) = 26.13$, $p < .001$] in the posttest when it is compared with the pretest results.

4.1.4.4 Factor 4: Training Diversity

Though from the above analyses the effects of the above training paradigms can be observed, the analyses per se do not provide any substantial and direct comparisons to ascertain which training paradigm outscores another or to what extent a combination of both perception and production training will benefit the subjects more than providing them with training in only one domain. In light of this, all the mean identification scores were submitted into a three-way repeated-measures ANOVA for analysis with Training Type (HVPT only [H], n = 17; HVPT with production [HP], n = 18; LVPT only [L], n = 15; LVPT with production [LP], n = 14; production only [P], n = 11) as a between-subject factor and Test (pretest and posttest) and Vowel (front, low, back) as repeated measures. Mean differences from the pretest to posttest are given in Table 4.4

TABLE 4.4

Mean Test Scores, Mean Percentage, and Mean Difference of all 5 Training Types across 3 Vowel Pairs in Perceptual Pretest and Posttest

		Perceptual Performance					
		Pretest	%	Posttest	%	Difference	%
	Total Score:	40		40			
/ɪ/-/iː/	H	27.88	69.71%	35.88	89.71%	8.00	20.00%
	S.D.:	2.50	6.24%	2.71	6.78%		
	HP	28.89	72.22%	35.72	89.31%	6.83	17.08%
	S.D.:	2.85	7.12%	2.67	6.69%		
	L	28.67	71.67%	35.27	88.17%	6.60	16.50%
	S.D.:	3.18	7.94%	2.81	7.04%		
	LP	27.79	69.46%	34.21	85.54%	6.43	16.07%
	S.D.:	3.04	7.61%	3.89	9.72%		
	P	27.27	68.18%	28.64	71.59%	1.36	3.41%
	S.D.:	5.14	12.85%	4.70	11.74%		
/e/-/æ/	H	24.71	61.76%	35.24	88.09%	10.53	26.32%
	S.D.:	3.90	9.75%	3.54	8.86%		
	HP	23.39	58.47%	37.11	92.78%	13.72	34.31%
	S.D.:	4.47	11.18%	2.11	5.28%		
	L	23.47	58.67%	36.27	90.67%	12.80	32.00%
	S.D.:	5.74	14.36%	3.22	8.04%		
	LP	24.79	61.96%	37.43	93.57%	12.64	31.61%
	S.D.:	2.67	6.66%	1.91	4.78%		
	P	22.45	56.14%	26.82	67.05%	4.36	10.91%
	S.D.:	6.07	15.18%	4.12	10.30%		
/ʊ/-/uː/	H	26.88	67.21%	33.76	84.41%	6.88	17.21%
	S.D.:	3.71	9.27%	3.40	8.50%		
	HP	27.44	68.61%	34.11	85.28%	6.67	16.67%
	S.D.:	4.18	10.44%	3.10	7.76%		
	L	26.40	66.00%	36.00	90.00%	9.60	24.00%
	S.D.:	4.37	10.93%	4.64	11.61%		
	LP	26.71	66.79%	35.71	89.29%	9.00	22.50%
	S.D.:	3.36	8.40%	4.08	10.21%		
	P	25.55	63.86%	27.45	68.64%	1.91	4.77%
	S.D.:	3.88	9.71%	2.21	5.52%		

In addition to the significant main effects of Test [$F(1,70) = 547.08$, $p < .001$], Vowel [$F(2,140) = 10.73$, $p < .001$] and Training Type [$F(4,70) = 11.50$, $p < .001$] which were due to the training types' overall improvement from the pretest to posttest across the three vowel pairs, the Test × Training Type [$F(4,70) = 12.82$, $p < .001$] and Test × Vowel [$F(2,140) = 26.53$, $p < .001$] interactions were also significant.

Planned comparisons with Bonferroni correction on the Test × Training Type interaction showed that all the groups improved from the pretest to posttest with significance. Figure 4.7 a, b and c on the next page illustrate the performance across training types and the three vowel pairs: The H type had an increase of 21.17% [$F(1,16) = 158.22$, $p < .001$]; the HP type had an increase of 22.90% [$F(1,17) = 210.53$, $p < .001$]; the L type had an increase of 24.17% [$F(1,14) = 128.72$, $p < .001$]; the LP type had an increase of 23.39% [$F(1,13) = 145.91$, $p < .001$] whereas the P type only had an increase of 6.36% [$F(1,10) = 10.90$, $p = .008$], indicating that all training types could help the subjects improve the perception of the vowel pairs to different extent. Also, all the training types had no significant differences in their performance in the pretest ($p = .609$); yet, in the posttest, significant differences were observed only between the four training types with the P type, i.e. H > P [$F(1,26) = 53.24$, $p < .001$], HP > P [$F(1,27) = 95.44$, $p < .001$], L > P [$F(1,24) = 67.77$, $p < .001$] and LP > P [$F(1,23) = 60.92$, $p < .001$], meaning that except the P type, all the four training types did not have significant differences between each other, i.e. no support can be given to claim which training type was be more effective than the other one in terms of the performance after training.

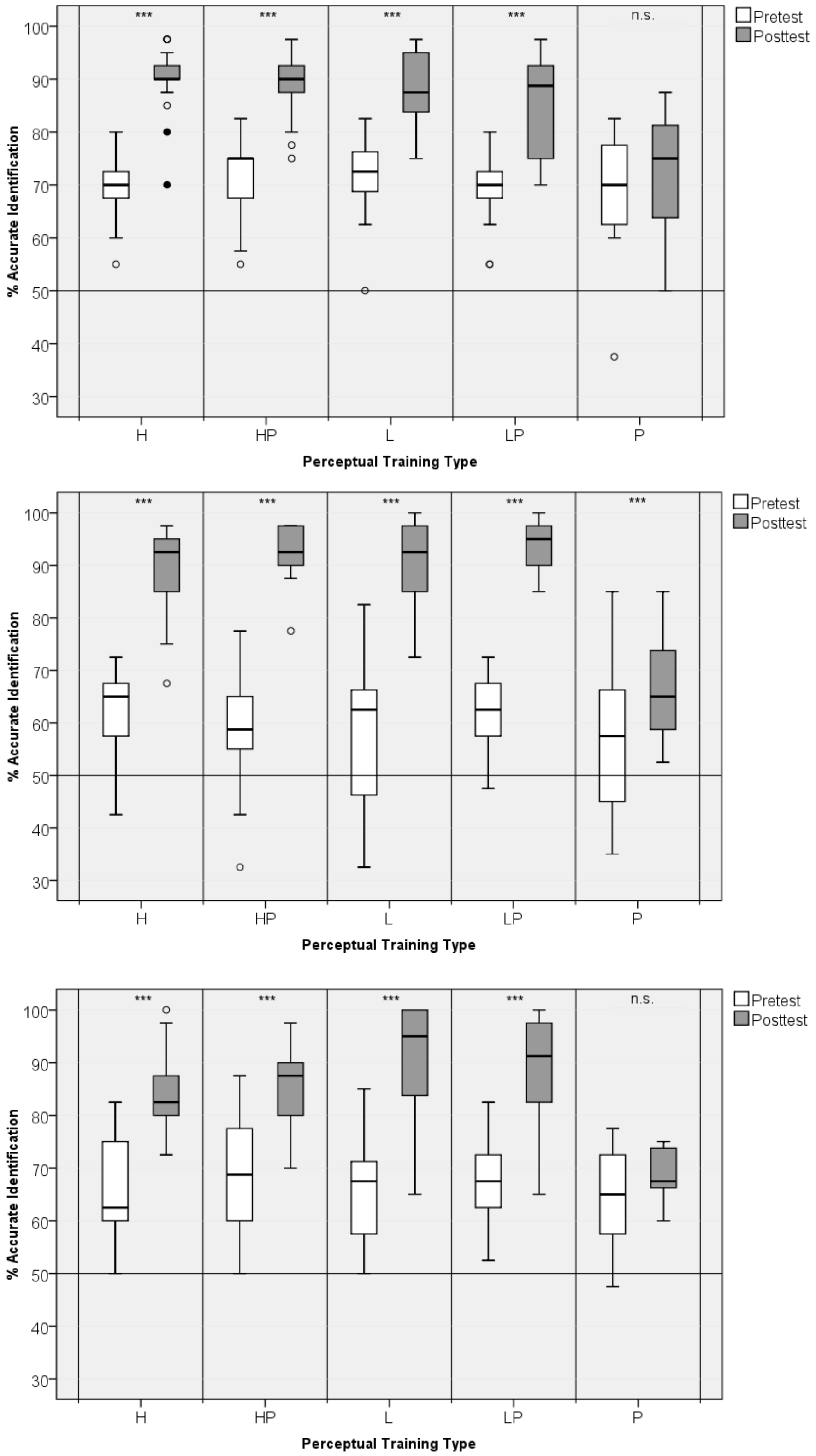


Figure 4.7a, b, c Mean percentages of correct identification of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from top to down) respectively among the 5 training types from the pretest (white boxes) to the posttest (dark boxes) with significant differences (*** = *p* < .001; ** = *p* < .005). The circles represent extreme outliers. The horizontal line indicates the chance level (50%) performance.

4.1.4.5 Summary of the Results of Perceptual Learning

Both HVPT and LVPT were shown to be beneficial to the subjects' learning of the three target vowel pairs to a similar extent as no significant differences could be observed between the two training approaches. It appears not to be in line with previous studies which have shown that HVPT is more efficient than LVPT as HVPT provides a wider variety of tokens with phonetic information crucial for the perpetual learning. The effects of the two training intensity levels were too similar that no significant differences were obtained. Production training was also useful for the perceptual learning of the low vowel pair only. Meanwhile, an addition of production training on perceptual training, be it HVPT or LVPT, does not appear to benefit the subjects more in the perpetual learning of the target vowels as the results across training types from the pretest to posttest are very close. Given the above results reporting the effects of various factors on the perceptual learning based on the pretest and posttest results, the generalization effects of the learning of the two groups is thus of our further interest as it provides a more complete picture of whether the perceptual learning can be transferred to new words which had not appeared in the training or tests before and/or produced by a new speaker whose voice had not been heard by the subjects. In the Section 4.1.5 below, we are going to look at the results in generalization.

4.1.5 Results in Generalization of Perceptual Learning to New Speaker/Word

Besides gauging the efficacy of the training paradigms through comparing the perceptual performance from the pretest to the posttest, the generalization of learning (perception of new words/speakers) is certainly one important indicator that measures the external validity of the training. The validation is important to, particularly, a language program or curriculum designer, since the generalization of learning can give researchers more support to confirm whether a training paradigm is useful. To access generalization, I compared the difference between the scores in the pretest and different Tests of Generalization[13] upon various factors – perceptual training, perceptual training intensity, production training, training diversities, which will be shown in the following sections.

4.1.5.1 FACTOR 1: PERCEPTUAL TRAINING ON GENERALIZATION

Figure 4.8 on the next page shows the differences of the mean accurate identification scores between the pretest with the TGs. A three-way repeated-measures ANOVA with Perception Training (HVPT, n = 17; LVPT, n = 15; Control, n = 10) as between-subjects factor and TG (three mean difference scores between TG1/TG2/TG3 with the pretest) and Vowel (front, low, back) as repeated measures revealed a significant main effect of TG [$F(2,78) = 7.81$, $p < .001$], due to the differences in the extent of improvement across the three TGs: the subjects had higher scores in TG2 than in TG1 for 3.41% ($p = .019$) and higher scores in TG3 than TG1 for 4.41% ($p = .002$). A main effect of Vowel [$F(2,78) = 4.78$, $p = .011$] was observed as well, because the low

---

[13] To recapitulate the details of the three Tests of Generalization: TG1 represents the test with new tokens produced by a new speaker; TG2 represents the test with new tokens produced by a familiar speaker whereas TG3 represents the test with familiar tokens produced by a new speaker.

vowel pair was consistently more accurately identified than the front vowel pair for 7.46% ($p < .001$) . There was also a main effect of Perception Training [$F(2,39) = 30.27$, $p < .001$], since the HVPT group improved more than the LVPT group for 7.29% ($p = .005$) and the control group for 19.05% ($p < .001$) and the LVPT performed better than the control group for 11.76% ($p < .001$). Yet, there were no significant two-way interactions between TG and Perception Training ($p = .073$), Vowel and Perception Training ($p = .628$) or TG and Vowel ($p = .317$), and neither was there a three-way interaction between TG, Perception and Vowel ($p = .204$).

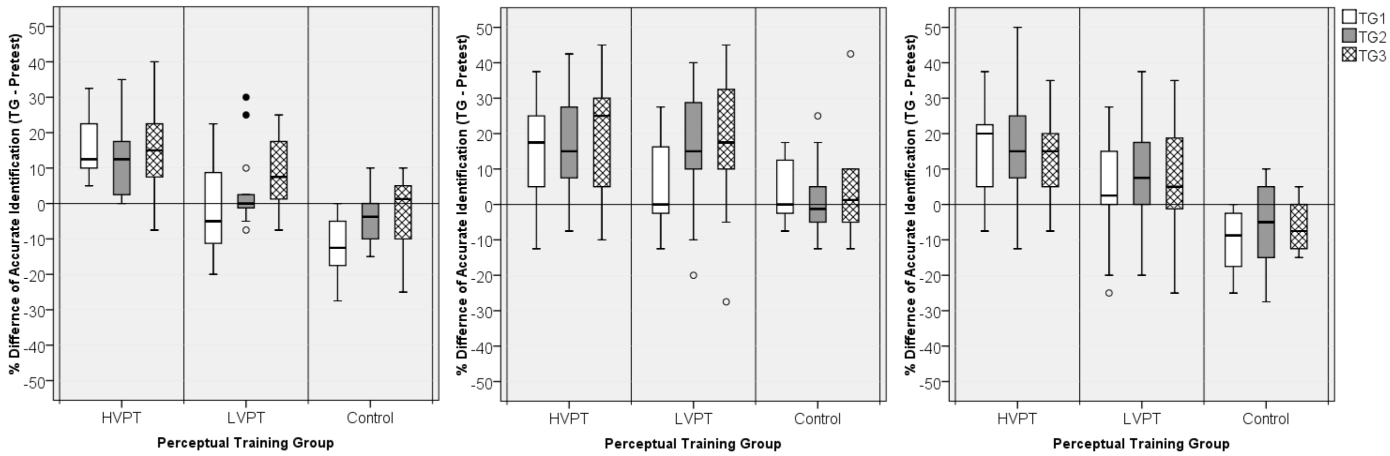


Figure 4.8 Difference (TG – Pretest) between the mean percentages of correct identification of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ respectively (from left to right) across the three perceptual training conditions across the three TGs. TG1 is represented by the white boxes, the dark boxes represent TG2 while the patterned boxes represent TG3. The circles represent extreme outliers. The horizontal line indicates the time when no difference between the TG and Pretest (0% of difference) is observed.

Note that the significant main effect of TG has included the scores of the control group. It will be fairer to compare only the difference scores in TGs between HVPT and LVPT, before drawing conclusion on how much generalization occurred. A three-way repeated-measures ANOVA with Perception Training (HVPT, n = 17; LVPT, n = 15) as between-subjects factor and TG (three mean difference scores between TG1/TG2/TG3 with the pretest) and Vowel (front, low, back) as repeated measures revealed a significant main effect of TG [$F(2,60) = 5.34$, $p = .007$] since TG3 was significantly higher than TG1 for 4.07% ($p = .007$). The interactions TG × Perception Training [$F(2,60) = 4.35$, $p = .017$] and TG × Vowel [$F(4,120) = 2.76$, $p = .031$] were also robust. Planned comparisons with Bonferroni correction on TG × Perception Training interaction revealed that no significant differences between TGs in the HVPT group; but TG2 difference scores were great than TG1 for 6.39% and TG3 greater than TG1 for 7.56%, both with significance, within the LVPT group. This interaction means that the HVPT subjects' identification scores when listening to new words and new speakers were as accurate as with familiar words and familiar speakers; yet, the same did not happen to the LVPT group.

Contrasting the results obtained in posttest in which HVPT and LVPT showed no significant differences in the performance, the generalization results indicated that on top of the observation that both HVPT and LVPT approaches were effective in modifying the perception of the three vowel pairs, HVPT was also more effective than LVPT under different speaker and word conditions. It is worth paying attention to the result that the generalization effects were extended only to the HVPT group, but not to the LVPT group.

4.1.5.2 Factor 2: Perceptual Training Intensity on Generalization

Figure 4.9 a and b illustrate the generalization performance (by using the scores in

TG – scores in pretest) of the two perpetually trained groups (HVPT and LVPT) with different intensity levels across the three vowel pairs. The analysis of the perceptual performance from the pretest to posttest already showed that intensity did not play a significant role in determining the performance of the subjects. Two separate three-way ANOVAs were submitted to check whether intensity levels have any effects on the generalization of learning.

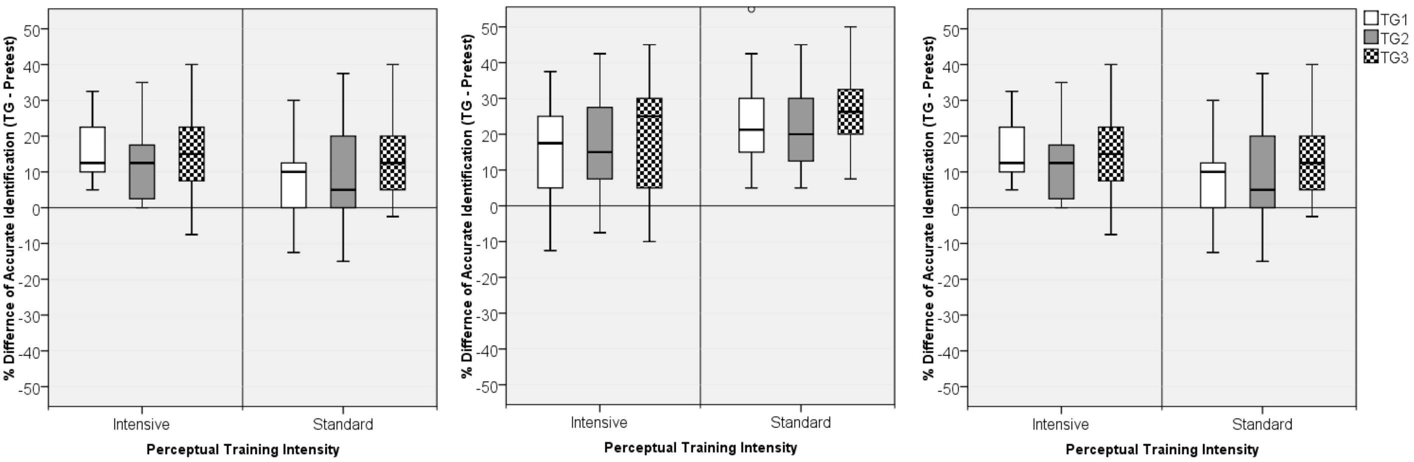


Figure 4.9a Difference (TG – Pretest) between the mean percentages of correct identification of the HVPT groups with different intensity levels across the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right). The circles represent outliners.

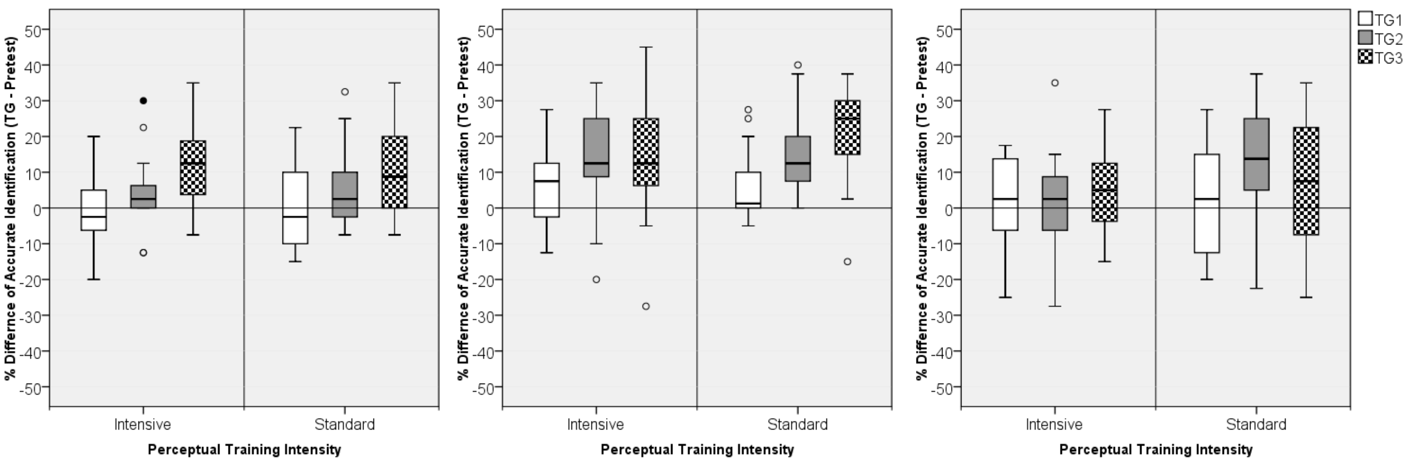


Figure 4.9b Difference (TG – Pretest) between the mean percentages of correct identification of the LVPT groups with different intensity levels across the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right). The circles represent outliners.

The first three-way repeated-measures ANOVA for the HVPT group with Perception Training Intensity (Intensive, n = 9; Standard, n = 8) as between-subjects factor and TG (three mean difference scores between TG1/TG2/TG3 with the pretest) and Vowel (front, low, back) as repeated measures showed that no main effects were robust at all: Perceptual Training Intensity (p = .077), Vowel (p = .723), TG (p = .525), indicating that the generalization of learning in different speaker and word conditions, or across training intensity levels or across the three target vowel pairs did not significantly differ from each other. The interactions TG × Vowel (p = .133), TG × Perceptual Training Intensity (p = .357), Vowel × Perceptual Training Intensity (p = .365) and the TG × Vowel × Perceptual Training Intensity (p = .382) were not significant either. It is indicated that generalization to new words or speakers was observed among the HVPT groups, although not affected by the intensity levels.

For the LVPT groups, another three-way repeated-measures ANOVA with Perception Training Intensity (Intensive, n = 8; Standard, n = 7) as between-subjects factor and TG (three mean difference scores between TG1/TG2/TG3 with the pretest) and Vowel (front, low, back) as repeated measures was submitted for analysis. The ANOVA showed significant main effects of TG [$F(2,26) = 7.52$, p = .003], since the subjects had higher difference scores in TG3 than in TG1 for 7.52%, $t(28) = 5.97$, p = .011, and TG2 than in TG1 for 6.54%, $t(28) = 4.71$, p = .017. Yet, Vowel (p = .109) and Perceptual Training Intensity (p = .667) was not significant main effects, and neither were the interactions TG × Vowel (p = .186), TG × Perceptual Training Intensity (p = .369), Vowel × Perceptual Training Intensity (p = .786) and TG × Vowel ×

Perceptual Training Intensity (*p* = .112). Consistent with the results when Factor 1 was analyzed, the failure in generalization among the LVPT group was observed.

4.1.5.3 FACTOR 3: EXPLICIT PRODUCTION TRAINING ON GENERALIZATION

Figure 4.10 displays the performance of the group with production training and the control group in all three TGs:

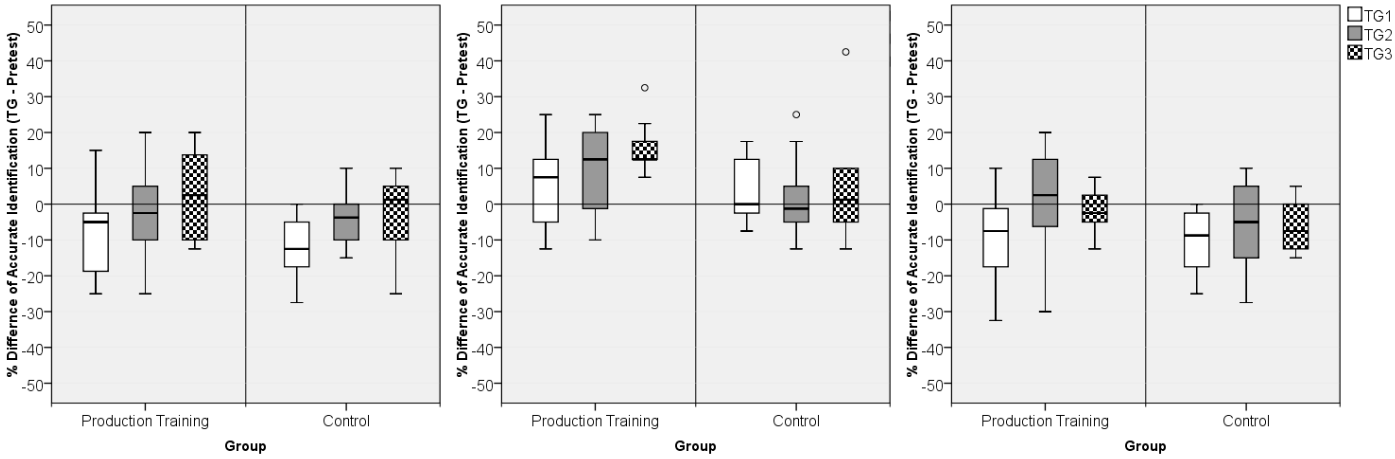


Figure 4.10 Difference (TG – Pretest) between the mean percentages of correct identification of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ respectively (from left to right) of the group with production training and the control group across the three TGs. TG1 is represented by the white boxes, the dark boxes represent TG2 while the patterned boxes represent TG3. The circles represent extreme outliers. The horizontal line indicates the time when no difference between the TG and Pretest (i.e. 0% of difference) is observed.

To investigate the effect of production training on the generalization effects of perceptual learning, a three-way repeated-measures ANOVA with Production Training (With, n = 11; Without, n = 10) as between-subjects factor and TG (three mean difference scores between TG1/TG2/TG3 with the pretest) and Vowel (front, low, back) as repeated measures was submitted to answer the question. The main effects of Production Training [$F(1,19) = 5.71$, $p = .027$] was significant, as the group with production training had higher difference scores than the control group across the target vowel pairs and the TGs for 4.90%. There was also a significant main effect of TG [$F(2,38) = 9.02$, $p = .001$], as the improvement in TG3 was more than TG1 for 7.06%, $t(20) = 3.75$, $p = .001$, and TG2 more than TG1 for 4.80%, $t(20) = 2.85$, $p = .010$. The significant main effect of Vowel [$F(2,38) = 20.44$, $p < .001$] was due to the fact that the subjects had higher difference scores for the low vowel pair than the front pair for 11.07%, $t(20) = 3.77$, $p = .001$, and the back vowel pair for 12.42%, $t(20) = 4.95$, $p < .001$. Yet, no interaction was robust: TG × Production Training ($p = .536$), Vowel × Production Training ($p = .798$), TG × Vowel ($p = .324$) and TG × Vowel × Production Training ($p = .487$). Since differences between the TGs were discovered, t-tests were administered on TG difference scores of only the production training group. The analysis revealed that the difference between TG2 and TG1 was not significant ($p = .062$); but the difference scores in TG3 was significantly higher than that in TG1 for 8.86%, $t(10) = 3.42$, $p = .007$. The results gave preliminary indication that production training could only improve the perceptual accuracy of words in new speaker condition but not in new word condition.

4.1.5.4 FACTOR 4: TRAINING DIVERSITY ON GENERALIZATION

In a previous section, the statistical analysis on the effects of training diversity on the performance after training, i.e. pretest vs. posttest, exhibited that perceptual training – regardless of its type – benefited the subjects' perceptual learning; yet, production training had only very mild effects on their learning. This section further examines the factor of training diversity on the generalization of perceptual learning. Figures 4.11 a, b and c on the next page cover the results of TGs across different training types and vowel pairs.

A three-way repeated-measures ANOVA with Training Type (HVPT only [H], n = 17; HVPT with production [HP], n = 18; LVPT only [L], n = 15; LVPT with production [LP], n = 14; production only [P], n = 11) as a between-subject factor and TG (three mean difference scores between TG1/TG2/TG3 with the pretest) and Vowel (front, low, back) as repeated measures was computed to analyze the results. The ANOVA revealed a significant main effect of Training Type [$F(4,70) = 9.55$, $p < .001$], in which the H type performed significantly better than the L type for 7.29% [$F(1,30) = 9.69$, $p = .004$], the LP type for 8.34% [$F(1,29) = 10.91$, $p = .003$] and the P type for 14.15% [$F(1,26) = 40.89$, $p < .001$]; while the HP type performed significantly better than the L type for 5.62% [$F(1,31) = 5.16$, $p = .030$], the LP type for 6.66% [$F(1,30) = 6.33$, $p = .017$] and the P type for 12.49% [$F(1,27) = 26.39$, $p < .001$] across the vowel pairs and TGs. There was also a significant main effect of TG [$F(2,140) = 21.79$, $p < .001$], due to the higher difference scores obtained in TG2 than in TG1, $t(74) = 3.23$, $p = .002$, and in TG2 than in TG3, $t(74) = 2.46$, $p = .016$. Vowel [$F(2,140) = 16.98$, $p < .001$] was also a robust main effect, since the subjects showed more improvement throughout the TGs in the low vowel pair than the back vowel pair for 10.33%, $t(74) = 6.26$, $p < .001$, and the front vowel pair for 9.97%, $t(74) = 5.51$, $p < .001$. There were also significant TG × Training Type [$F(8,140) = 3.42$, $p = .001$] and TG × Vowel [$F(4,280) = 6.03$, $p < .001$] interactions.

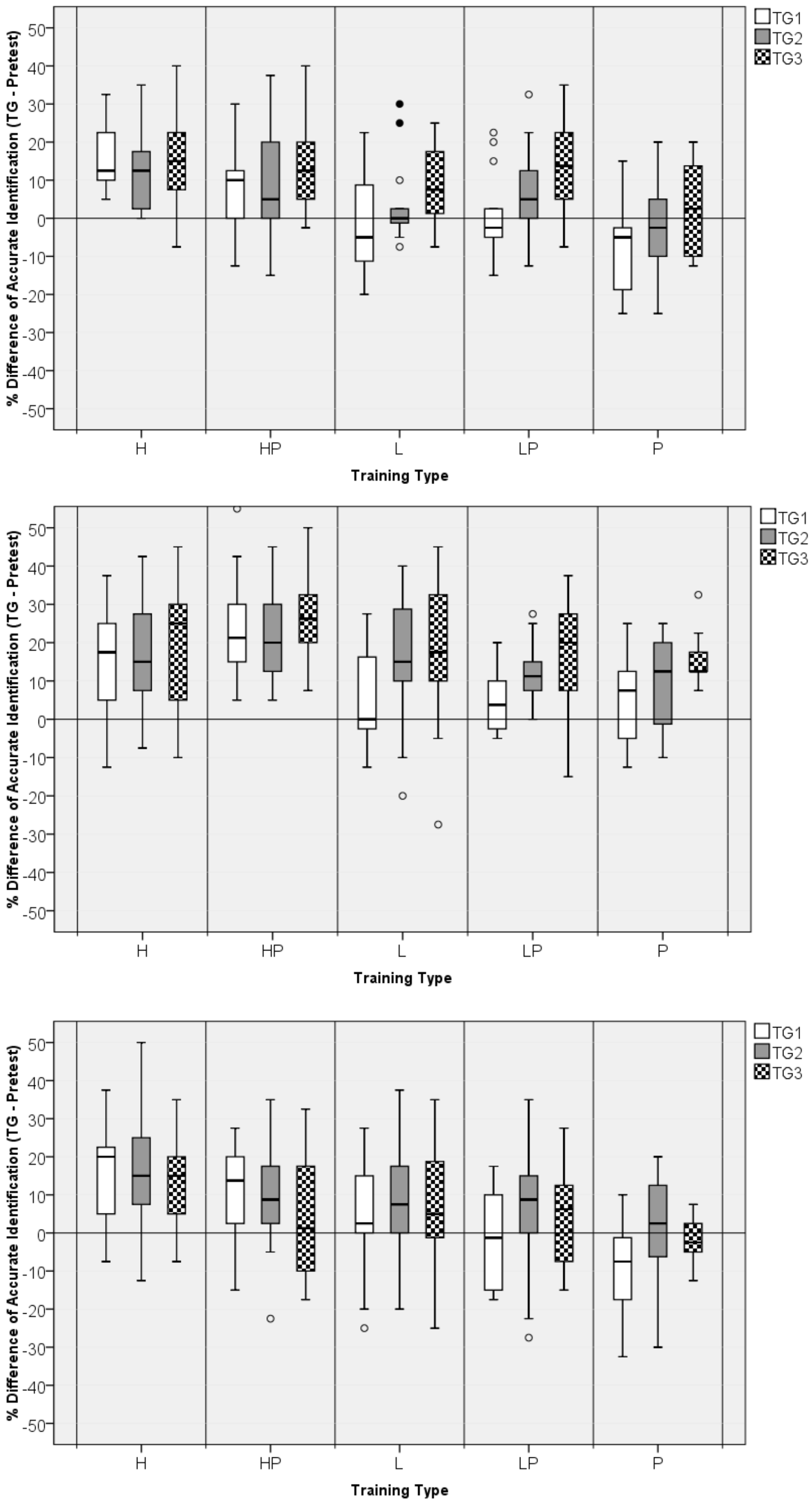


Figure 4.11 a, b, c Mean difference percentages (TG – Pretest) of correct identification of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from top to down) respectively among the 5 training types. The front vowel pair is represented by the white boxes, the dark boxes represent the low vowel pair while the patterned boxes represent the back vowel pair. The circles represent extreme outliers. The horizontal line indicates the time when no difference between the TG and Pretest (0% of difference) is observed.

Planned comparisons with Bonferroni correction were carried out to explore only the interaction between TG and Training Type. For both the H and the HP types, no significant differences were discovered among the three TGs, indicating that the degree of learning extent in the three TGs was highly comparable and that learning can be generalized to both new word and new speaker condition. Yet, the L, LP and P types all had different performances in the three TGs: for the L type, TG3 > TG1 for 7.56%, $t(14) = 3.69$, $p = .002$, TG2 > TG1 for 6.39%, $t(14) = 3.21$, $p = .006$; for the LP type, TG3 > TG1 for 10.83%, $t(13) = 4.76$, $p < .001$; TG2 > TG1 for 6.91%, $t(13) = 3.34$, $p = .005$; and for the P type, TG3 > TG1 for 8.86%, $t(10) = 3.42$, $p = .007$. All these analyses revealed that learning in the perceptual modality could not be generalized to new word or speaker condition.

Concerning the differences between training types in different TGs, it is found that in TG1, the H and HP types performed better than all other three groups: H > L for 11.75% [$F(1,30) = 24.71$, $p < .001$], H > LP for 14.06% [$F(1,29) = 33.60$, $p < .001$] and H > P for 18.81% [$F(1,26) = 48.07$, $p < .001$] whereas HP > L for 10.20% [$F(1,31) = 16.20$, $p < .001$], HP > LP for 12.51% [$F(1,30) = 23.06$, $p < .001$] and HP > P for 17.26% [$F(1,27) = 34.94$, $p < .001$]. In TG2 and TG3, the picture was simpler: the H and HP types only had more learning than the P type. For TG2, H > P for 13.12% [$F(1,26) = 18.95$, $p < .001$] and HP > P for 10.70% [$F(1,27) = 13.36$, $p = .001$]; for TG3, H > P for 10.53% [$F(1,26) = 18.43$, $p < .001$] and HP > P for 9.47% [$F(1,27) = 10.10$, $p = .004$]. These results indicated that all the training types had different performance under new speaker and word conditions, the H and HP types appeared to be more effective training types than the others in terms of the generalization performance gauged by comparing the TG scores and the pretest scores. Moreover, the subjects in general did better in TG2 and TG3 than in TG1, showing that new tokens produced by a new speaker were generally a harder task for the subjects. These will be further discussed in the next chapter.

Planned comparisons with Bonferroni correction on TG × Vowel interaction revealed that across all TGs, the low vowel pair was consistently more correctly identified than the other two pairs. In TG1, the difference score between TG1 and the pretest for the low vowel pair was 7.86% more than the front pair, $t(74) = 3.79$, $p < .001$; and 6.97% more than the back pair, $t(74) = 3.18$, $p = .002$. In TG2, the performance of the low vowel pair was 10.00% more than the front pair, $t(74) = 4.82$, $p < .001$; while 7.48% more than the back pair, , $t(74) = 3.25$, $p = .002$. In TG3, the low vowel pair was 8.86% more correctly identified than the front pair, $t(74) = 4.22$, $p < .001$ and 13.93% more than the back pair, , $t(74) = 46.42$, $p < .001$.

4.1.5.5 Summary of the Generalization Effect

The above statistical analysis on the generalization of learning delineates a rather different picture with the results obtained in Section 4.1.4 which concerns about the results of perceptual learning by simply comparing the pretest and posttest results. In contrast to the results in perceptual learning showing that both HVPT and LVPT were useful training paradigm which had no significant differences in their efficacy, it was found that the HVPT group had more significant improvement in the generalization tests than the LVPT and control groups. Meanwhile, the HVPT (intensive) group also showed more robust generalization than the HVPT (standard) group, whereas no significant differences were found between the LVPT groups with different intensity levels. Subjects with production training could generalize the perceptual learning in general as well. However, when different training types were compared, only the H and HP showed significantly better performance than other groups.

In sum, perceptual training benefits the subjects even when the perceptual tokens were new and/or were produced by a new voice, indicating that the learning could be generalized; yet, the training with high stimulus variability which had consistently greater performance was more effective than other paradigms.

4.1.6 Summing up the Results in Perceptual Identification Tests

The above statistical analyses shed some light on the interpretation of the effectiveness and the generalizability of the training approaches on improving the perceptual identification of the three target vowel pairs. In general, subjects with training, be it HVPT, LVPT, production training or a mixture of both, did considerably better than those without. The HVPT group also did better than the LVPT group, particularly in the generalization tests. The production training was slightly useful in improving the perceptual categorization of the vowel contrasts. Perceptual training intensity, as one main focus of the present study, was not a highly influential factor in improving the perceptual performance of the subjects. All these factors will be discussed in detail in the next chapter. The results of the production tests were also worth examining before interpreting the results in the perceptual aspect, for it provides a more comprehensive picture of the efficacy of the training approaches. If implicit transfer of perceptual learning to production can be traced, implications to links between the perception and production of sound categories emerged can be deduced. The next section elucidates the results and statistical analyses of the production performance of all the groups, which includes the posttest and Test of Contextualization. Acoustic analyses will also be presented to support and validate the production results.

4.2 Production Performance

Whether there is any transfer of perceptual learning to production or to what extent explicit production training and a combination of both perceptual and production training help the production of the target vowel pairs are of core concern in the present study. Evaluated by two ways[14], the production results of all the subjects will be presented in detail in this Section.

Firstly, the production scores were obtained via three phonetic transcription exercises conducted by three different phonetically-trained researchers (including native speakers). The data will be the basis for major comparison and statistical analysis for it represents the perceptible change in the intelligibility of the productions indicated based on the transcriptions. Secondly, all the productions were submitted to acoustic measurement of which the F1, F2 and duration values were attained for comparisons. The F1 and F2 values, being the meaningful frequency components important for distinguishing vowels, together allows more quantitative measurement on the vowel produced by the subjects, so that the change (if any) in the production after intervention can be gauged. The change in the production of vowel duration can also provide evidence to how perceptual training will influence the subjects' production of duration difference between the contrasts.

A presentation of a preview and an overview of the production results will be shown in 4.2.1 and 4.2.2 respectively. Following this will be the choice of statistical analysis in 4.2.3, and the results of perceptual training in production in which the effects

---

[14] The evaluation procedures of the production data have been elucidated in detail in Section 3.4.2.

of four factors in investigation in 4.2.4. Section 4.2.5 states the effects of the four factors on the results of the Test of Contextualization, giving more details on the transfer of perceptual learning on production at the sentence level. Before closing this section, a detailed acoustic analysis of productions will be presented in 4.2.6, giving more solid insights to the actual production performance of the subjects.

4.2.1 Preview of the Overall Production Performance

As a preview of the entire section, Table 4.5 on the next page serves as an overview of the results of the statistical analyses conducted that compared the scores obtained in the production pretest, posttest and the Test of Contextualization. The statistical analysis of the four factors, namely perceptual training, perpetual training intensity, production training and training diversity, that were involved in the experiment and the specific results that were important to answering the research questions will be displayed in the table as a prelude of understanding the data.

4.2.2 Overall Production Performance

Tables 4.6 a, b and c on the next several pages (for a detailed version with individual subject's scores and percentages, see Appendix K) list the overall results of the production performance of all the ten experimental and control groups by showing the averaged scores and mean percentages of accurate production in the pretest, the posttest and the Test of Contextualization (TC) across the three target vowel pairs. This allows a first glance of the performance before and after training.

TABLE 4.5

Overview of the Statistical Results of the Effects of Different Factors on the Transfer of Perceptual Learning to the Production Domain

Factors in investigation	**Results (only important ones related to RQs will be shown)**	**Statistically significant?**
A. Training effectiveness (Pretest vs. Posttest)		
1. Perceptual Training (HVPT vs. LVPT vs. Control)	a. HVPT > LVPT b. HVPT > Control c. LVPT > Control	✓ ✓ ✓
2. Perceptual Training Intensity (Standard vs. Intensive)	c. HVPT: Standard ~ Intensive d. LVPT: Standard ~ Intensive	✗ ✗
3. Production Training (With vs. Without)	With > Without	✓
4. Training Diversity (Perceptual Training vs. Perceptual Training + Production Training vs. Production Training)	a. HVPT + Production > HVPT only b. HVPT + Production > LVPT only c. HVPT + Production > LVPT + Production d. HVPT + Production > Production only	✓ ✓ ✓ ✓
B. Contextualization		
1. Perceptual Training (HVPT vs. LVPT vs. Control)	a. HVPT > LVPT b. HVPT > Control c. LVPT > Control	✓ ✓ ✗
2. Perceptual Training Intensity (Standard vs. Intensive)	a. HVPT: Standard ~ Intensive b. LVPT: Standard ~ Intensive	✗ ✗
3. Production Training (With vs. Without)	With > Without	✓
4. Training Diversity (Perceptual Training vs. Perceptual Training + Production Training vs. Production Training)	a. HVPT + Production > HVPT only b. HVPT + Production > LVPT only c. HVPT + Production > LVPT + Production d. HVPT + Production > Production only	✓ ✓ ✓ ✓

✓ = Yes

✗ = No

TABLE 4.6a

Mean Scores and Mean Percentage of the Accurate Production of the Vowel Pair /ɪ/-/iː/ *in 3 Production Tests (Pretest, Posttest, Test of Contextualization) across different Groups*

	Group Name	Perceptual Training	Training Intensity	Production Training	Total Score:	Pretest 20	%	Posttest 20	%	TC 30	%
Trained	HSP	HVPT	Standard	Yes	*Average:*	9.44	47.22%	17.22	86.11%	24.67	82.22%
					S.D.:	0.76	3.78%	2.17	10.84%	6.00	20.00%
	HIP	HVPT	Intensive	Yes	*Average:*	9.89	49.44%	18.22	91.11%	24.78	82.59%
					S.D.:	1.55	7.76%	1.98	9.91%	2.72	9.07%
	HSN	HVPT	Standard	No	*Average:*	10.00	50.00%	15.56	77.78%	22.22	74.07%
					S.D.:	2.36	11.78%	3.00	14.99%	4.07	13.56%
	HIN	HVPT	Intensive	No	*Average:*	10.25	51.25%	15.88	79.38%	23.75	79.17%
					S.D.:	0.46	2.31%	3.68	18.41%	4.43	14.77%
	LSP	LVPT	Standard	Yes	*Average:*	9.71	48.57%	15.00	75.00%	21.14	70.48%
					S.D.:	2.05	10.27%	3.46	17.31%	5.12	17.06%
	LIP	LVPT	Intensive	Yes	*Average:*	9.57	47.86%	14.71	73.57%	18.86	62.86%
					S.D.:	1.75	8.75%	3.29	16.46%	4.72	15.73%
	LSN	LVPT	Standard	No	*Average:*	9.86	49.29%	13.29	66.43%	17.00	56.67%
					S.D.:	2.18	10.91%	3.51	17.54%	5.41	18.04%
	LIN	LVPT	Intensive	No	*Average:*	9.50	47.50%	12.38	61.88%	18.50	61.67%
					S.D.:	2.27	11.34%	2.77	13.87%	4.75	15.84%
Control	COP	Control	N/A	Yes	*Average:*	10.00	50.00%	13.64	68.18%	20.36	67.88%
					S.D.:	1.04	5.18%	4.03	20.17%	5.71	19.02%
	CON	Control	N/A	No	*Average:*	9.40	47.00%	9.30	46.50%	14.75	49.17%
					S.D.:	0.99	4.96%	1.73	8.63%	8.26	27.54%

TABLE 4.6b

Mean Scores and Mean Percentage of the Accurate Production of the Vowel Pair /e/-/æ/ *in 3 Production Tests (Pretest, Posttest, Test of Contextualization) across different Groups*

						Production of /e/-/æ/					
	Group Name	Perceptual Training	Training Intensity	Production Training	Total Score:	Pretest 20	%	Posttest 20	%	TC 30	%
Trained	HSP	HVPT	Standard	Yes	*Average:*	10.33	51.67%	17.11	85.56%	23.78	79.26%
					S.D.:	0.52	2.59%	4.06	20.31%	5.06	16.88%
	HIP	HVPT	Intensive	Yes	*Average:*	10.00	50.00%	17.67	88.33%	24.22	80.74%
					S.D.:	2.80	14.00%	3.34	16.68%	6.16	20.55%
	HSN	HVPT	Standard	No	*Average:*	10.11	50.56%	15.11	75.56%	18.72	62.41%
					S.D.:	4.45	22.27%	3.15	15.75%	4.88	16.27%
	HIN	HVPT	Intensive	No	*Average:*	9.50	47.50%	15.50	77.50%	20.88	69.58%
					S.D.:	1.85	9.26%	4.50	22.52%	3.94	13.15%
	LSP	LVPT	Standard	Yes	*Average:*	9.43	47.14%	14.14	70.71%	19.57	65.24%
					S.D.:	3.01	15.03%	3.79	18.94%	6.14	20.46%
	LIP	LVPT	Intensive	Yes	*Average:*	9.86	49.29%	13.14	65.71%	15.71	52.38%
					S.D.:	2.76	13.81%	3.64	18.21%	5.63	18.78%
	LSN	LVPT	Standard	No	*Average:*	10.71	53.57%	13.29	66.43%	16.00	53.33%
					S.D.:	2.39	11.97%	3.73	18.65%	5.90	19.68%
	LIN	LVPT	Intensive	No	*Average:*	10.00	50.00%	12.88	64.38%	15.88	52.92%
					S.D.:	1.41	7.07%	3.40	17.00%	5.33	17.77%
Control	COP	Control	N/A	Yes	*Average:*	9.64	48.18%	13.55	67.73%	18.73	62.42%
					S.D.:	2.98	14.88%	4.41	22.04%	6.44	21.46%
	CON	Control	N/A	No	*Average:*	9.70	48.50%	9.30	46.50%	13.70	45.67%
					S.D.:	1.07	5.35%	2.70	13.48%	4.52	15.06%

TABLE 4.6c

Mean Scores and Mean Percentage of the Accurate Production of the Vowel Pair /ʊ/-/uː/ *in 3 Production Tests (Pretest, Posttest, Test of Contextualization) across different Groups*

						Production of /ʊ/-/uː/					
	Group Name	Perceptual Training	Training Intensity	Production Training	Total Score:	Pretest 10	%	Posttest 10	%	TC 20	%
Trained	HSP	HVPT	Standard	Yes	*Average:*	4.22	42.22%	7.56	75.56%	15.67	78.33%
					S.D.:	1.13	11.26%	1.31	13.09%	3.66	18.32%
	HIP	HVPT	Intensive	Yes	*Average:*	4.33	43.33%	7.22	72.22%	17.00	85.00%
					S.D.:	0.89	8.86%	0.99	9.91%	1.96	9.80%
	HSN	HVPT	Standard	No	*Average:*	4.56	45.56%	6.11	61.11%	12.89	64.44%
					S.D.:	0.76	7.56%	1.46	14.58%	2.38	11.88%
	HIN	HVPT	Intensive	No	*Average:*	4.38	43.75%	6.13	61.25%	14.00	70.00%
					S.D.:	0.52	5.18%	1.25	12.46%	3.42	17.11%
	LSP	LVPT	Standard	Yes	*Average:*	4.57	45.71%	6.14	61.43%	15.57	77.86%
					S.D.:	0.66	6.60%	1.25	12.46%	3.64	18.20%
	LIP	LVPT	Intensive	Yes	*Average:*	4.43	44.29%	5.86	58.57%	12.29	61.43%
					S.D.:	0.62	6.21%	1.22	12.15%	3.57	17.84%
	LSN	LVPT	Standard	No	*Average:*	4.57	45.71%	5.86	58.57%	11.29	56.43%
					S.D.:	0.60	6.03%	1.23	12.29%	3.70	18.48%
	LIN	LVPT	Intensive	No	*Average:*	4.53	45.33%	5.93	59.33%	10.33	51.67%
					S.D.:	0.64	6.40%	1.39	13.87%	3.75	18.77%
Control	COP	Control	N/A	Yes	*Average:*	4.45	44.55%	5.91	59.09%	12.00	60.00%
					S.D.:	0.92	9.16%	1.07	10.69%	2.36	11.78%
	CON	Control	N/A	No	*Average:*	4.60	46.00%	4.70	47.00%	9.40	47.00%
					S.D.:	0.53	5.35%	0.71	7.07%	1.69	8.45%

Figures 4.11 a, b and c visualize all the data. All the experimental groups have a considerable amount of improvement in general. For front vowel pair /ɪ/-/iː/, all four groups trained under HVPT had a high degree of improvement, ranging from 27.78% to 41.67%. The four LVPT groups also improved in the production, with three of them improved over 14% and one group around 27%; the group with production training had a gain of 18.18% while the control group had very similar performance from the pretest to the posttest.

Concerning the low vowel pair /e/-/æ/, the four HVPT groups had improved for more than 25%, with one group reaching even 38%. The four LVPT groups also had more accurate production after training; all had over 12% of improvement. The group with only production training also improved for 19.55% but the control group had not much change in their performance in terms of the averaged scores across subjects.

A similar picture can be observed in the back vowel pair /ʊ/-/uː/. The four groups trained under HVPT improved for over 15.56%, with one group reaching 33.33%; the four LVPT groups gained from 12.86% to 15.71% whereas the group with only production training had more accurate production for 14.55%. Again, the control group had only very slight change in the performance.

From the above brief overview, the four HVPT groups appeared to perform better than the LVPT groups and the group with only production training. Also, those perceptually trained groups with an addition of production training showed more improvement than the groups without. Yet, no regular pattern can be found for the performance of the groups with different intensity levels.

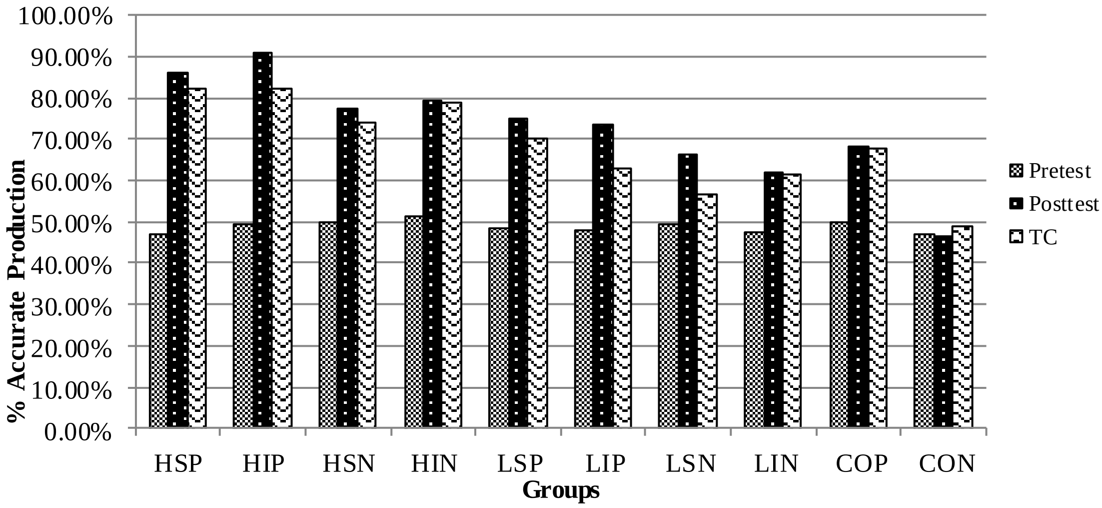


Figure 4.12a Mean percentages of accurate production performance of the vowel pair /ɪ/-/iː/ for the 9 training groups and control group in the pretest, the posttest and Test of Contextualization (TC)

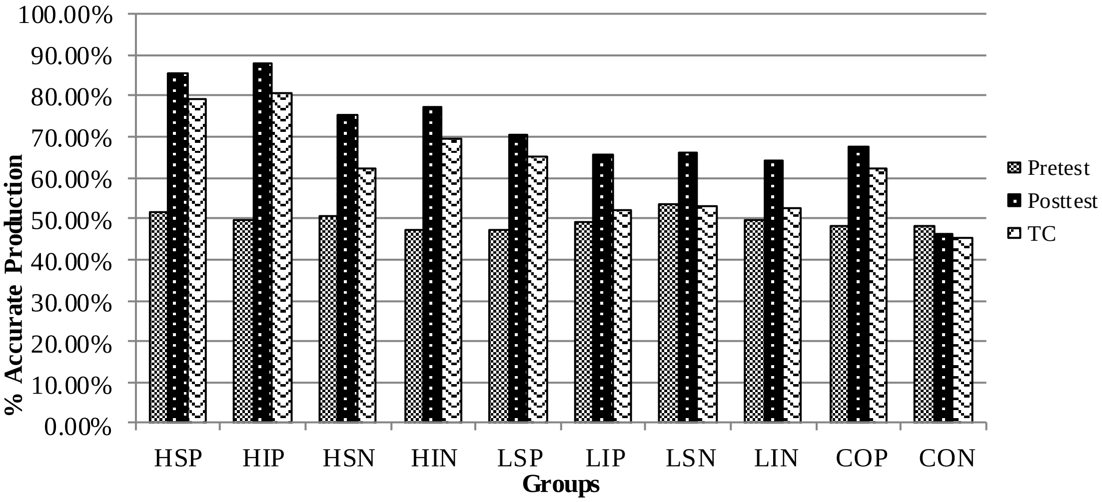


Figure 4.12b Mean percentages of accurate production performance of the vowel pair /e/-/æ/ for the 9 training groups and control group in the pretest, the posttest and Test of Contextualization (TC)

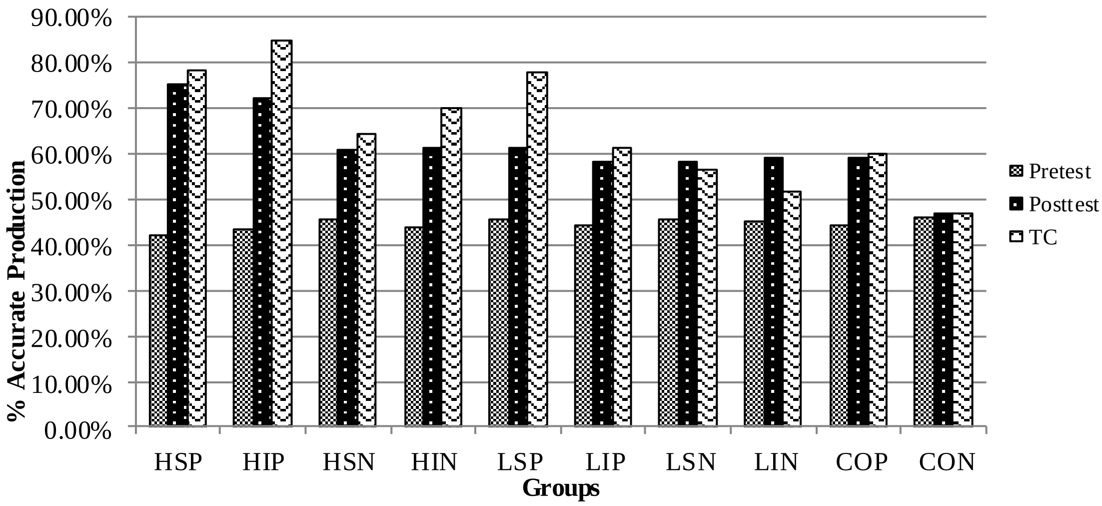


Figure 4.12c Mean percentages of accurate production performance of the vowel pair /ʊ/-/uː/ for the 9 training groups and control group in the pretest, the posttest and Test of Contextualization (TC)

Test of Contextualization (TC), which was not designed to be compared within-groups but between-groups, was an extra test aiming to offer further insights on whether the production performance can be maintained under a more naturalistic context after perceptual training. Presented as the mean percentages, the TC results of the ten groups showed that the production of the three vowel pairs in a context was over the chance level, 50% (some were up to 85%) for both the HVPT and LVPT groups. The group with solely production training also showed a very consistent pattern of accurate production of around 60% across the target vowel pairs. Nevertheless, almost all mean percentages in TC across all groups were lower than those they obtained in the posttest, hinting that the effects on production learning was context-dependent.

The confusion matrices in the next few pages summarize the production pretest and posttest performance of different groups. For the sake of easier understanding, the scores were collapsed into: 1) the HVPT groups (HSN, HIN), 2) the HVPT with production training groups (HSP, HIP), 3) the LVPT groups (LSN, LIN), 4) the LVPT with production training groups (LSP, LIP), 5) the production training group and 6) the control group. The productions were pooled across the target English productions in rows mapping to their actual performance in columns. The tables show the sum of responses and the overall percentages in the pretest and the posttest across the groups. The highlighted and bolded numbers are the productions made by the subjects which were the same as the targets. Unless the responses were the counterpart of a vowel pair contrast, those that are less than 2% were not shown.

TABLE 4.7a

Confusion Matrix showing the averaged scores and percentages of the responses in the Production Pretest and Posttest of the three vowel pairs by the subjects trained under HVPT with production training (N=18)

			Production (Pre-/Post-test) Responses						Sum
			/ɪ/	/iː/	/e/	/æ/	/ʊ/	/uː/	
Target English Vowels	/ɪ/	Pre	**1.00**	8.89					9.89
			10.00%	88.88%					98.88%
		Post	**8.67**	1.33					10.00
			86.67%	13.33%					100.00%
		Diff.	***7.67***	*-7.56*					
			76.67%	*-75.55%*					
	/iː/	Pre	1.00	**8.67**					9.67
			10.00%	**86.67%**					96.67%
		Post	0.83	**9.06**					9.89
			8.33%	**90.56%**					98.89%
		Diff.	*-0.17*	***0.39***					
			-1.67%	***3.89%***					
	/e/	Pre			**7.17**	2.67			9.83
					71.67%	26.67%			98.33%
		Post			**9.28**	0.72			10.00
					92.78%	7.22%			100.00%
		Diff.			***2.11***	*-1.94*			
					21.11%	*-19.44%*			
	/æ/	Pre			6.72	**3.00**			9.72
					67.22%	**30.00%**			97.22%
		Post			1.89	**8.11**			10.00
					18.89%	**81.11%**			100.00%
		Diff.			*-4.83*	***5.11***			
					-48.33%	***51.11%***			
	/ʊ/	Pre					**0.50**	4.11	4.61
							10.00%	87.69%	97.69%
		Post					**2.50**	1.83	4.33
							50.00%	50.00%	100.00%
		Diff.					***2.00***	*-2.27*	
							40.00%	*-45.47%*	
	/uː/	Pre					1.56	**3.78**	5.33
							23.33%	**75.56%**	98.89%
		Post					0.33	**4.89**	5.22
							2.22%	**97.78%**	100.00%
		Diff.					*-1.22*	***1.11***	
							-24.44%	***22.22%***	

TABLE 4.7b

Confusion Matrix showing the averaged scores and percentages of the responses in the Production Pretest and Posttest of the three vowel pairs by the subjects trained under HVPT (without production training) subjects (N=17)

Target English Vowels		Production (Pre-/Post-test) Responses						Sum
		/ɪ/	/iː/	/e/	/æ/	/ʊ/	/uː/	
/ɪ/	Pre	**2.20** **22.01%**	6.90 72.10%					9.10 94.12%
	Post	**5.77** **57.71%**	4.64 41.70%					10.41 99.41%
	Diff.	***3.57*** ***35.69%***	*-2.26* *-22.55%*					
/iː/	Pre	4.94 12.53%	**7.92** **79.24%**	0.76 7.65%				13.63 99.41%
	Post	2.03 0.56%	**9.94** **99.44%**	0 0.00%				11.97 100.00%
	Diff.	*-2.91* *-29.11%*	***2.02*** ***20.21%***	*-0.76* *-7.65%*				
/e/	Pre		0.83 8.29%	**7.47** **74.65%**	1.65 16.47%			9.94 99.41%
	Post		0.00 0.00%	**9.71** **97.08%**	0.14 1.41%			9.85 98.50%
	Diff.		*-0.83* *-8.29%*	***2.24*** ***22.43%***	*-1.51* *-15.06%*			
/æ/	Pre			7.48 74.83%	**2.34** **23.40%**			9.82 98.24%
	Post			4.40 44.03%	**5.60** **55.97%**			10.00 100.00%
	Diff.			*-3.08* *-30.80%*	***3.26*** ***32.57%***			
/ʊ/	Pre					**0.59** **11.81%**	4.23 88.19%	4.82 100.00%
	Post					**1.94** **38.89%**	2.24 61.11%	4.19 100.00%
	Diff.					***1.35*** ***27.08%***	*-1.99* *-39.72%*	
/uː/	Pre					2.43 21.32%	**3.88** **77.50%**	6.30 98.82%
	Post					0.99 16.53%	**4.17** **83.47%**	5.17 100.00%
	Diff.					*-1.43* *-28.68%*	***0.30*** ***5.97%***	

TABLE 4.7c.

Confusion Matrix showing the averaged scores and percentages of the responses in the Production Pretest and Posttest of the three vowel pairs by the subjects trained under LVPT with production training (N=14)

Target English Vowels			Production (Pre-/Post-test) Responses /ɪ/	/iː/	/e/	/æ/	/ʊ/	/uː/	Sum
	/ɪ/	Pre	**1.29** **12.86%**	8.07 83.57%					9.36 96.43%
		Post	**5.43** **54.29%**	4.50 45.00%					9.93 99.29%
		Diff.	***4.14*** ***41.43%***	*-2.79* *-27.86%*					
	/iː/	Pre	0.86 8.57%	**8.36** **83.57%**	0.79 7.86%				10.00 100.00%
		Post	0.14 1.43%	**9.43** **94.29%**	0.14 4.29%				9.71 100.00%
		Diff.	*-2.50* *-25.00%*	***1.07*** ***10.71%***					
	/e/	Pre		0.64 6.43%	**7.50** **75.00%**	1.86 18.57%			10.00 100.00%
		Post		0.86 8.57%	**7.50** **75.00%**	1.50 15.00%			9.86 98.57%
		Diff.		*0.21* *2.14%*	***0.00*** ***0.00%***	*-0.36* *-3.57%*			
	/æ/	Pre			7.71 77.14%	**2.14** **21.43%**			9.86 98.57%
		Post			3.79 37.86%	**6.14** **61.43%**			9.93 99.29%
		Diff.			*-3.93* *-39.29%*	***4.00*** ***40.00%***			
	/ʊ/	Pre					**0.21** **4.29%**	4.57 95.71%	4.79 100.00%
		Post					**1.50** **30.00%**	3.50 70.00%	5.00 100.00%
		Diff.					***1.29*** ***25.71%***	*-1.07* *-21.43%*	
	/uː/	Pre					0.71 14.29%	**4.29** **85.71%**	5.00 100.00%
		Post					0.43 8.57%	**4.50** **90.00%**	4.93 98.57%
		Diff.						***0.21*** ***4.29%***	

TABLE 4.7d.

Confusion Matrix showing the averaged scores and percentages of the responses in the Production Pretest and Posttest of the three vowel pairs by the subjects trained under LVPT (without production training) subjects (N=15)

			Production (Pre-/Post-test) Responses						Sum
			/ɪ/	/iː/	/e/	/æ/	/ʊ/	/uː/	
Target English Vowels	/ɪ/	Pre	**1.58**	8.21	0				9.79
			15.80%	83.53%	0.00%				99.33%
		Post	**4.08**	5.65	0.27				10.00
			40.80%	56.53%	2.67%				100.00%
		Diff.	***2.50***	*-2.56*	*0.27*				
			25.00%	*-25.57%*	*2.67%*				
	/iː/	Pre	1.24	**8.10**	0.67				10.00
			12.35%	**80.98%**	6.67%				100.00%
		Post	0.85	**8.75**	0.40				10.00
			8.50%	**87.50%**	4.00%				100.00%
		Diff.	*-0.39*	***0.65***	*-0.27*				
			-3.85%	***6.52%***	*80.83%*				
	/e/	Pre		0.40	**8.95**	0.59			9.93
				4.00%	**89.46%**	5.87%			99.33%
		Post		0	**7.71**	2.15			9.87
				0.00%	**77.14%**	21.52%			98.67%
		Diff.		*-0.40*	***-1.23***	*1.57*			
				-4.00%	***-12.32%***	*15.65%*			
	/æ/	Pre			8.59	**1.41**			10.00
					85.89%	**14.11%**			100.00%
		Post			4.63	**5.37**			10.00
					46.34%	**53.66%**			100.00%
		Diff.			*-3.96*	***3.96***			
					-39.55%	***39.55%***			
	/ʊ/	Pre					**0.17**	4.83	5.00
							3.33%	96.67%	100.00%
		Post					**1.45**	3.55	5.00
							28.95%	71.05%	100.00%
		Diff.					***1.28***	*-1.28*	
							25.62%	*-25.62%*	
	/uː/	Pre					0.61	**4.39**	5.00
							12.29%	**87.71%**	100.00%
		Post					0.53	**4.45**	4.98
							10.67%	**88.95%**	99.62%
		Diff.					*0.06*	***0.06***	
							1.24%	***1.24%***	

TABLE 4.7e.

Confusion Matrix showing the averaged scores and percentages of the responses in the Production Pretest and Posttest of the three vowel pairs by the subjects receiving production training only (N=11)

Target English Vowels		Production (Pre-/Post-test) Responses						Sum
		/ɪ/	/iː/	/e/	/æ/	/ʊ/	/uː/	
/ɪ/	Pre	**1.64**	7.91	0.00				9.55
		16.36%	79.09%	0.00%				95.45%
	Post	**4.55**	4.27	0.45				9.27
		45.45%	49.09%	4.55%				99.09%
	Diff.	***2.91***	*-3.64*	*0.45*				
		29.09%	*-30.00%*	*4.55%*				
/iː/	Pre	0.91	**8.36**	0.55				9.82
		9.09%	**83.64%**	5.45%				98.18%
	Post	0.18	**9.09**	0.73				10.00
		1.82%	**90.91%**	7.27%				100.00%
	Diff.	*-0.73*	***0.73***	*0.18*				
		-7.27%	***7.27%***	*1.82%*				
/e/	Pre		1.36	**7.82**	0.82			10.00
			13.64%	**78.18%**	8.18%			100.00%
	Post		1.36	**8.00**	0.55			9.91
			13.64%	**80.00%**	5.45%			99.09%
	Diff.		*0.00*	***0.18***	*-0.27*			
			0.00%	***1.82%***	*-2.73%*			
/æ/	Pre			7.91	**1.82**			9.73
				79.09%	**18.18%**			97.27%
	Post			4.18	**5.55**			9.73
				41.82%	**55.45%**			97.27%
	Diff.			*-3.73*	***3.73***			
				-37.27%	***37.27%***			
/ʊ/	Pre		0			**0.45**	4.27	4.73
			0.00%			**9.09%**	89.09%	98.18%
	Post		0.18			**1.82**	3.00	5.00
			3.64%			**36.36%**	60.00%	100.00%
	Diff.		*0.18*			***1.36***	*-1.27*	
			3.64%			***27.27%***	*-29.09%*	
/uː/	Pre					1.00	**4.00**	5.00
						20.00%	**80.00%**	100.00%
	Post					0.91	**4.09**	5.00
						18.18%	**81.82%**	100.00%
	Diff.					*0.09*	***0.09***	
						1.82%	***1.82%***	

TABLE 4.7f.

Confusion Matrix showing the averaged scores and percentages of the responses in the Production Pretest and Posttest of the three vowel pairs by the control subjects (N=10)

			Production (Pre-/Post-test) Responses						Sum
			/ɪ/	/iː/	/e/	/æ/	/ʊ/	/uː/	
Target English Vowels	/ɪ/	Pre	**1.40**	8.40	0.00				9.80
			14.00%	84.00%	0.00%				98.00%
		Post	**1.30**	8.10	0.60				10.00
			13.00%	81.00%	6.00%				100.00%
		Diff.	***-0.10***	*-0.70*	*0.60*				
			-1.00%	*-7.00%*	*6.00%*				
	/iː/	Pre	1.10	**8.00**	0.50				9.60
			11.00%	**80.00%**	5.00%				96.00%
		Post	1.00	**8.00**	0.60				9.60
			10.00%	**80.00%**	6.00%				96.00%
		Diff.	*-0.10*	***0.00***	*0.10*				
			-1.00%	***0.00%***	*1.00%*				
	/e/	Pre			**8.80**	1.10			9.90
					88.00%	11.00%			99.00%
		Post			**7.90**	1.90			9.80
					79.00%	19.00%			98.00%
		Diff.			***-0.90***	*0.80*			
					-9.00%	*8.00%*			
	/æ/	Pre			8.80	**0.90**			9.70
					88.00%	**9.00%**			97.00%
		Post			8.50	**1.40**			9.90
					85.00%	**14.00%**			99.00%
		Diff.			*-0.30*	***0.50***			
					-3.00%	***5.00%***			
	/ʊ/	Pre					**0.00**	4.30	4.30
							0.00%	86.00%	86.00%
		Post					**0.40**	4.30	4.70
							8.00%	86.00%	94.00%
		Diff.					***0.40***	*0.00*	
							8.00%	*0.00%*	
	/uː/	Pre					0.40	**4.60**	5.00
							8.00%	**92.00%**	100.00%
		Post					0.70	**4.30**	5.00
							14.00%	**86.00%**	100.00%
		Diff.					*0.30*	***-0.30***	
							6.00%	***-6.00%***	

4.2.3 STATISTICAL ANALYSIS: 5-WAY ANOVA VS. 3-WAY ANOVA

Same as the choice adopted in analyzing the perceptual data, a three-way ANOVA in lieu of a five-way will be utilized in the analysis of production scores. Besides the fact that the results obtained are similar for the both the three-way and five-way, it is also for the sake of bringing more succinct presentation and comprehension of the data. The five-way ANOVA with all the factors included is in Appendix L as reference.

4.2.4 RESULTS OF TRANSFER OF LEARNING TO THE PRODUCTION DOMAIN

4.2.4.1 FACTOR 1: PERCEPTUAL TRAINING APPROACHES

One of the goals of the present study is to compare the effectiveness of the two perpetual training approaches, HVPT and LVPT, on the modification of learning the three target vowel pairs. Besides perceptual learning, the transfer of learning to the production domain is of high relevance to the understanding of the efficacy of the perceptual training paradigms on production modification, let alone the exploration of the link between perception and production. To achieve this, the production scores were fed into a three-way repeated-measures ANOVA with Perceptual Training Type (HVPT, LVPT and Control) as between-subject factor, Test (pretest and posttest), and Vowel (front, low and back) as repeated-measures. Before reading the results of the ANOVA, readers may look at Table 4.8 on the next page which shows the averaged scores and percentages of the productions of the subjects across HVPT ($n = 17$), LVPT ($n = 15$) and the control group ($n = 10$) in the pretest and posttest.

The ANOVA revealed a significant main effect of Perceptual Training [$F(2,39)$ =

Table 4.8

Mean Test Scores, Mean Percentage, and Mean Difference of all the groups in Production Pretest and Posttest

		Production Performance					
		Pretest	%	Posttest	%	Difference	%
	Total Score:	10		10			
/ɪ/-/iː/	HVPT	5.06	50.63%	32.22	78.58%	27.15	27.95%
	S.D.:	0.70	7.05%	6.68	16.70%		
	LVPT	4.84	48.39%	6.42	64.15%	1.58	15.76%
		1.11	11.13%	1.57	15.70%		
	Control	4.70	47.00%	4.65	46.50%	-0.05	-0.50%
		0.50	4.96%	0.86	8.63%		
/e/-/æ/	HVPT	4.90	49.03%	7.65	76.53%	2.75	27.50%
		1.58	15.76%	1.91	19.14%		
	LVPT	5.18	51.79%	6.54	65.40%	1.36	13.62%
		0.95	9.52%	1.78	17.82%		
	Control	4.85	48.50%	4.65	46.50%	-0.20	-2.00%
		0.53	5.35%	1.35	13.48%		
/ʊ/-/uː/	HVPT	4.47	44.65%	6.12	61.18%	1.65	16.53%
		0.64	6.37%	1.35	13.52%		
	LVPT	4.55	45.52%	5.90	58.95%	1.34	13.43%
		0.62	6.22%	1.31	13.08%		
	Control	4.60	46.00%	4.70	47.00%	0.10	1.00%
		0.53	5.35%	0.71	7.07%		

12.05, p <.001], since both the HVPT and the LVPT groups performed better than the control group, and the HVPT group also outperformed LVPT. The main effect of Test [F(1,39) = 58.92, p <.001] was also robust, due to the overall improvement from the pretest to posttest across groups and vowels. Vowel [F(2,78) = 5.14, p = .008] was also a significant main effect, as the front and low vowel pairs were in general more accurately produced than the back vowel pair across groups and tests. Only the interaction between Test × Perceptual Training [F(2,39) = 17.41, p <.001] was robust; Vowel × Perceptual Training (p = .161), Test × Vowel (p = .546) or the three-way Test × Vowel × Perceptual Training (p = .440) interactions all failed to reach significance.

Planned comparisons with Bonferroni correction conducted on the Test × Perceptual Training interaction showed that only the HVPT group and LVPT group improved their production of the three vowel pairs with significance from the pretest to the posttest: HVPT for 23.92% [*F*(1,16) = 128.46, *p* <.001] and LVPT for 14.44% [*F*(1,14) = 41.95, *p* <.001] whereas the control group had a non-significant and slight decline of 0.50% (*p* = .456). Meanwhile, no significant differences between groups in the pretest (all at *p* = 1.00) was observed, which sets a fair ground for comparison of their performance. In the posttest, the HVPT group outperformed the LVPT group for 9.17% [*F*(1,30) = 5.69, *p* = .024] and control group for 25.39% [*F*(1,25) = 69.85, *p* <.001] and the LVPT group outscored the control group for 16.22% [*F*(1,23) = 15.62, *p* =.001]. The following box plots visualize the production performance of all the groups before and after training across all three target vowel pairs:

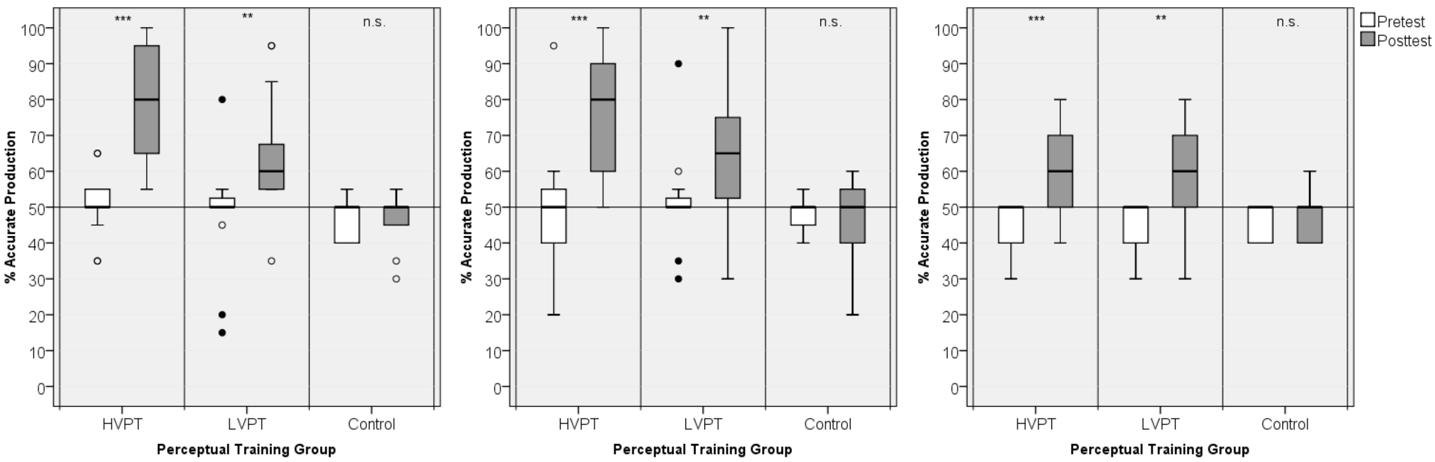


Figure 4.13 Mean percentages of accurate production of the trained groups (HVPT & LVPT) with significant difference between the pretest (white boxes) and the posttest (dark boxes) [*** = *p* < .001; ** = *p* < .005] and the control group with no significant differences [n.s. = non-significant] across /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right). Black dots and circles represent outliners.

4.2.4.2 FACTOR 2: PERCEPTUAL TRAINING INTENSITY

The previous analysis shows that the subjects' modification of production can be observed in both the HVPT and LVPT groups. A comparison of the intensity levels of the two interventions will yield further understanding of how this factor may contribute to the transfer of perceptual learning. The following two figures give a first glance of the performance of the two groups with different intensity levels:

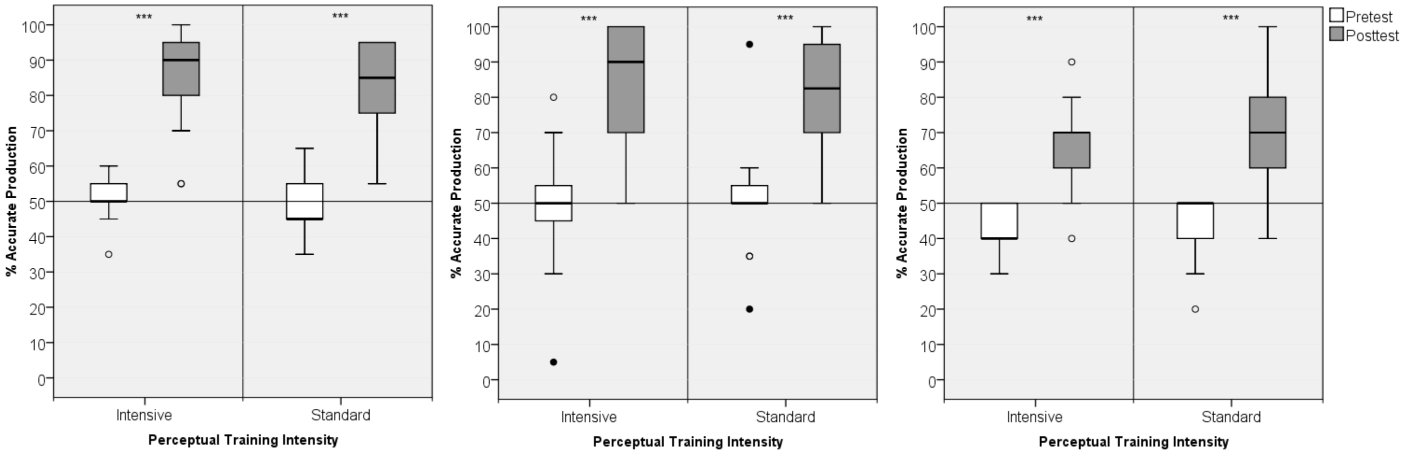


Figure 4.14a (From left to right) Mean percentages of accurate production of /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ respectively by the HVPT group with different perceptual training intensities. All showed significant differences between the pretest (white boxes) and posttest (dark boxes) performance (*** = $p < .001$). The black dots and circles represent outliners. The horizontal line indicates the chance level (50%) performance.

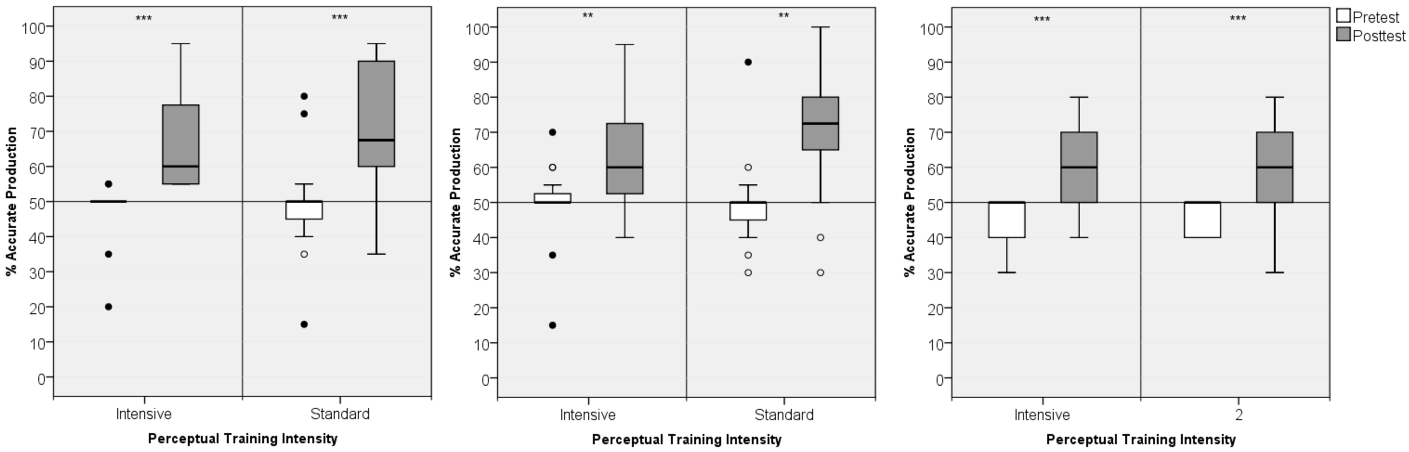


Figure 4.14b (From left to right) Mean percentages of accurate production of /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ respectively by the LVPT group with different perceptual training intensities. All showed significant differences between the pretest (white boxes) and posttest (dark boxes) performance (*** = $p < .001$; (*** = $p < .005$). The black dots and circles represent outliners. The horizontal line indicates the chance level (50%) performance.

Two separate three-way ANOVAs were conducted. For the HVPT groups, the three-way repeated-measures ANOVA for the HVPT group with Perception Training Intensity (Intensive, n = 8; Standard, n = 9) as between-subjects factor and Test (pretest, posttest) and Vowel (front, low, back) as repeated measures showed only a significant main effect of Test [$F(1,15) = 59.11$, $p <.001$], since the HVPT groups with the two intensity levels improved from the pretest to posttest across the three vowel pairs for 23.99% ($p < .001$). Vowel [$F(2,30) = 6.92$, $p = .003$] was a significant main effect as well, due to the better production accuracy rates observed among the front and low vowel pairs. Perceptual Training Intensity ($p = .745$) was not a robust effect, neither were the interactions of Test × Vowel ($p =.186$), Test × Perceptual Training Intensity ($p = .702$), Vowel × Perceptual Training Intensity ($p = .936$) or Test × Vowel × Perceptual Training Intensity ($p = .942$). All these showed that intensity level was not a major influential factor impacting the production results of the subjects.

Concerning the LVPT groups, the same three-way ANOVA with Perception Training Intensity (Intensive, n = 8; Standard, n = 7) as between-subjects factor and Test (pretest, posttest) and Vowel (front, low, back) as repeated measures showed a significant main effect of Test [$F(1,13) = 24.46$, $p < .001$], due to the significant changes in performance from the pretest to the posttest. The main effects of Vowel ($p = .552$) and Perceptual Training Intensity was not robust ($p = .552$); the interactions Test × Vowel ($p = .917$), Test × Perceptual Training Intensity ($p = .960$), Vowel × Perceptual Training Intensity ($p = .854$) and Test × Vowel × Perceptual Training Intensity ($p = .891$) were not significant either. This analysis again indicates that perceptual training intensity did not affect perceptual learning and hence the transfer to production, but any training in itself already contributed to the learning of the vowel contrasts.

4.2.4.3 FACTOR 3: EXPLICIT PRODUCTION TRAINING

Statistical analysis on the effects of production training on perceptual learning revealed only significant effects on the modification of the low vowel pair. The other two vowel pairs had demonstrable improvement, but statistically speaking they were not robust. It is thus interesting to gauge also the direct effect of production training on the production modification of the three target vowel pairs. Their performances of the three vowel pairs are shown in Figure 4.15:

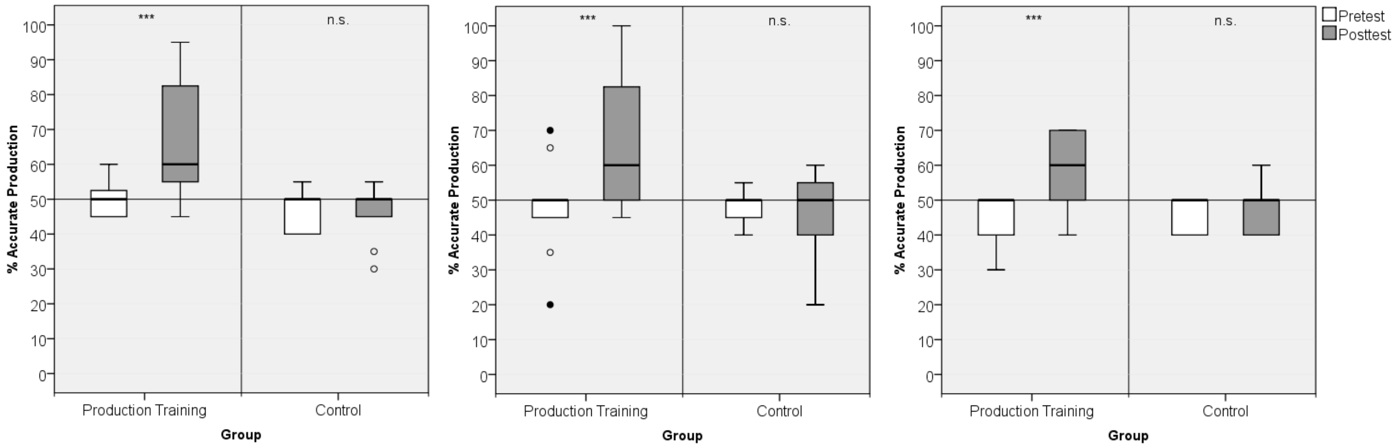


Figure 4.15 (From left to right) Mean percentages of accurate production of /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ respectively by the groups with and without production training. All experimental group showed significant differences between the pretest (white boxes) and posttest (dark boxes) performance (*** = $p < .001$). The black dots and circles represent outliners. The horizontal line indicates the chance level (50%) performance.

A three-way repeated-measures ANOVA with Production Training (With, n = 11; Without, n = 10) as between-subjects factor and Test (pretest, posttest) and Vowel (front, low, back) as repeated measures yielded significant main effects of Production Training [$F(1,19) = 9.21$, $p = .007$] and Test [$F(1,19) = 18.77$, $p < .001$], due to the overall improvement of the production of the target vowel pairs from the pretest to the posttest. There was also significant Test × Production Training interaction [$F(1,19) = 21.05$, $p < .001$], but Vowel ($p = .224$), Vowel × Production Training ($p = .436$), Test × Production Training ($p = .722$) and Test × Vowel × Production Training ($p = .647$) were not robust.

Planned comparisons with Bonferroni correction on the interaction Test × Production Training interaction revealed no significant differences in the pretest between the group with production pretest and the group without ($p = .856$). Only the performance in the posttest was robust [$F(1,19) = 15.44$, $p < .001$]. In addition, only the production training group demonstrated significant improvement for 17.42% from the pretest to posttest [$F(1,10) = 22.83$, $p < .001$] but the control group had very mild fluctuation for 0.5% after training ($p = .862$). The group with production also performed significantly better than the one without for 35.73% [$F(1,19) = 17.31$, $p < .001$]. As can be seen from the above analysis and the figures, production training can benefit the subjects in improving the vowel pairs to a large extent, even though highly varied individual differences can be observed. The fact that production training can only significantly benefit the subjects in the perception of only the low vowel pair, but that the results indicate that all three vowel pairs received gains in the production accuracy is worth more investigation and will be delved in Chapter 5.

4.2.4.4 Factor 4: Training Diversity

In the present design of the study, some groups were only given perceptual training whereas some were given an addition of explicit production training. Analysis above has shown that both the perceptual training and the production training alone can benefit the subjects' perception and production of the target vowel contrasts, though to different extents. The results of different groups will be compared again to shed light on the effectiveness of these training approaches. Figure 4.16a - c on the next page gives three box plots showed the percentage of accurate production for the three vowel pairs for all training types in the pretest and posttest.

To pursue this, a three-way repeated-measures ANOVA with Training Type (HVPT only [H], n = 17; HVPT with production [HP], n = 18; LVPT only [L], n = 15; LVPT with production [LP], n = 14; production only [P], n = 11) as a between-subject factor and Test (pretest, posttest) and Vowel (front, low, back) as repeated measures were submitted to compare the results of the types. The ANOVA yielded significant main effects of Test [$F(1,70) = 266.06$, $p < .001$], Vowel [$F(2,140) = 20.73$, $p = .001$] and Training Type ($p = .002$). Two interactions, Test × Training Type [$F(4,70) = 8.22$, $p < .001$] and Test × Vowel [$F(2,140) = 4.22$, $p = .017$], were significant as well, due to the overall improvement from the pretest to posttest across types and vowels. Vowel × Training Type ($p = .746$) and the three-way Test × Training Type × Vowel ($p = .893$) interactions were not significant.

Planned comparisons with Bonferroni correction conducted on the Test × Training Type interaction revealed that all subjects trained under different training types improved significantly from the pretest to posttest: the H type for 23.922% [$F(1,16) = 62.28$, $p < .001$], the HP type for 35.83% [$F(1,17) = 148.20$, $p < .001$], the L type for

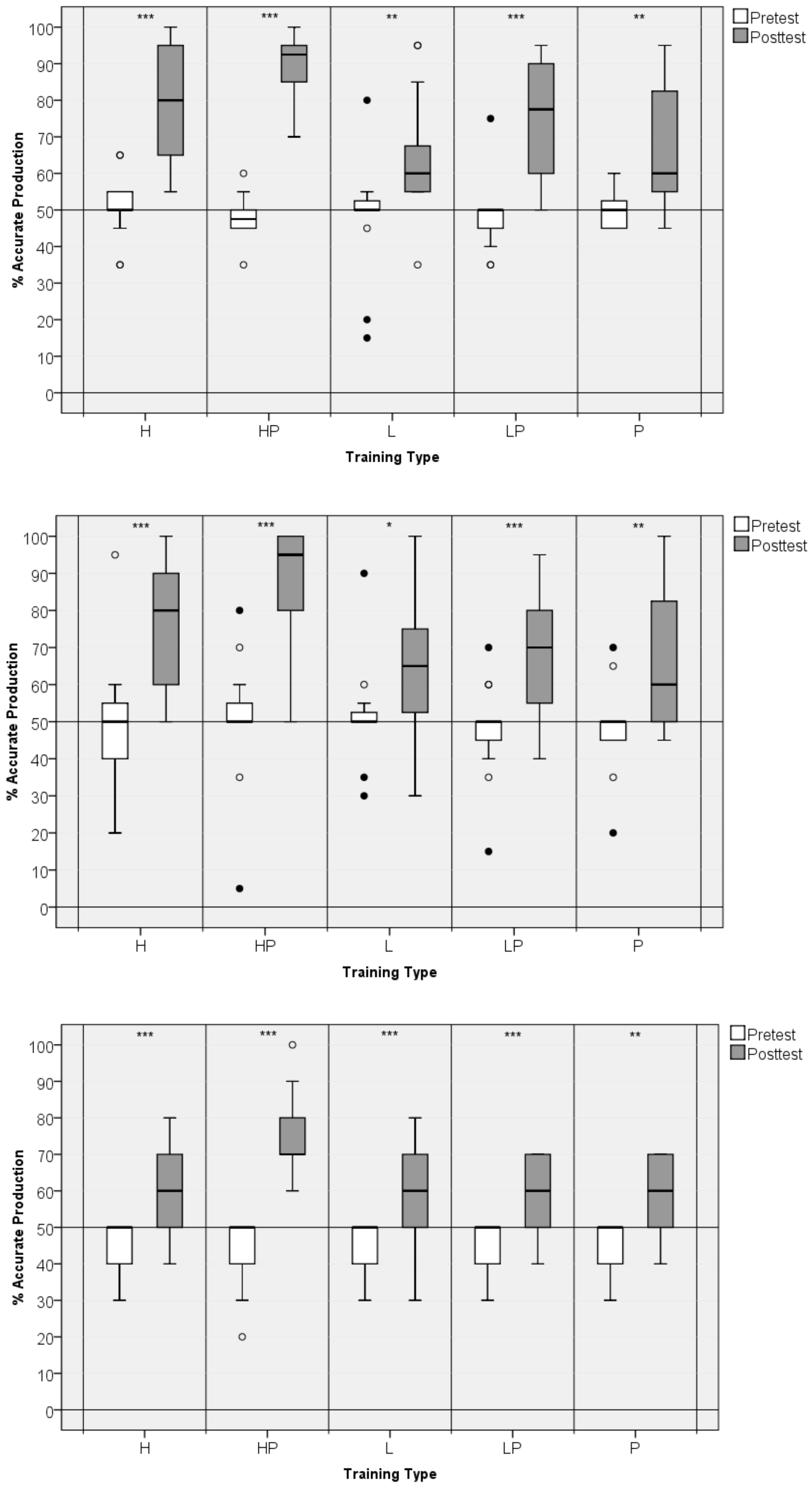


Figure 4.16a, b, c Mean percentages of accurate production of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from top to down) among the 5 training types. Significant differences were found between the pretest (white boxes) and the posttest (dark boxes) (*** = $p < .001$; ** = $p < .005$; * = $p < .050$). The circles and black dots represent outliers. The horizontal line shows the chance level (50%) performance.

14.44% [$F(1,14) = 26.49$, $p < .001$], the LP type for 20.36% [$F(1,13) = 56.31$, $p < .001$] and the P type for 17.42% [$F(1,10) = 22.83$, $p = .001$]. Meanwhile, no significant differences between types were observed at the pretest (all at $p = 1.00$). In the posttest, only the HP type outperformed other groups with significance: the HP type outperformed the H, L, LP and P type for 11.09% [$F(1,33) = 10.85$, $p = .002$], 20.26 % [$F(1,31) = 25.14$, $p < .001$], 15.65% [$F(1,30) = 19.93$, $p < .001$] and 18.15% [$F(1,27) = 15.25$, $p < .001$] respectively. The analysis revealed that all training types were useful in modifying the production of subjects, with the HP type being the most outstanding type.

For the Test × Vowel interaction, planned comparisons with Bonferroni correction also showed that all the groups improved the production accuracy rates of all three pairs of vowels from the pretest to posttest with significance: the front vowel pair for 25.63% [$F(1,70) = 171.28$, $p < .001$], the low vowel pair for 23.34% [$F(1,70) = 96.43$, $p < .001$] and the back vowel pair for 18.23% [$F(1,70) = 111.75$, $p < .001$]. Also, at the pretest, the accuracy rates of the front vowel pair was already significantly higher than the back vowel (for 4.62%, $p = .001$) while the low vowel pair was also higher than the back vowel (for 5.13%, $p = .022$); while at the posttest, the subjects still had higher production accuracy rate in the front vowel pair than the back (for 12.02%, $p < .001$) as well as the low vowel pair than the back (for 10.24%, $p < .001$), meaning that the difficulty in correctly producing the back vowels before and after training was higher.

4.2.4.5 Summary of the Results of Production Learning

The above analyses showed evidence that learners with production training can improve the production of the three target vowel contrasts, while perception training alone can also facilitate accurate production of the vowel pairs by having the perceptual learning transferred to the production domain, with HVPT being a more effective paradigm than LVPT in helping to improve the subjects' production. An addition of production training on a perceptual training (although it is only applicable to HVPT) benefits the subjects more as well. Yet, perceptual training intensity was not an influential factor in affecting the learning. Besides comparing the results of the pretest and the posttest, the scores in the Test of Contextualization can give further support to whether the training paradigms are useful in helping the production of the three target vowels at the sentence level. The results will be shown in the following section.

4.2.5 RESULTS IN PRODUCTION CONTEXTUALIZATION

Both the production pretest and posttest are not a measure of the naturalistic speech production of the subjects; Test of Contextualization (TC) was thus devised as a supplementary test to provide more information on the sentence-level production performance of the subjects after the training. The four factors that were included in the analyses above will also be included in the statistical computing.

4.2.5.1 FACTOR 1: PERCEPTUAL TRAINING ON CONTEXTUALIZATION

Figure 4.17 on the next page displays the TC performance of the groups across the vowel pairs. Transfer of perceptual learning to the production domain through the use of the two perceptual training paradigms, HVPT and LVPT, was detected after comparing

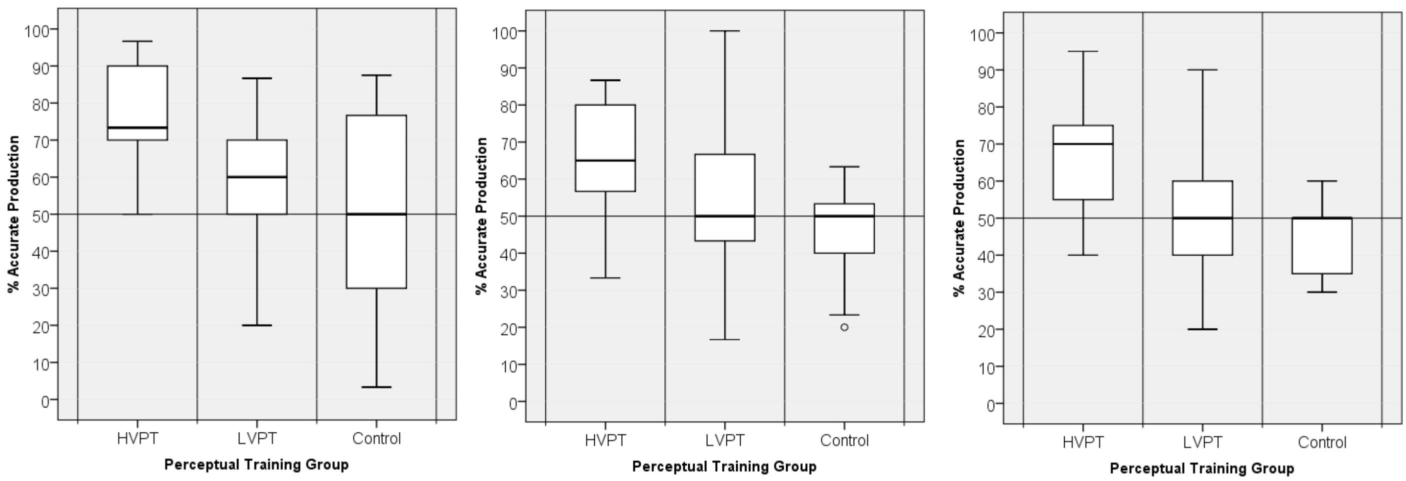


Figure 4.17 Mean percentages of accurate production of the of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right) in TC among the HVPT, the LVPT and the control group. The circle represents outliers.

the results of the production pretest and posttest, with HVPT being a more effective approach than LVPT. Now we further examine the transfer of perceptual learning to the production at sentence-level to gauge the extent to which the training can benefit the subjects.

A two-way repeated-measures ANOVA with Perceptual Training Type (HVPT, n = 17; LVPT, n = 15; Control, n = 10) as between-subject factor and Vowel (front, low and back) as repeated-measures revealed only a significant main effect of Perceptual Training Type [$F(2,39) = 12.74$, $p < .001$], which was because the HVPT group performed better than the LVPT for 15.70% [$F(1,30) = 11.94$, $p = .002$] and the control group for 22.49% [$F(1,25) = 26.80$, $p < .001$]. Even though the performance between the LVPT group and the control group was not significant ($p = .406$), the LVPT group still performed 7.43% better than the control group, which was a demonstrable amount that should not be neglected. However, the main effect of Vowel ($p = .104$) and the interaction Perceptual Training Type × Vowel were not significant ($p = .905$), indicating that the production performance at the sentence-level across the three vowel pairs were close to each other.

4.2.5.2 Factor 2: Perceptual Training Intensity on Contextualization

The two sets of box plots show the performance of the two perceptual training groups, i.e. the HVPT and the LVPT, with different intensity levels across the three vowel pairs:

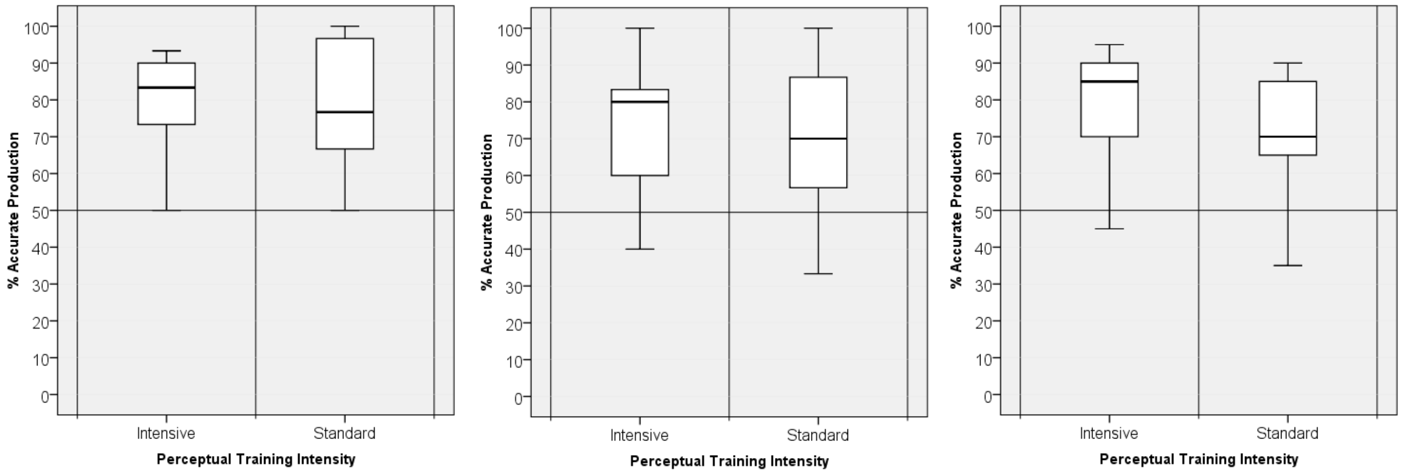


Figure 4.18a Mean percentages of accurate production of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right) in TC with different intensity levels in the HVPT groups.

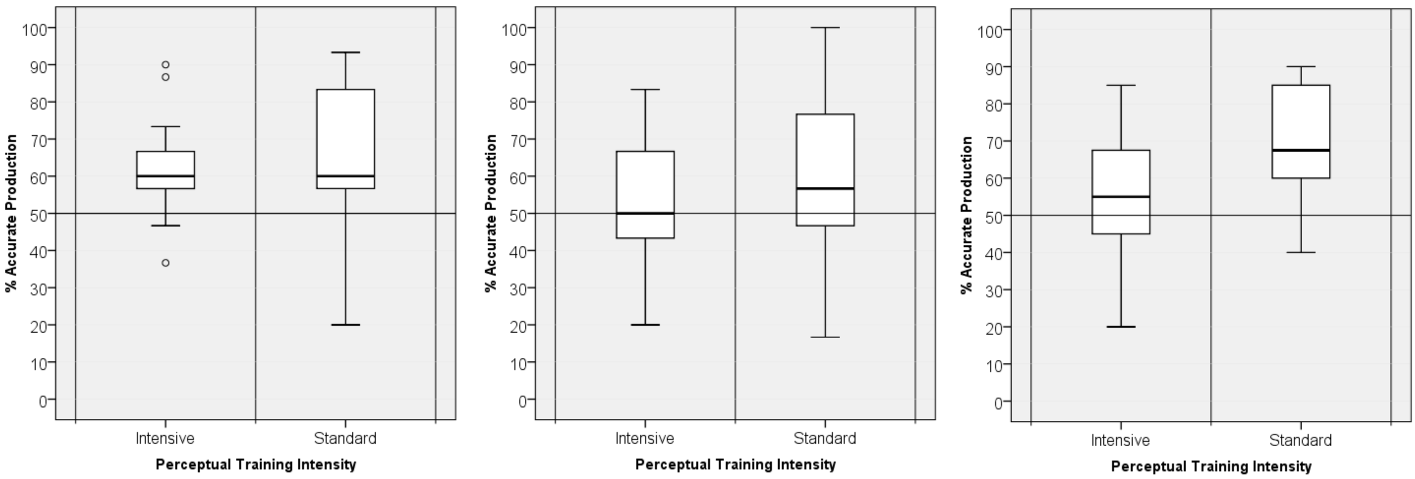


Figure 4.18b Mean percentages of accurate production of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right) in TC with different intensity levels in the LVPT groups. The circles represent outliners.

To investigate whether perceptual training intensity plays a role in impacting the production performance at the sentence level, the scores of accurate production in TC were submitted to two separate two-way repeated-measures of ANOVA with Perceptual Training Intensity (Intensive, Standard) as a between-subject factor and Vowel (front, low, back) as repeated measures for the HVPT and the LVPT groups.

The first ANOVA for the HVPT groups (Intensive, n = 9; Standard, n = 8) revealed no significant main effects: Vowel (p = .073) and Perceptual Training Intensity (p = .786). The interaction Perceptual Training Intensity × Vowel (p = .301) was not robust either. For the LVPT groups (Intensive, n = 8; Standard, n = 7), the ANOVA yielded similar results: Vowel (p = .167) and Perceptual Training Intensity (p = .214) were not significant and neither was the interaction Perceptual Training Intensity × Vowel (p = .281). Again, the analyses gave support that Perceptual Training Intensity was not an influential factor that would affect the production learning of the target vowel contrasts.

4.2.5.3 Factor 3: Explicit Production Training on Contextualization

Direct and explicit production training was found to be useful in improving the subjects’ production accuracy rates from the analysis in comparing the pretest and posttest results. It is worth checking the effects of contextualization as well. A two-way repeated-measures with Production Training (with, n = 11; without, n = 10) as a between-subject factor and Vowel (front, low, back) as repeated measures was submitted to compare the results of the TC so as to measure whether the production learning can be extended to the sentence level. The ANOVA yielded only a significant main effect of

Production Training [$F(1,19) = 7.86$, $p = .011$], since the group with production training performed better than the control group for 16.16% in general. Yet, Vowel ($p = .575$) and Vowel × Production Training ($p = .856$) were not robust. This finding shows that the production training can benefit the production accuracy only generally, but not for individual pairs. The following box plots display the mean percentages of accurate production in TC between the group with production training and the control group across the three target vowel pairs:

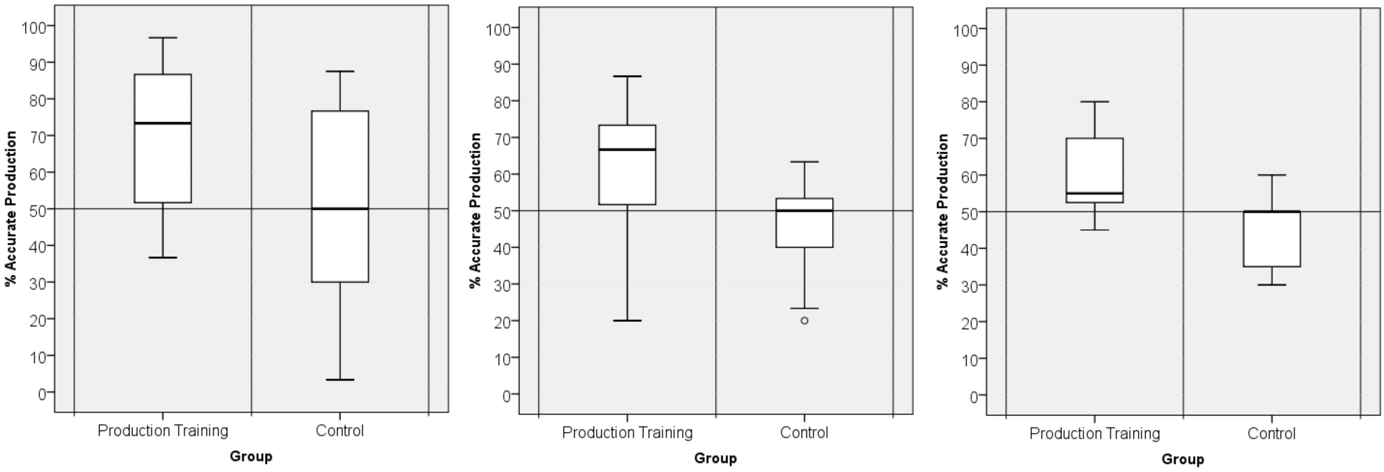


Figure 4.19 Mean percentages of accurate production of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from left to right) in TC of the group with production training and the group without.

4.2.5.4 FACTOR 4: TRAINING DIVERSITY ON CONTEXTUALIZATION

Figure 4.20a, b and c on the next page illustrate the production accuracy rates of the three vowel pair across the five training types. A two-way repeated-measures ANOVA with Training Type (HVPT only [H], n = 17; HVPT with production [HP], n = 18; LVPT only [L], n = 15; LVPT with production [LP], n = 14; production only [P], n = 11) as a between-subject factor and Vowel (front, low, back) as repeated measures were submitted to compare the results of TC across different training types. A significant

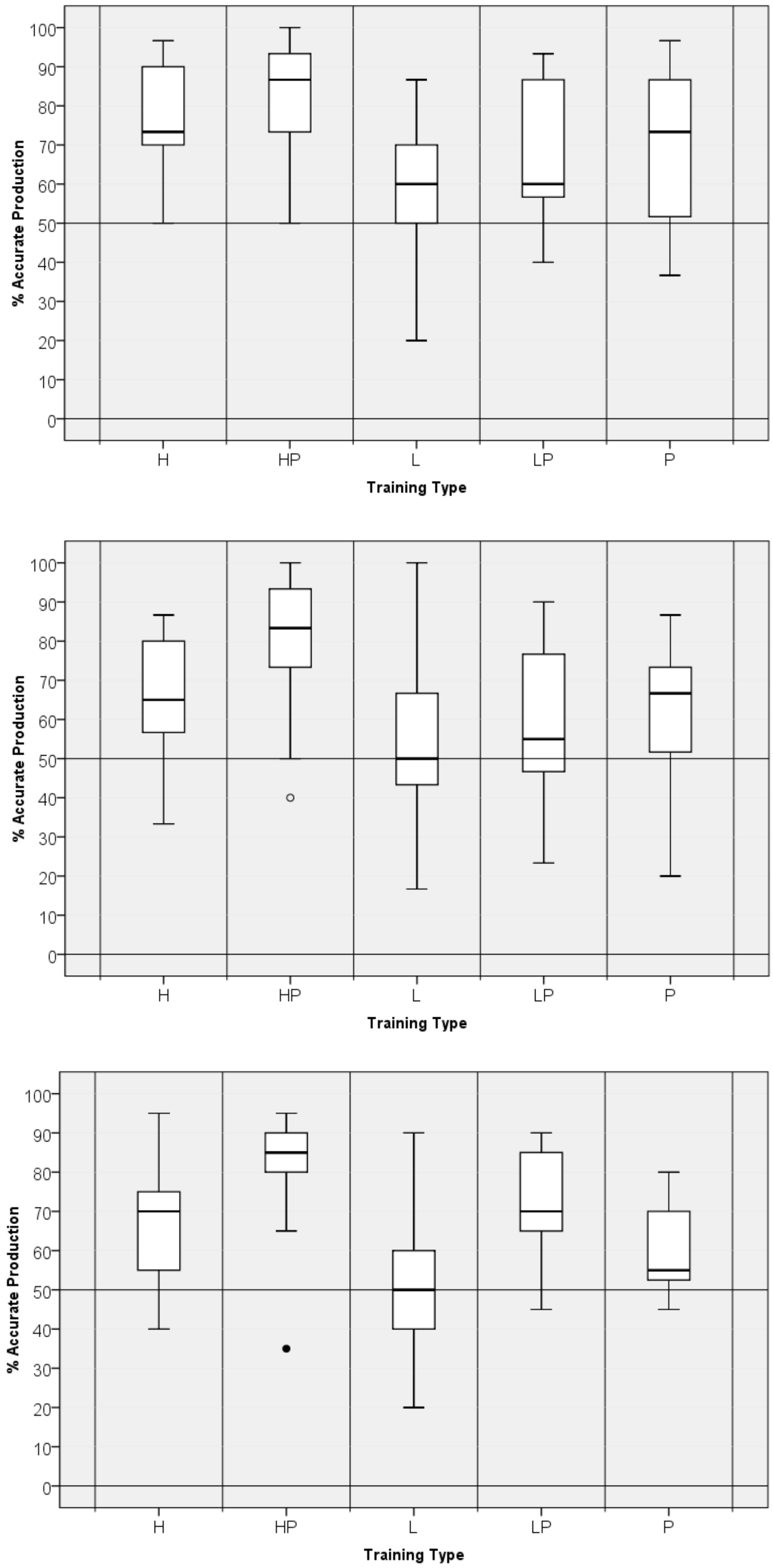


Figure 4.20a, b, c Mean percentages of accurate production of the three vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ (from top to down) among the 5 training types. The circles and black dots represent outliers. The horizontal line shows the chance level (50%) performance.

main effect of Training Type was discovered [$F(4,70) = 8.76$, $p = .001$], for the HP type demonstrated significantly better scores than the L type for 26.59% [$F(1,31) = 34.61$, $p < .001$], the LP type for 16.56% [$F(1,30) = 11.15$, $p = .002$] and the P type for 17.86% [$F(1,27) = 11.97$, $p = .002$]. There was also a significant main effect of Vowel [$F(2,140) = 5.15$, $p = .025$], but pairwise comparisons showed no robust differences between each vowel pair. Training Type × Vowel interaction was insignificant ($p = .051$). The analysis again shows that the HP type was a more effective approach than the others in general even at the production at the sentence level, whereas the L type appeared to be the least effective among all.

4.2.5.5 Summary of the Results in Test of Contextualization

Investigating the subjects' production performance at the sentence level, i.e. Test of Contextualization, offers a more complete picture of whether the transfer of perceptual learning can be transferred to the production domain and to what extent direct production training can benefit the learning of a non-native contrast. The above analyses show that the subjects trained under perceptual training paradigms, particularly HVPT, could demonstrate better and more accurate production of the target vowels at the sentence level than those who were trained under the LVPT and those without. Besides that the production training can directly help enhance the production accuracy of the subjects, it is found that if the subjects are provided with HVPT and an addition of production training, they will perform better than those with solely the perception or production paradigm. This claim appears not to be a novel discovery, since it is ordinary for learners who are given more training to perform better than those with less. Yet, those who were given LVPT with production training performed no less than those with only LVPT, but were not better than those with only the production training. These findings will be further discussed in detail in Chapter 5.

Before discussing the findings, it is worth looking at the productions made by the subjects from another perspective: their acoustic dimensions.

4.2.6 Acoustic Analysis of Production Performance (Pre vs. Post)

An acoustic analysis on the productions of the English vowels can offer an objective evaluation and more comprehensive understanding of the production performance of the subjects. Comparing the acoustic properties of the native speakers' vowel productions with those produced by the subjects in the pretest and the posttest can further confirm the transcription reliability and determine to what extent different training interventions had on modifying the subjects' production performance. For the sake of easier understanding and presentation, the subjects' performance in terms of the acoustic will be grouped according to six training types, i.e. groups trained under HVPT (H), HVPT with production training (HP), LVPT (L), LVPT with production training (LP), production training only (P) and the control group (C). All results in the statistical and acoustic measurements will be presented under these six training types. The mean values (with standard deviations shown in italics) of the first two formant frequencies and durations of the three target vowel pairs produced by both the native speakers and the subjects are firstly shown on the next several pages.

TABLE 4.9a

Duration and F1/F2 Frequencies for /ɪ/-/iː/ produced by Three Native English Speakers and Six Groups of Male Subjects (standard deviations in italics)

Native (male)	/ɪ/				/iː/			
F1	445				300			
F2	1850				2403			
Duration	149				198			
	Pretest		Posttest		Pretest		Posttest	
HP (male)								
F1	409	*58*	426	*36*	400	*39*	335	*42*
F2	2181	*69*	2052	*84*	2113	*153*	2177	*175*
Duration	156	*25*	122	*21*	181	*68*	217	*63*
H (male)								
F1	418	*44*	411	*69*	362	*55*	324	*40*
F2	2071	*139*	1947	*245*	2059	*193*	2296	*225*
Duration	140	*24*	140	*23*	204	*49*	255	*40*
LP (male)								
F1	395	*26*	390	*60*	372	*28*	312	*22*
F2	2154	*55*	2046	*96*	2110	*165*	2287	*48*
Duration	169	*33*	129	*22*	199	*57*	248	*72*
L (male)								
F1	391	*44*	374	*19*	406	*39*	354	*8*
F2	2232	*147*	2134	*126*	2208	*188*	2353	*185*
Duration	128	*25*	120	*10*	170	*26*	217	*38*
P (male)								
F1	427	*80*	417	*26*	441	*78*	368	*33*
F2	2114	*123*	1997	*51*	2071	*76*	2145	*91*
Duration	148	*51*	129	*66*	170	*86*	156	*74*
C (male)								
F1	482	*56*	389	*31*	419	*109*	346	*67*
F2	2249	*136*	2088	*76*	2266	*222*	2037	*200*
Duration	133	*14*	136	*21*	197	*39*	222	*23*

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

TABLE 4.9b

Duration and F1/F2 Frequencies for /ɪ/-/iː/ produced by Three Native English Speakers and Six Groups of Female Subjects (standard deviations in italics)

Native (female)	/ɪ/				/iː/			
F1	553				309			
F2	2346				2187			
Duration	197				229			
	Pretest		Posttest		Pretest		Posttest	
HP (female)								
F1	470	*43*	504	*49*	442	*64*	400	*44*
F2	2224	*355*	2081	*328*	2205	*415*	1992	*358*
Duration	163	*25*	127	*30*	211	*35*	269	*68*
H (female)								
F1	503	*47*	530	*104*	453	*52*	453	*92*
F2	2052	*227*	2134	*224*	1986	*310*	2129	*297*
Duration	131	*19*	132	*27*	172	*29*	224	*28*
LP (female)								
F1	467	*45*	515	*60*	462	*95*	429	*61*
F2	2368	*546*	2038	*285*	2265	*271*	2066	*305*
Duration	199	*73*	150	*23*	290	*138*	327	*112*
L (female)								
F1	467	*34*	530	*93*	462	*54*	458	*99*
F2	2034	*249*	2114	*201*	2137	*268*	2098	*295*
Duration	152	*31*	127	*26*	191	*37*	254	*66*
P (female)								
F1	479	*62*	484	*69*	471	*83*	414	*32*
F2	2217	*365*	2523	*641*	2147	*394*	2120	*417*
Duration	130	*27*	132	*23*	145	*28*	240	*54*
C (female)								
F1	469	*93*	531	*95*	532	*170*	444	*73*
F2	2122	*148*	2264	*480*	2164	*208*	3020	*1719*
Duration	127	*15*	130	*15*	156	*24*	209	*45*

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

TABLE 4.9c

Duration and F1/F2 Frequencies for /e/-/æ/ *produced by Three Native English Speakers and Six Groups of Male Subjects (standard deviations in italics)*

Native (male)	/e/				/æ/			
F1	661				891			
F2	1763				1625			
Duration	152				209			
	Pretest		Posttest		Pretest		Posttest	
HP (male)								
F1	761	*84*	718	*47*	766	*77*	802	*83*
F2	2261	*136*	1836	*125*	1811	*100*	1753	*131*
Duration	153	*27*	163	*41*	176	*54*	224	*48*
H (male)								
F1	718	*23*	665	*19*	754	*25*	757	*41*
F2	1746	*154*	1754	*98*	1737	*61*	1689	*141*
Duration	170	*17*	165	*18*	179	*27*	227	*34*
LP (male)								
F1	699	*59*	677	*48*	745	*53*	809	*66*
F2	1849	*35*	1838	*71*	1825	*41*	1684	*73*
Duration	203	*54*	175	*49*	217	*55*	267	*78*
L (male)								
F1	722	*48*	703	*17*	774	*36*	817	*53*
F2	1734	*110*	1792	*112*	1794	*56*	1642	*145*
Duration	176	*65*	138	*14*	156	*36*	207	*41*
P (male)								
F1	704	*68*	673	*57*	715	*67*	764	*92*
F2	1770	*121*	1771	*86*	1754	*62*	1706	*33*
Duration	169	*58*	143	*48*	158	*37*	251	*153*
C (male)								
F1	736	*117*	726	*52*	756	*85*	773	*66*
F2	1869	*82*	1831	*99*	1793	*48*	1734	*76*
Duration	163	*31*	160	*29*	203	*47*	177	*34*

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

TABLE 4.9d

Duration and F1/F2 Frequencies for /e/-/æ/ produced by Three Native English Speakers and Six Groups of Female Subjects (standard deviations in italics)

Native (female)	/e/				/æ/			
F1	778				938			
F2	2147				1846			
Duration	182				228			
	Pretest		Posttest		Pretest		Posttest	
HP (female)								
F1	860	*111*	902	*115*	918	*114*	938	*103*
F2	1976	*334*	1849	*355*	1886	*225*	1781	*183*
Duration	266	*139*	168	*54*	212	*42*	211	*50*
H (female)								
F1	936	*117*	902	*122*	942	*104*	970	*118*
F2	1755	*237*	1815	*244*	1858	*535*	1832	*151*
Duration	153	*21*	163	*31*	176	*23*	206	*26*
LP (female)								
F1	882	*148*	830	*84*	914	*92*	942	*100*
F2	1864	*252*	1981	*332*	1758	*262*	1649	*360*
Duration	236	*64*	212	*61*	242	*74*	303	*101*
L (female)								
F1	848	*131*	867	*71*	925	*102*	958	*111*
F2	1746	*180*	1993	*195*	1751	*231*	1726	*264*
Duration	180	*38*	158	*45*	194	*36*	210	*50*
P (female)								
F1	874	*112*	830	*132*	913	*58*	971	*100*
F2	1834	*268*	1950	*38*	1855	*210*	1908	*100*
Duration	173	*40*	155	*58*	197	*45*	206	*45*
C (female)								
F1	884	*100*	800	*58*	889	*81*	853	*44*
F2	1840	*285*	1855	*348*	2009	*97*	2013	*191*
Duration	142	*35*	160	*38*	160	*19*	193	*21*

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

TABLE 4.9e

Duration and F1/F2 Frequencies for /ʊ/-/uː/ produced by Three Native English Speakers and Six Groups of Male Subjects (standard deviations in italics)

Native (male)	/ʊ/				/uː/			
F1	394				363			
F2	1546				1405			
Duration	147				200			
	Pretest		Posttest		Pretest		Posttest	
HP (male)								
F1	501	*50*	464	*34*	479	*69*	433	*40*
F2	1219	*139*	1408	*333*	1458	*313*	1568	*332*
Duration	172	*59*	179	*36*	178	*50*	248	*86*
H (male)								
F1	416	*19*	420	*81*	384	*26*	385	*39*
F2	1309	*194*	1347	*257*	1296	*169*	1407	*301*
Duration	179	*14*	198	*74*	206	*29*	257	*41*
LP (male)								
F1	455	*38*	465	*26*	408	*24*	435	*54*
F2	1207	*41*	1454	*86*	1430	*329*	1658	*242*
Duration	196	*27*	200	*44*	214	*46*	264	*89*
L (male)								
F1	489	*26*	441	*89*	391	*31*	444	*81*
F2	1225	*58*	1443	*220*	1345	*316*	1622	*388*
Duration	130	*27*	153	*8*	159	*19*	188	*41*
P (male)								
F1	578	*108*	531	*48*	589	*93*	474	*60*
F2	1315	*133*	1174	*72*	1299	*194*	1503	*387*
Duration	184	*82*	146	*63*	185	*73*	155	*55*
C (male)								
F1	535	*535*	457	*457*	443	*443*	535	*535*
F2	1469	*351*	1346	*201*	1554	*260*	1469	*351*
Duration	180	*39*	165	*27*	186	*35*	180	*39*

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

TABLE 4.9f

Duration and F1/F2 Frequencies for /ʊ/-/uː/ produced by Three Native English Speakers and Six Groups of Female Subjects (standard deviations in italics)

Native (female)	/ʊ/				/uː/			
F1	451				339			
F2	1723				1894			
Duration	108				143			
	Pretest		Posttest		Pretest		Posttest	
HP (female)								
F1	508	*41*	498	*37*	471	*56*	445	*45*
F2	1510	*198*	1437	*186*	1559	*184*	1537	*144*
Duration	202	*46*	231	*71*	225	*48*	274	*67*
H (female)								
F1	516	*56*	506	*52*	478	*48*	454	*29*
F2	1486	*117*	1407	*158*	1833	*869*	1489	*143*
Duration	161	*21*	188	*36*	183	*50*	224	*35*
LP (female)								
F1	527	*80*	518	*34*	485	*61*	460	*39*
F2	1477	*198*	1427	*168*	1440	*201*	1405	*82*
Duration	241	*94*	245	*81*	264	*89*	315	*73*
L (female)								
F1	497	*72*	512	*47*	499	*61*	464	*29*
F2	1701	*136*	1388	*142*	1480	*190*	1445	*88*
Duration	181	*42*	179	*46*	202	*47*	247	*60*
P (female)								
F1	488	*72*	464	*35*	452	*55*	434	*28*
F2	1456	*149*	1508	*176*	1460	*204*	1554	*183*
Duration	180	*45*	208	*22*	213	*76*	243	*55*
C (female)								
F1	556	*78*	483	*46*	533	*113*	459	*42*
F2	1413	*273*	1323	*176*	1537	*316*	1476	*191*
Duration	144	*20*	176	*32*	181	*49*	222	*61*

Note: F1 and *F2* values are measured in Hertz while duration in millisecond

The first two formant frequencies (*F1*[15] and *F2*[16]) and the vowel duration of the productions made by all six groups of subjects and the six native speakers who produced the training stimuli in the study were measured by using the *Praat* analysis software (Boersma & Weenink, 2002). The third formant frequency (*F3*) was not of interest since the first two formant frequencies can already provide adequate information to distinguish all English vowels (except the rhotic vowel /ɝ/ in General American accent, which is not the focus of the present investigation) from one another (Ladefoged, 2005). The formant frequency measurements of each vowel, measured at the vowel midpoint, were estimated by the formant tracking function in *Praat*. The temporal measurements of the vowels were measured from the onset to the offset of periodic energy in F2 from dual spectrogram and waveform displays.

Separate three-way repeated-measures ANOVAs with Training Type (HVPT only [H], n = 17; HVPT with production [HP], n = 18; LVPT only [L], n = 15; LVPT with production [LP], n = 14; production only [P], n = 11; control group [C], n = 10) as a between-subject factor, Test (pretest, posttest) and Vowel (6 vowels: /ɪ/, /iː/, /e/, /æ/, /ʊ/, /uː/) as repeated-measures were carried out on F1 and F2 values and vowel durations of the productions made by male and female subjects.

For F1 values, the ANOVA on the male subjects yielded significant main effects of Test [$F(1,23) = 6.32$, $p = .019$], Vowel [$F(5,115) = 2565.69$, $p < .001$] and Training Type [$F(5,23) = 3.35$, $p = .020$]. There were also significant Test × Vowel [$F(5,115) = 5.90$, p

[15] The first formant frequency of vowels (F1) corresponds to the vowel openness: the more open a vowel is, the higher F1 values are.
[16] The second formant frequency (F2) represents vowel frontness/backness. Front vowels have higher F2 frequencies than back vowels.

$< .001$] and Test × Vowel × Training Type [$F(25,115) = 1.94$, $p = .010$] interactions. However, Test × Training Type ($p = .832$), Vowel × Training Type ($p = .067$) were not robust. Post-hoc pairwise comparisons using the Bonferroni correction on Test × Vowel interaction separated the vowels as follows: in the pretest, /ɪ/, /iː/, /ʊ/ and /uː/ had similar F1 values whereas /e/ and /æ/ had similar F1 values, meaning that the members of these two groups of vowels were similar to each other in terms of vowel openness. In the posttest, except that /ɪ/ and /uː/ had similar F1 values, the other four vowels were significantly different from each other, meaning that the subjects began to differentiate the vowel in terms of vowel openness. Also, vowels /iː/ and /e/ became less open while /æ/ became more open after training (all at $p < .001$). Post-hoc pairwise comparisons using the Bonferroni correction on Test × Vowel × Training Type interaction further showed that there were no significant differences among training types across all six vowels in the pretest, meaning that all training types were well-matched prior to training. In the posttest, the H, HP and P types differentiated the vowel as follows: /ɪ/, /ʊ/ and /uː/ had similar F1 values; the other three vowels, /iː/, /e/ and /æ/ were significantly different from one another in terms of vowel openness (all at $p < .001$). The L and LP types only separated /e/ and /æ/ as having different mouth openness (all at $p < .001$); /ɪ/, /iː/, /ʊ/ and /uː/ were similar. The control group showed that /ɪ/, /iː/, /ʊ/ and /uː/ had similar F1 values while /e/ and /æ/ also had similar vowel openness.

Another ANOVA on the F1 values of vowels produced by female subjects showed that only the main effect of Vowel [$F(5,115) = 781.33$, $p < .001$] and Test × Vowel [$F(5,250) = 6.48$, $p < .001$] interaction were robust. Nevertheless, the main effects of

Test (p = .447), Training Types (p = .693), Test × Training (p = .540), Type Vowel × Training Type (p = .732) and Test × Vowel × Training Type (p = .905) were not significant. Post-hoc pairwise comparisons using the Bonferroni correction on Test × Vowel interaction revealed that in the pretest, all groups separated the vowels as follows: /ɪ/, /iː/ and /uː/ had similar F1 values, while /e/, /æ/ and /ʊ/ had different vowel openness. In the posttest, except that /ɪ/ and /ʊ/ had similar F1 values, all other four vowels distinguish themselves from one another in terms of vowel openness (all at $p < .001$), meaning that the subjects attempted to separate the vowels in the posttest.

For F2, the ANOVA on the male subjects' productions yielded significant main effect of Vowel [$F(5,115) = 141.54$, $p < .001$] and Test × Vowel [$F(5,115) = 2.67$, $p = .026$] interaction. The main effect of Training Type (p = .670) and Test (p = .441) were not robust, and neither were Test × Training Type (p = .052), Vowel × Training Type (p = .359) and Test × Vowel × Training Type (p = .561). Post-hoc pairwise comparisons using the Bonferroni correction on Test × Vowel interaction separated the vowels in the pretest: /ɪ/, /iː/, /e/ and /æ/ had similar F2 values, while /ʊ/ and /uː/ also had similar F2 values; in the posttest, only /ɪ/ and /iː/ had similar values and were the most front vowels, while other four vowels had significantly different F2 values with others, meaning that the subjects began to differentiate the vowels in terms of vowel frontness/backness. Vowels /ɪ/ (p = .003) and /æ/ (p = .011) were produced fronter in the posttest than in the pretest; whereas vowel /uː/ (p = .022) became backer in the posttest.

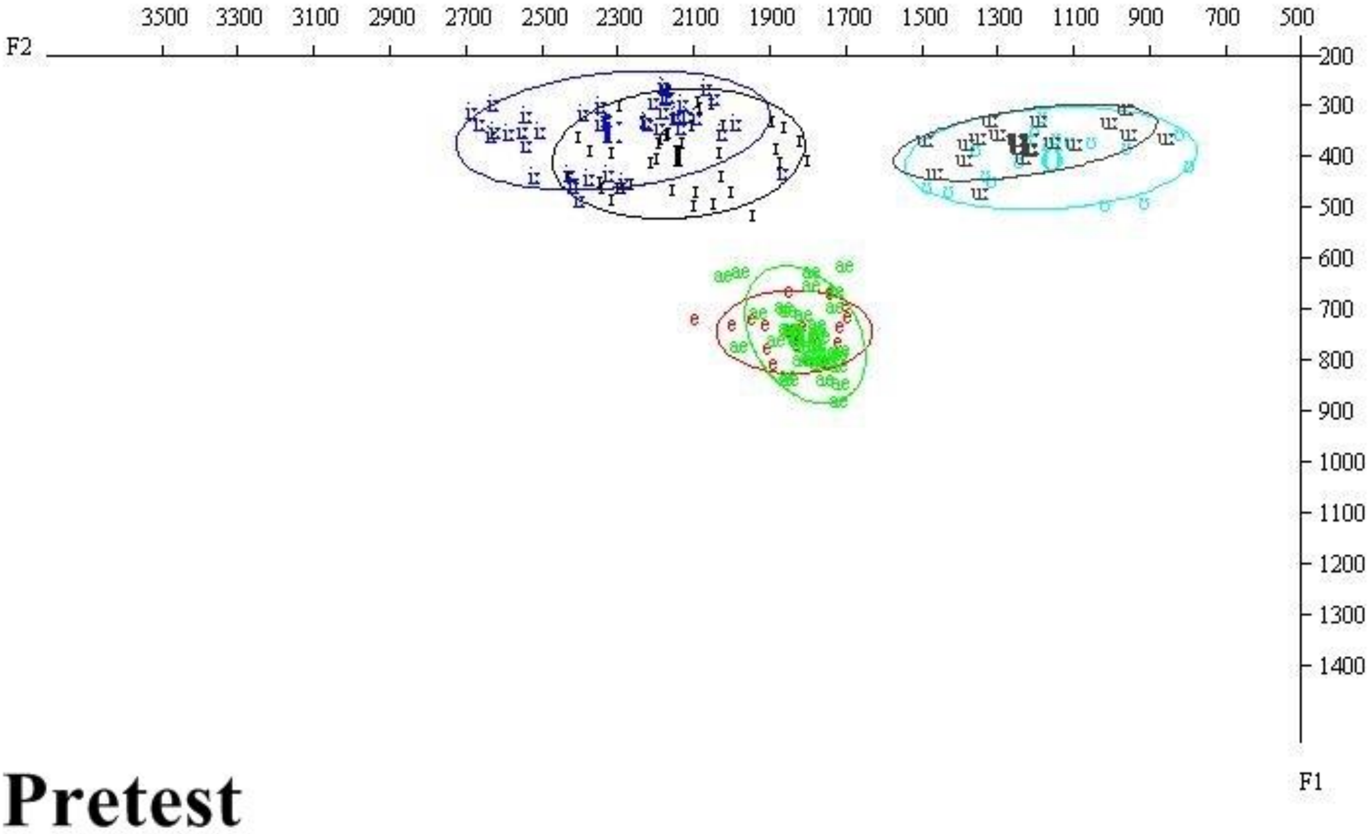
For female subjects, the ANOVA submitted to the compare the F2 values of the productions yielded only a significant main effect of Vowel [$F(5,250) = 60.49$, $p < .001$].

Post-hoc pairwise comparisons using the Bonferroni correction separated the vowels: /ɪ/ and /iː/; /e/ and /æ/; and finally /ʊ/ and /uː/ in general. Since other main effects and interactions were not robust, it appears that the F2 values of the productions made by the female subjects did not differ significantly in the pretest and posttest.

Concerning the vowel duration, the ANOVA on the male subjects' productions showed that the main effects of Test [$F(1,23) = 4.78$, $p = .039$] and Vowel [$F(5,115) = 32.26$, $p < .001$] were significant. There was also robust Test × Vowel [$F(5,115) = 10.30$, $p < .001$] interaction. The main effect of Training Type ($p = .422$) and Test × Training Type ($p = .766$), Vowel × Training Type ($p = .410$) and Test × Vowel × Training Type ($p = .229$) were not significant. Post-hoc pairwise comparisons using the Bonferroni correction on Test × Vowel showed that in the pretest, /ɪ/ was the shortest vowel, while /iː/, /e/, /æ/, /ʊ/ and /uː/ had not significant differences in durations among themselves. In the posttest, /ɪ/ was still the shortest vowel, followed by /e/, /ʊ/, /iː/, /uː/ and /æ/. The vowel duration of the vowel /ɪ/ was shorter ($p = .002$) while those of the vowels /iː/ ($p = .002$), /æ/ ($p = .001$) and /uː/ ($p = .013$) were longer compared with the productions in the pretest. These showed that the subjects attempted to differentiate the vowels by durations in the posttest.

For the duration of the female subjects' vowel productions, the ANOVA revealed significant effects of Training Type [$F(5,50) = 5.18$, $p = .001$], Test [$F(1,50) = 4.79$, $p = .033$] and Vowel [$F(5,250) = 56.30$, $p < .001$]. There was robust Test × Vowel interaction as well. However, Test × Training Type ($p = .828$), Vowel × Training Type ($p = .148$) and Test × Vowel × Training Type ($p = .125$) were not significant. Post-hoc pairwise comparisons using the Bonferroni correction on Test × Vowel showed that in the pretest, the vowel /ɪ/ was the shortest among all, but the vowel durations of /iː/, /e/, /æ/, /ʊ/ and /uː/ had no significant differences. In the posttest, /ɪ/ was still the shortest vowel, followed by /e/, /ʊ/, /æ/, /iː/ and /uː/. Meanwhile, when compared to the pretest, the length of the vowels /ɪ/ (p = .002), /e/ (p = .017), and /ʊ/ (p = .005) were shorter in the posttest; whereas /iː/ ($p < .001$), /æ/ (p = .008), and /uː/ ($p < .001$) were longer in the posttest.

To better illustrate and compare the production performance of the subjects, the F1-F2 space plots with the six target English vowels produced by the six groups of subjects in the pretest and the posttest are provided in Figures 4.21a – l. They are put in juxtaposition (above: pretest; below: posttest) for comparison. From the figures, there was a pairing of the vowels (/ɪ/ with /iː/, /e/ with /æ/, and /ʊ/ with /uː/) in the pretest across all six groups. The F1 and F2 frequencies of the vowel pairs are so close that they overlapped and are merged to become almost a single vowel. Recall also the results of the transcription evaluation task on the subjects' productions, previous documented evidence (e.g. Bolton & Kwok, 1990; Chan, 2010; Hung, 2002; Meng et al., 2007) reporting that the two long and short vowel pairs (/ɪ/-/iː/ and /ʊ/-/uː/) tend to be indistinguishable and that /e/ is always a substitution for /æ/ among Hong Kong Cantonese ESL learners are both confirmed. Moreover, when the pretest productions are compared to the native productions of the Cantonese vowels which are the predicted assimilation targets, it is evident that the conflations of /ɪ/ with /iː/, /e/ with /æ/, and /ʊ/ with /uː/ are the manifestation of neutralization through using the respective three



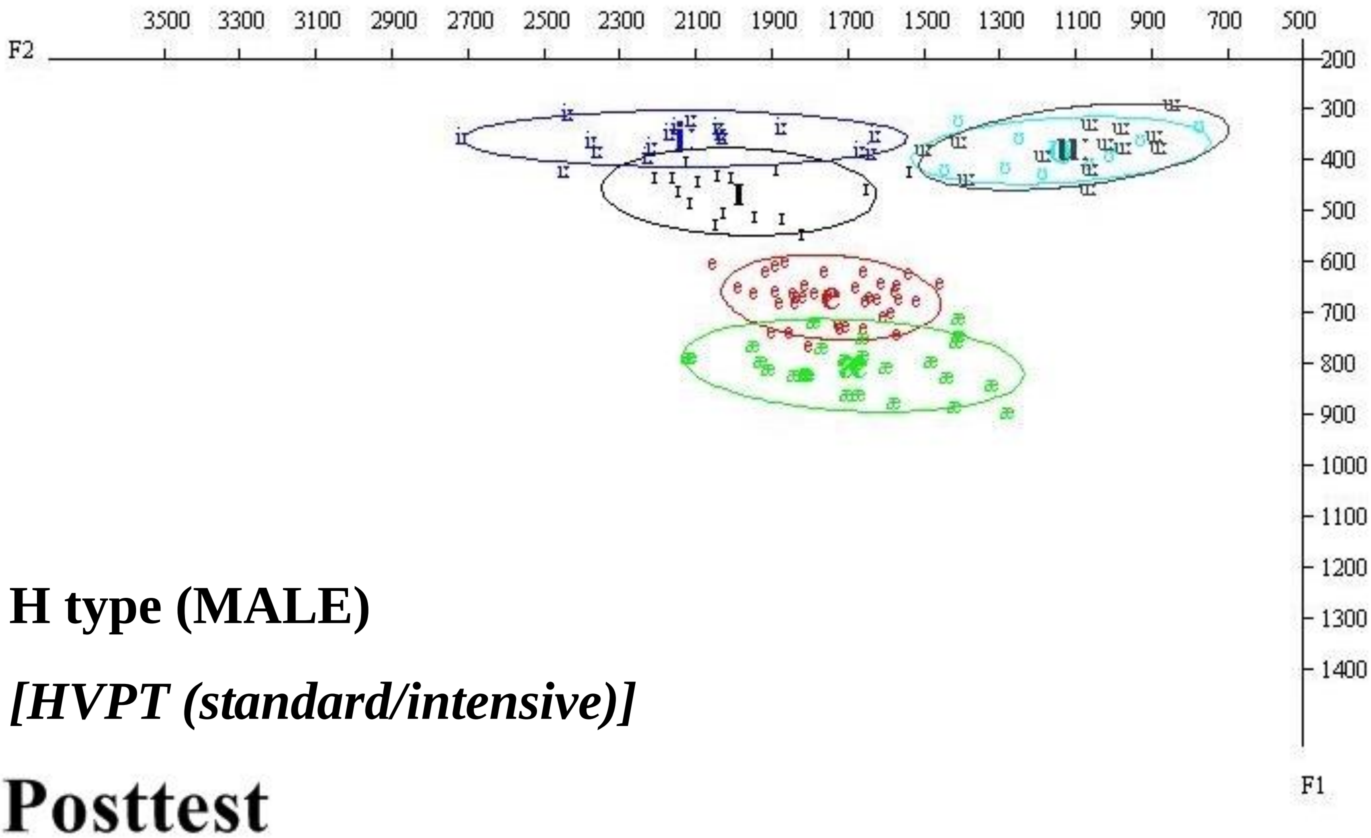


Figure 4.21a F1-F2 space plot of the male subjects in the H type [HVPT (standard/intensive)] in the pretest (above) and the posttest (below).

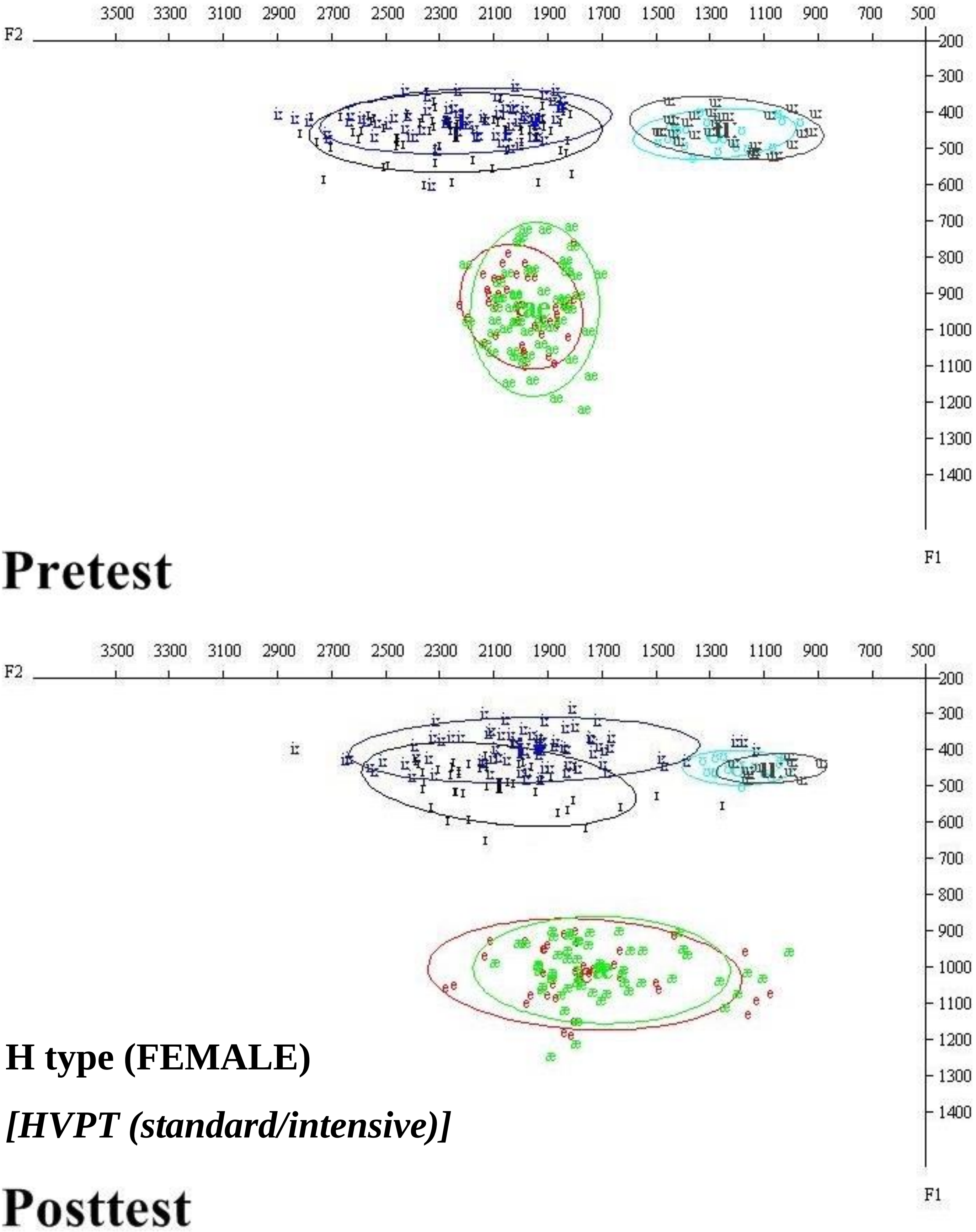
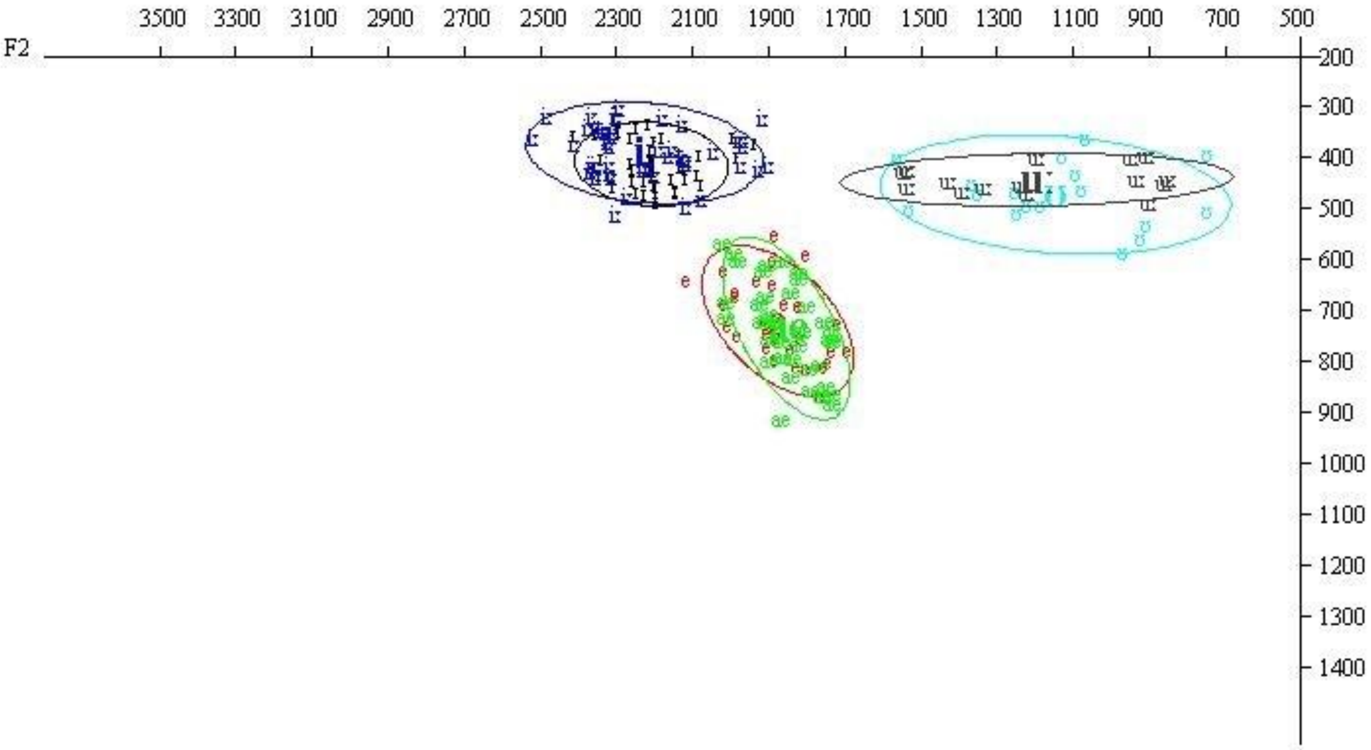
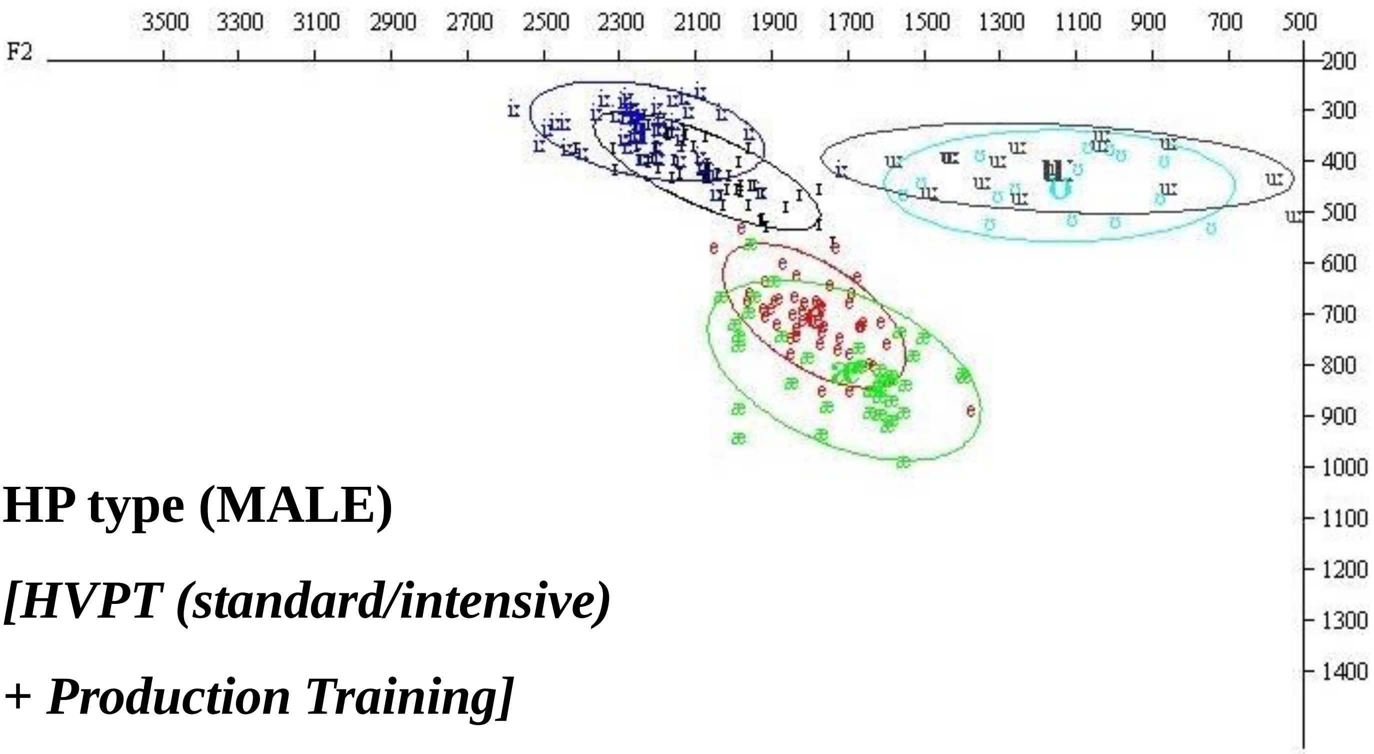


Figure 4.21b F1-F2 space plot of the female subjects in the H type [HVPT (standard/intensive)] in the pretest (above) and the posttest (below).



Pretest



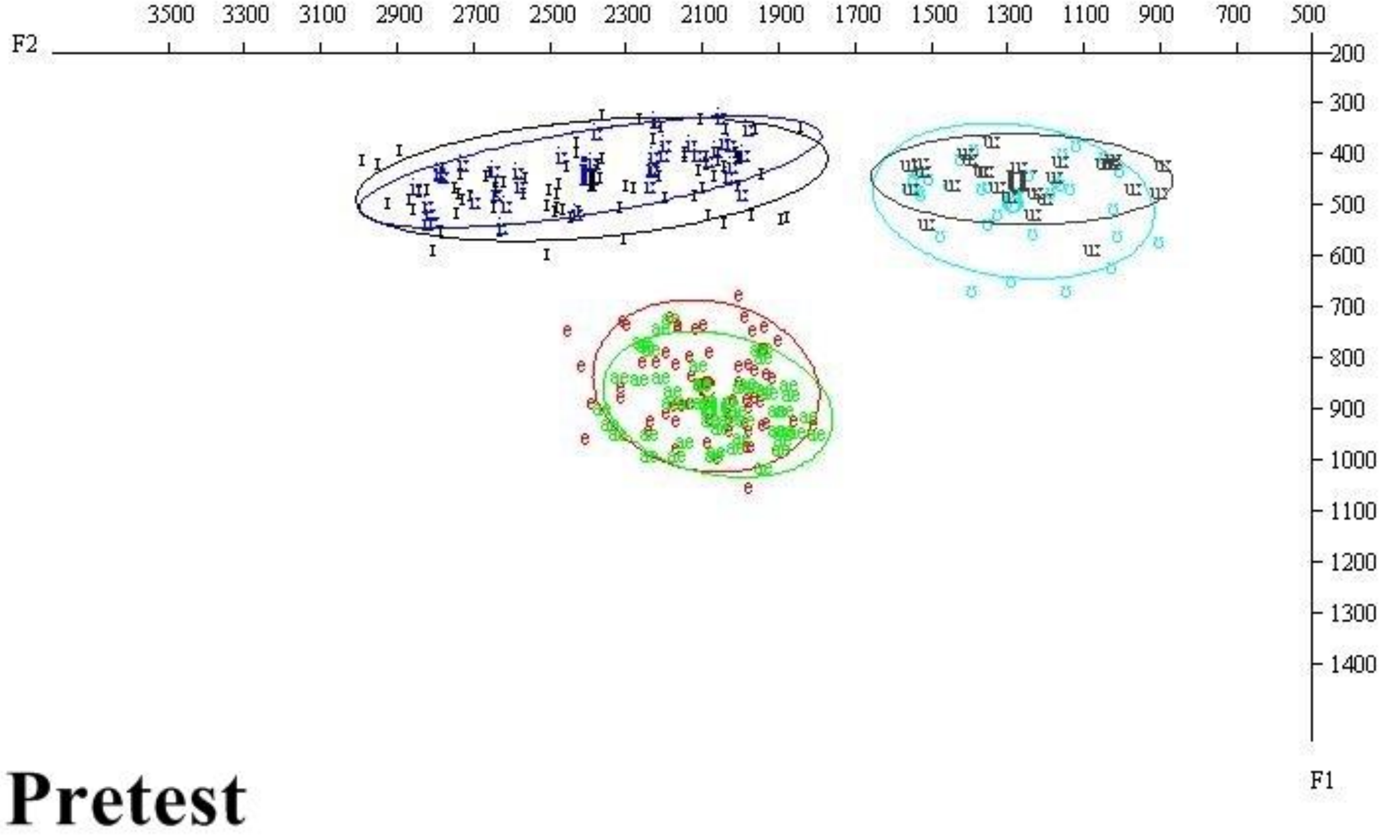
HP type (MALE)

[HVPT (standard/intensive)

+ Production Training]

Posttest

Figure 4.21c F1-F2 space plot of the male subjects in the HP type [HVPT (standard/intensive) + Production Training] in the pretest (above) and the posttest (below).



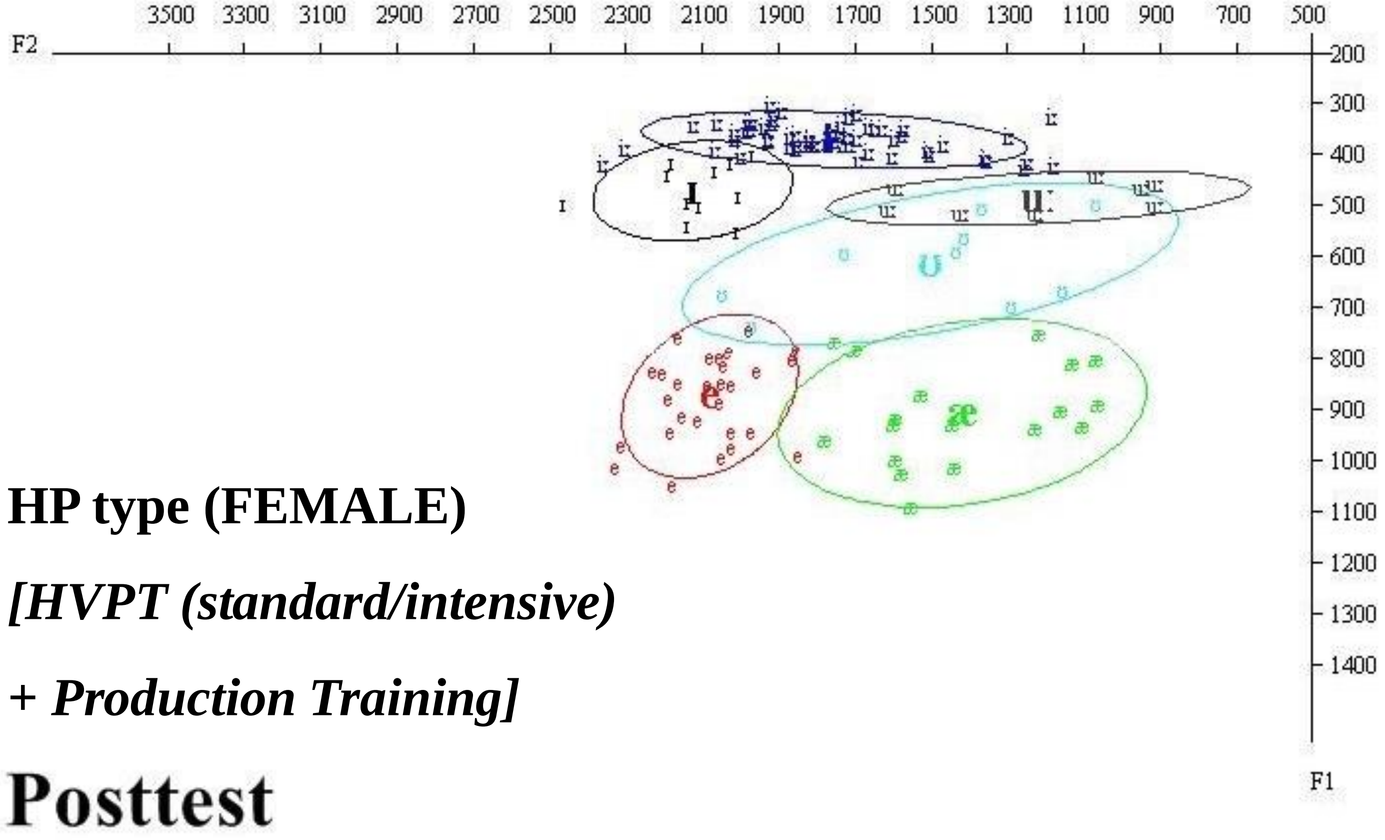


Figure 4.21d F1-F2 space plot of the male subjects in the HP type [HVPT (standard/intensive) + Production Training] in the pretest (above) and the posttest (below).

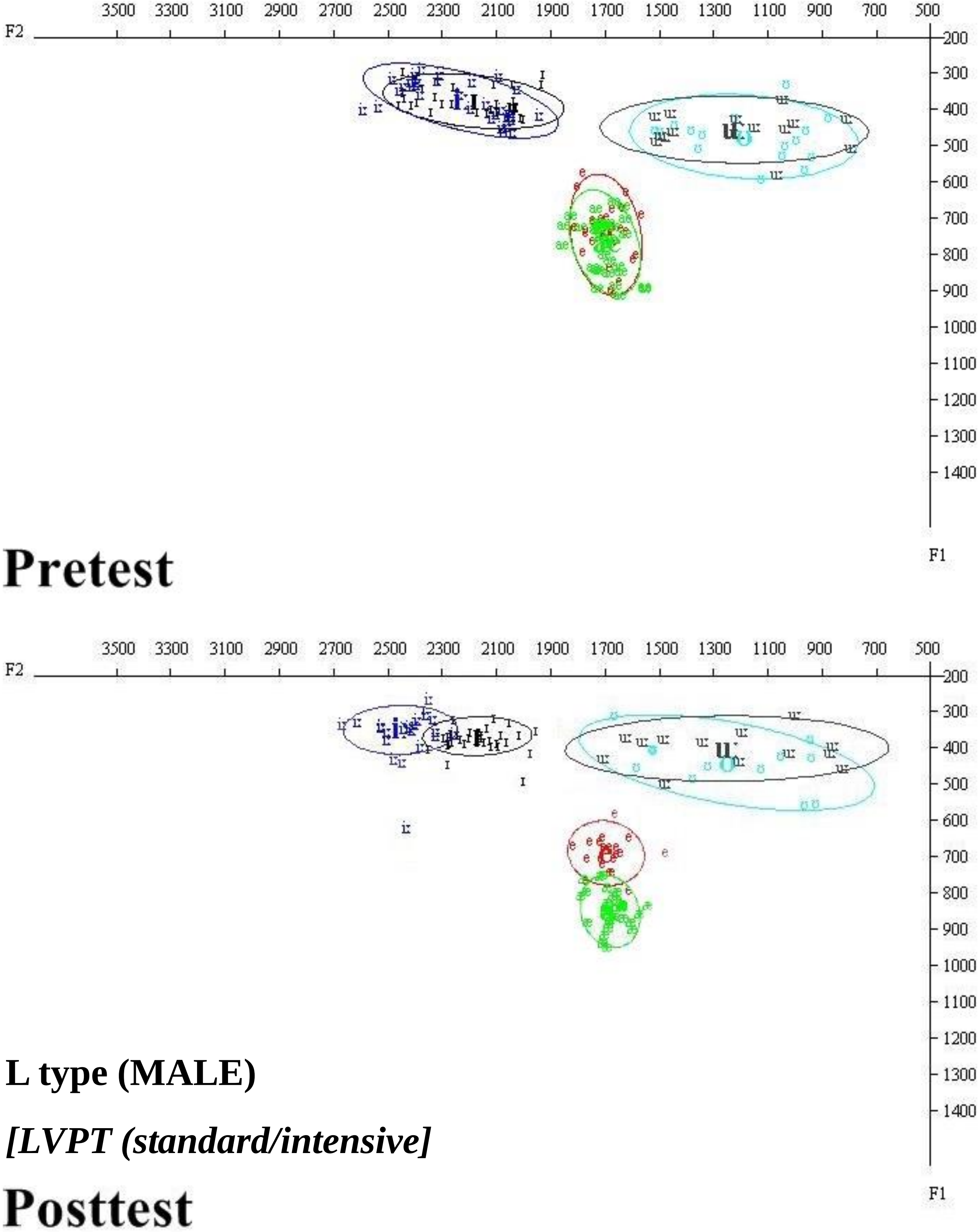


Figure 4.21e F1-F2 space plot of the male subjects in the L type [LVPT (standard/intensive)] in the pretest (above) and the posttest (below).

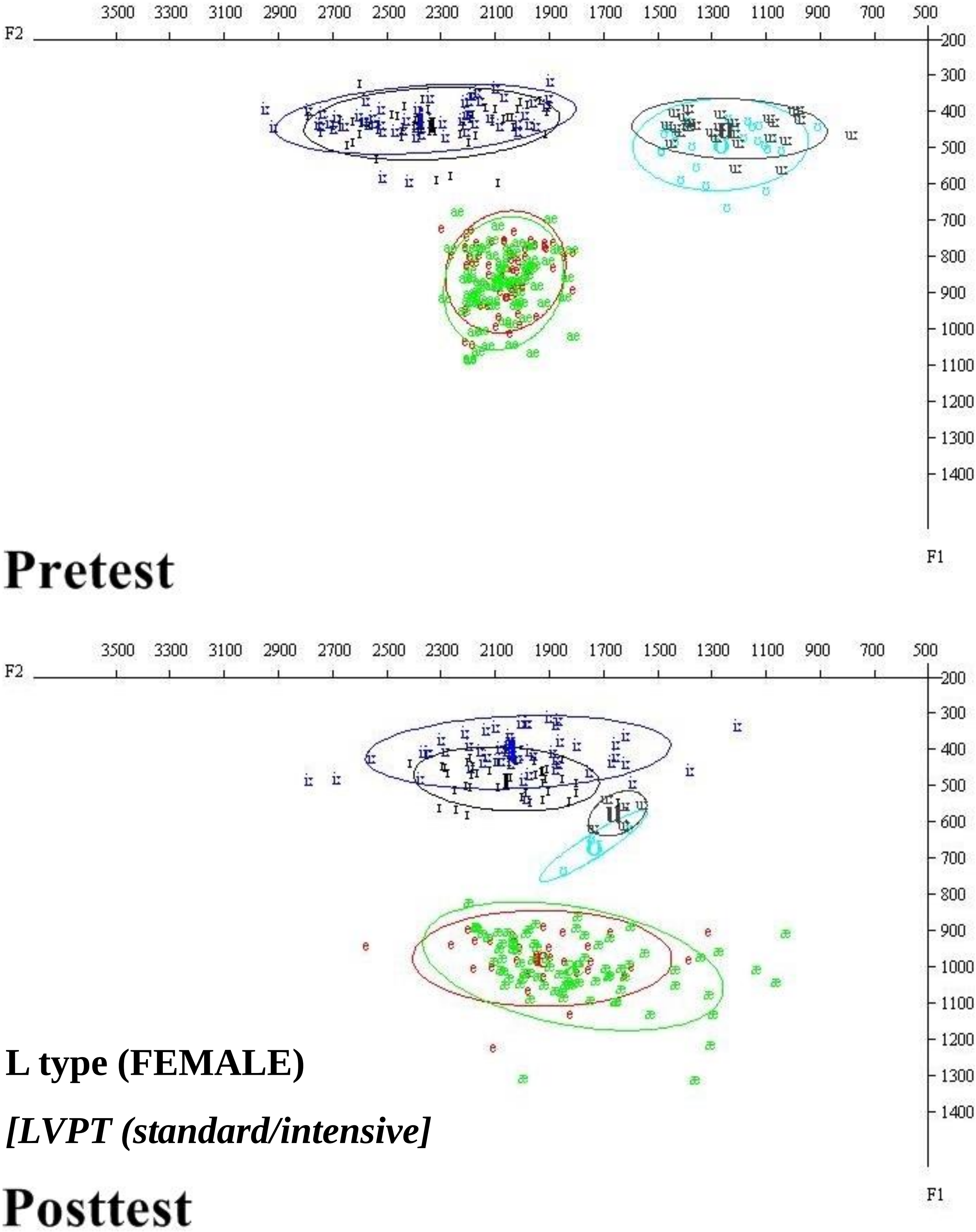
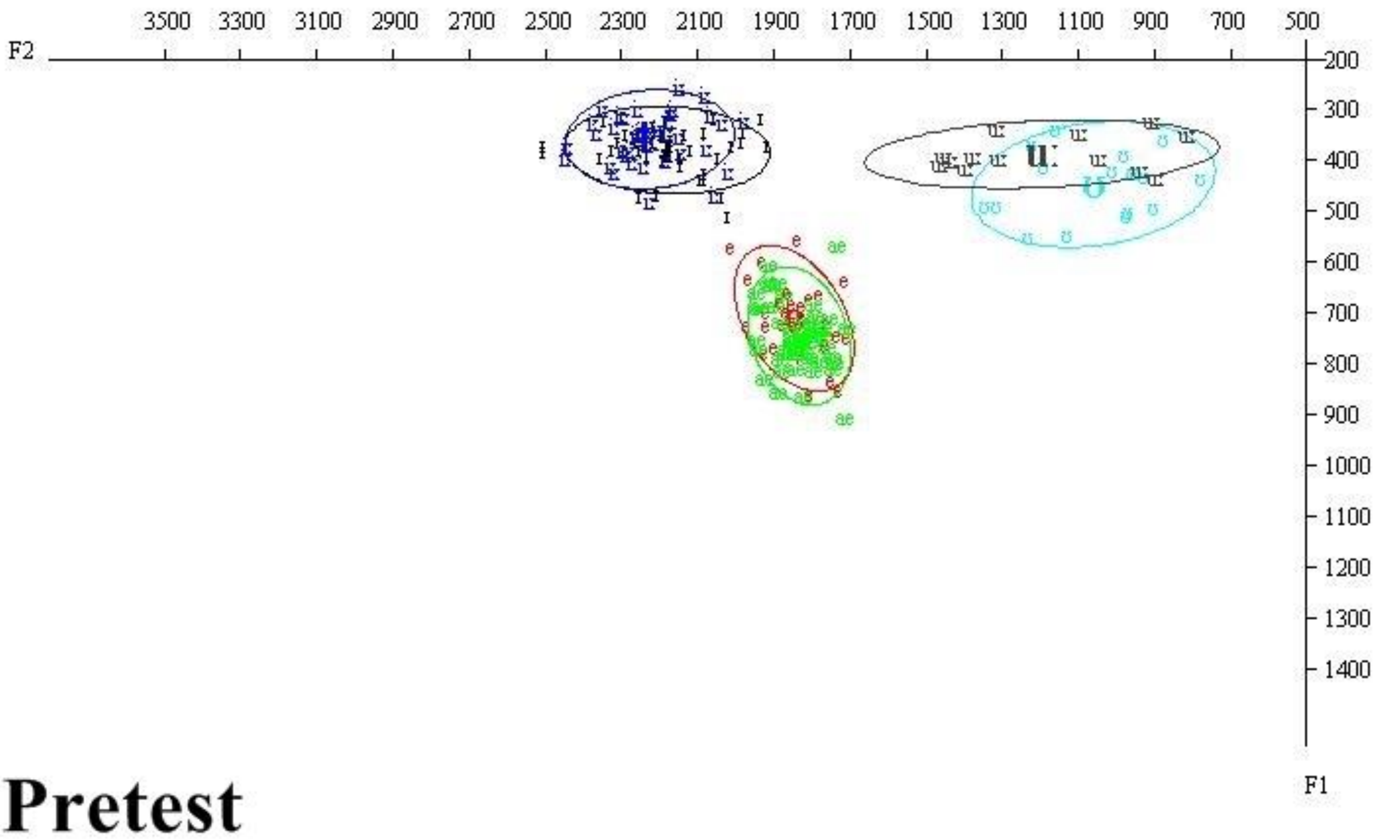


Figure 4.21f F1-F2 space plot of the female subjects in the L type [LVPT (standard/intensive)] in the pretest (above) and the posttest (below).



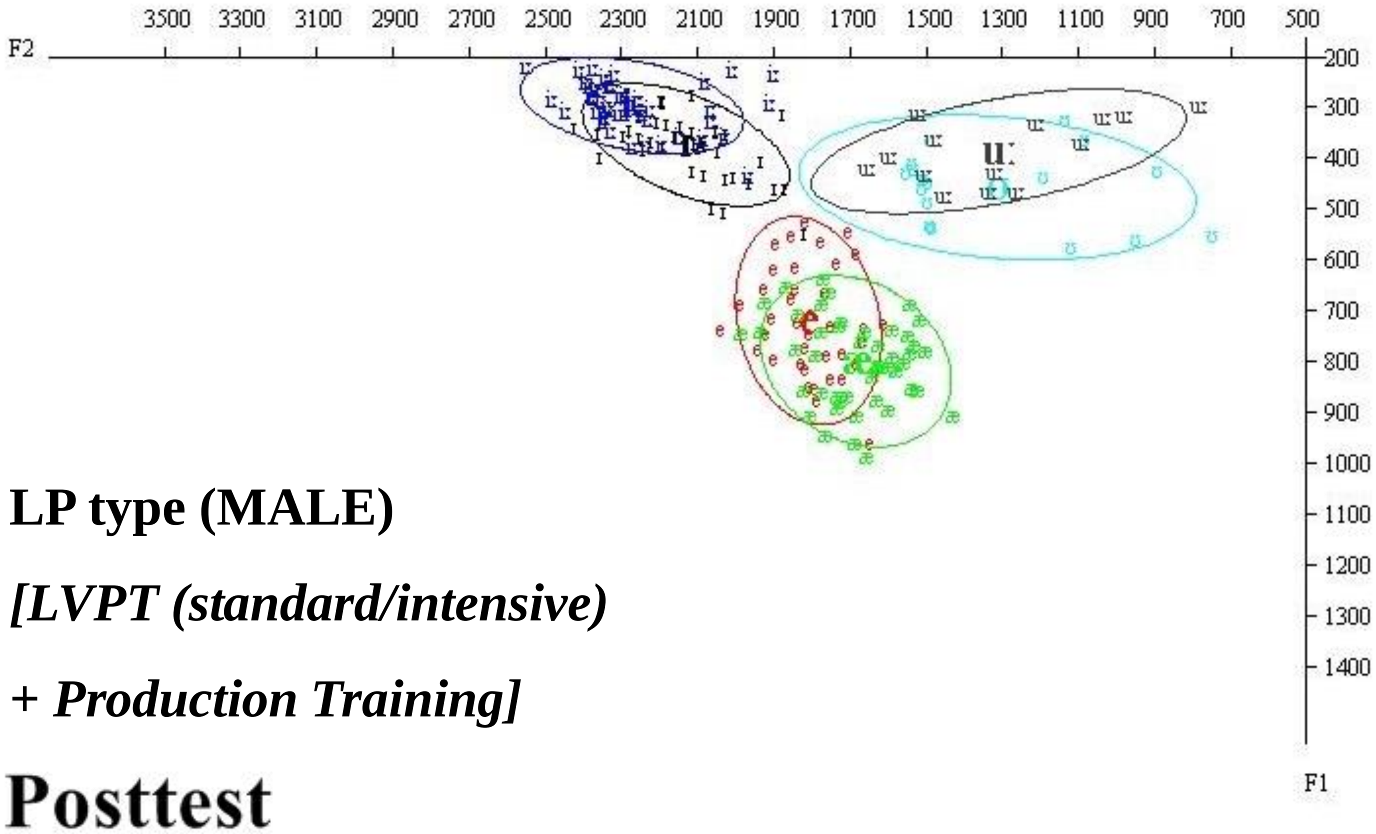


Figure 4.21g F1-F2 space plot of the male subjects in the HP type [LVPT (standard/intensive) + Production Training] in the pretest (above) and the posttest (below).

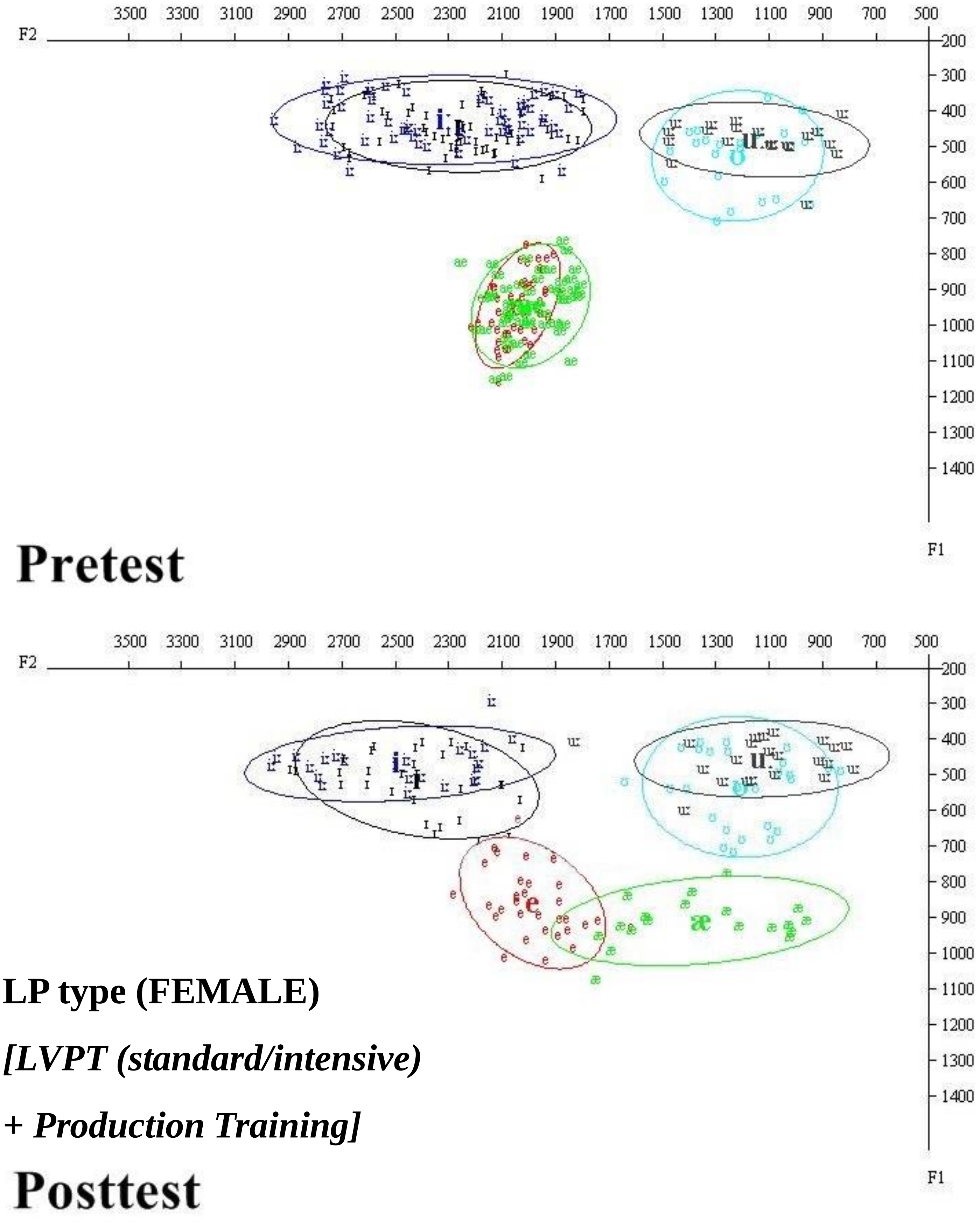


Figure 4.21h F1-F2 space plot of the female subjects in the HP type [LVPT (standard/intensive) + Production Training] in the pretest (above) and the posttest (below).

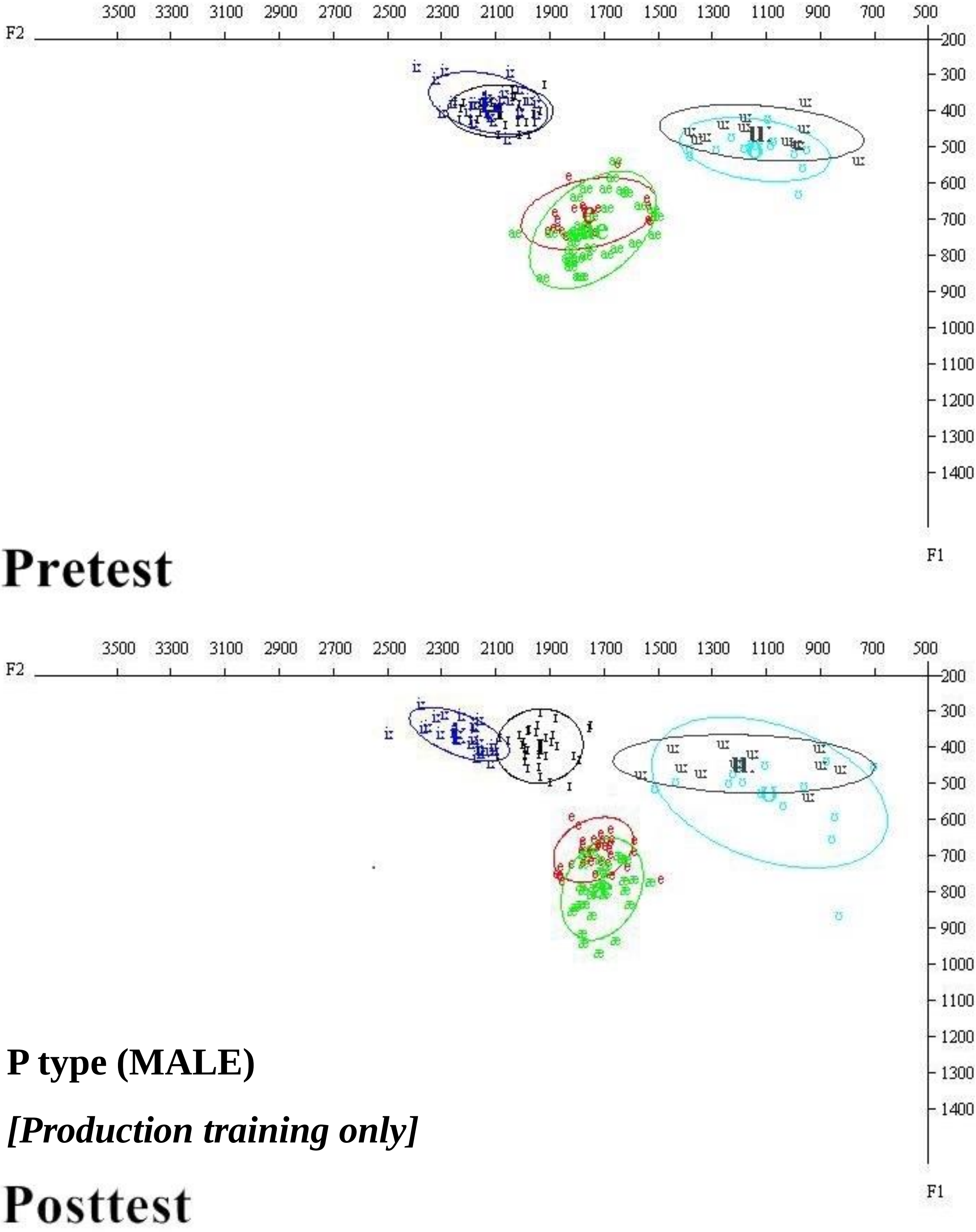


Figure 4.21i F1-F2 space plot of the male subjects in the P type [Production Training only] in the pretest (above) and the posttest (below).

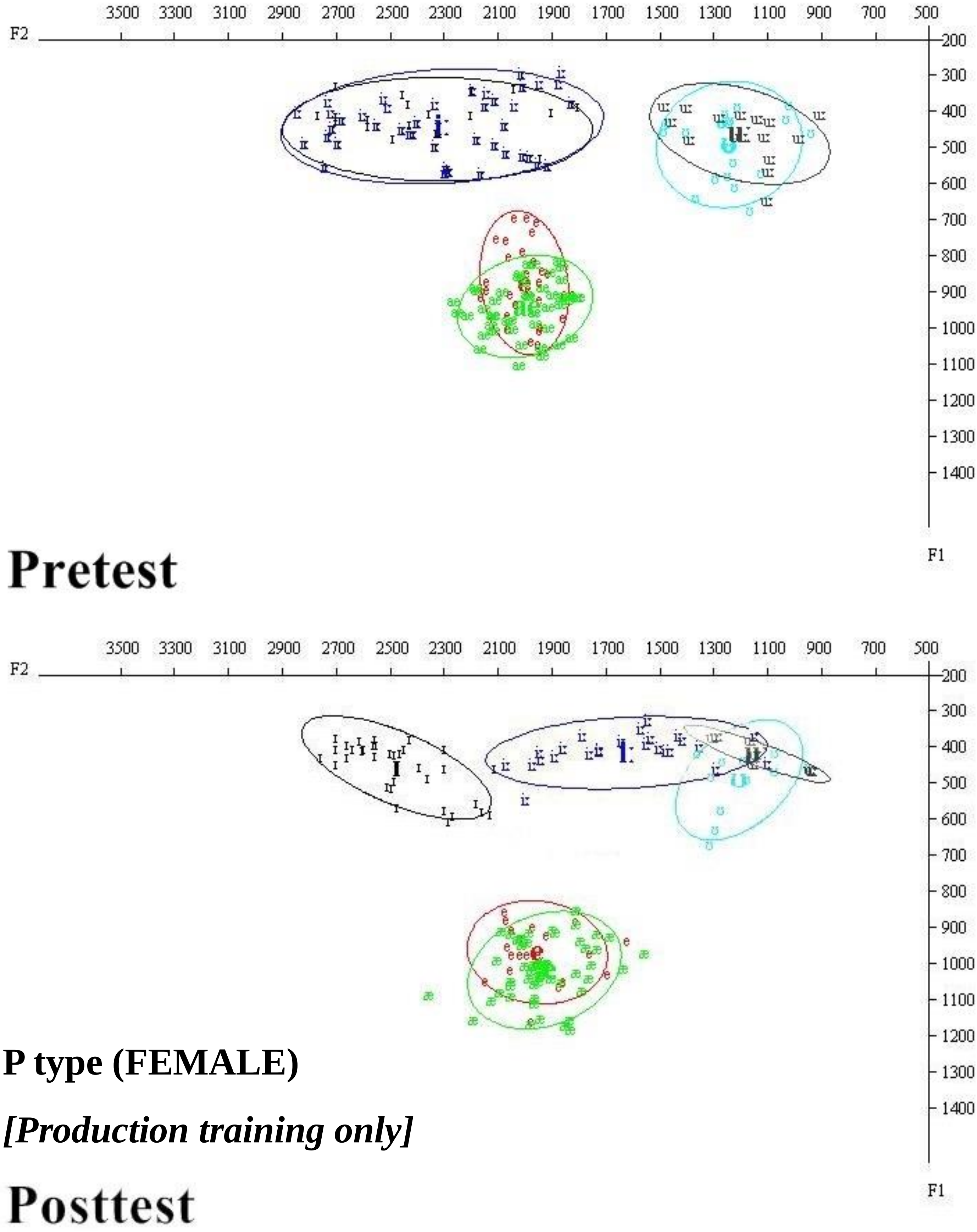


Figure 4.21j F1-F2 space plot of the female subjects in the P type [Production Training only] in the pretest (above) and the posttest (below).

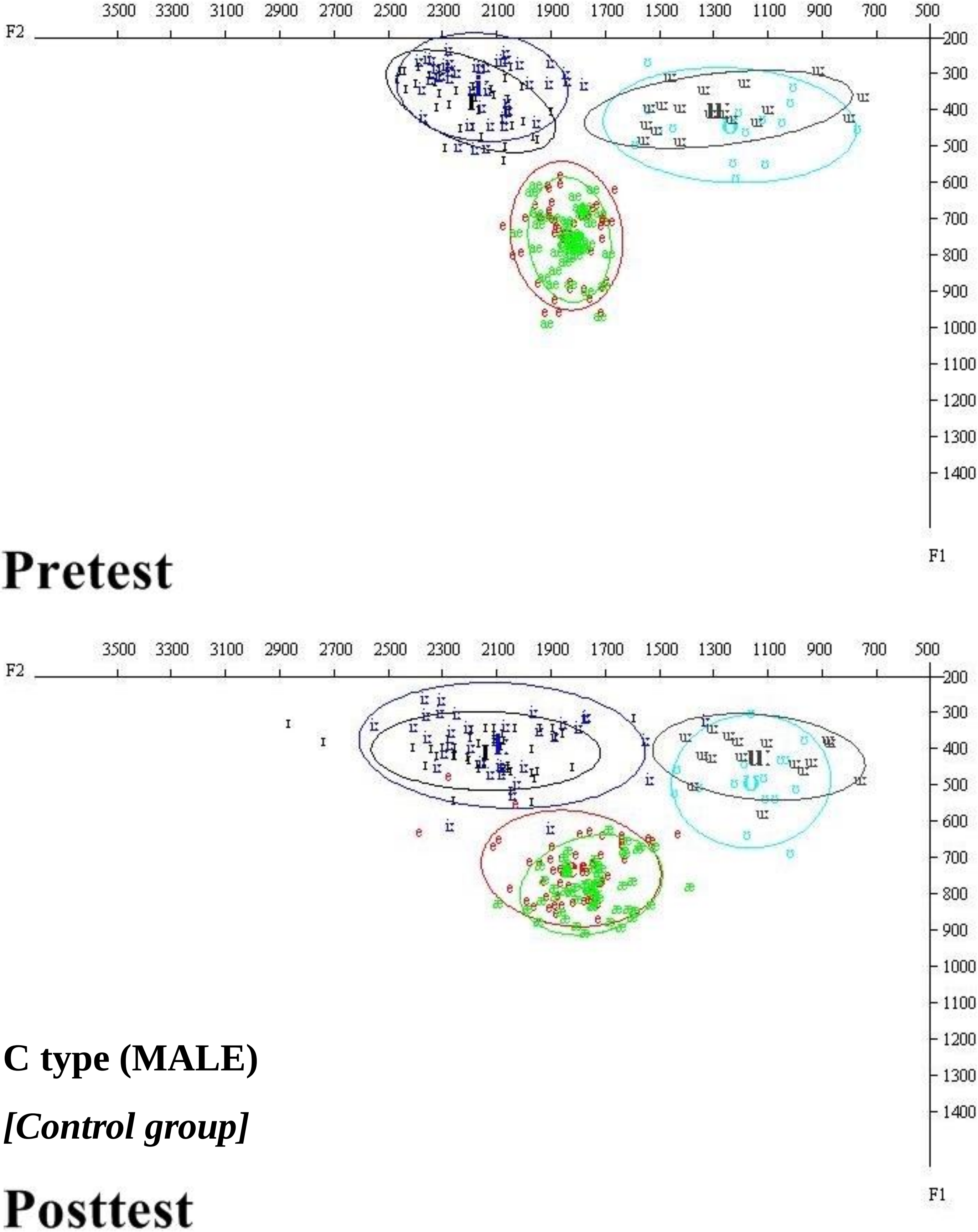


Figure 4.21k F1-F2 space plot of the male subjects in the C type [Control group] in the pretest (above) and the posttest (below).

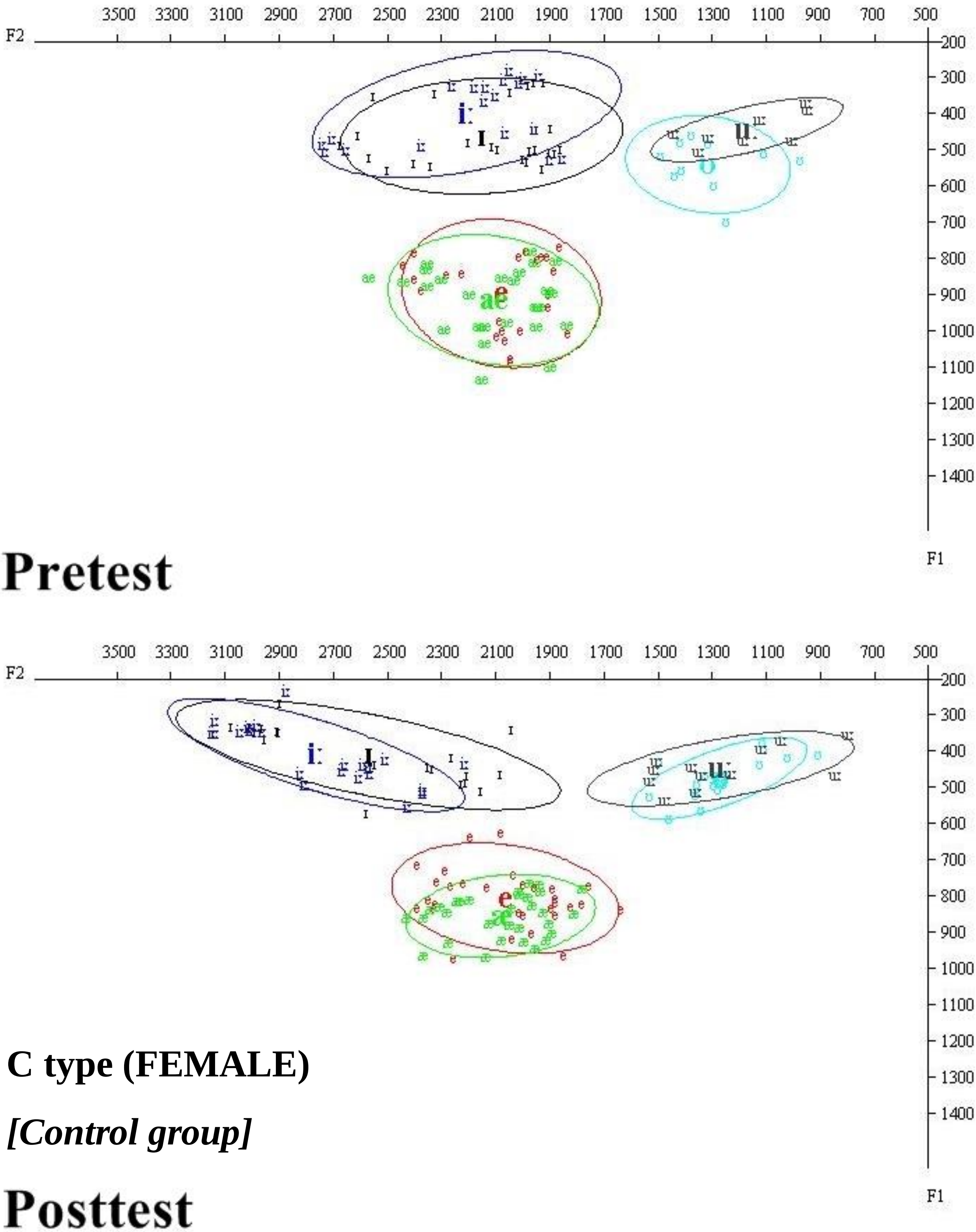


Figure 4.21l F1-F2 space plot of the female subjects in the C type [Control group] in the pretest (above) and the posttest (below).

Cantonese vowels, /i/, /ɛ/ and /u/, to substitute for the three target vowel pairs. The fact that the five trained groups started to show signs of separating the vowel pairs in terms of the F1-F2 clusters or even became able to distinguish two vowels in a pair in terms of vowel quality means that the productions have begun differentiating the English vowels, not just in terms of the duration (as reported in the statistical analysis), but also in terms of the vowel quality, although the degree of which may not be very large. However, it is worth noting that, as pointed out by Hillenbrand and Nearey (1999), spectral information that classifies speech sounds successfully may not mean that human listeners also classify them by relying on the same cues. Since spectral change is not measured in the present experiment, information obtained from acoustic analysis should be taken as supplementary information, though indicative and informative, for the understanding of the production data.

4.2.7 Summing up the Results in Production Learning

In conclusion, statistical analyses of the production performance of the three target vowel pairs showed that groups with training performed better than the control group in the posttest, with the HVPT groups outperforming the LVPT groups. The HVPT groups with an addition of production training also outscored the other groups in the posttest and TC; yet, the same efficacy was not observed in the groups with both the LVPT and the production training. Still, production training alone could benefit the subjects. Nevertheless, no significant differences were observed in the performance of the subjects trained under different perceptual training intensities, implying that this factor may not be influential to the production learning of the subjects. Acoustic analyses have also shown that the production of the three pairs of vowels after training were not as conflated as before, showing signs that the subjects began to separate the vowels when the spectral measurements are considered. The durational differences from the pretest to posttest also revealed that the subjects also took duration as a cue to distinguish the vowel pairs.

4.3 Overall Summary of the Chapter

Chapter 4 has presented both the perceptual and production performance of the subjects, with the data also analyzed by statistical means. Preliminarily, the perceptual training approaches seemed to be playing a weighty role in the improvement of both the perceptual and the production of vowel contrasts, yet training intensity was not an influential factor in the difference in performance. Stimulus variability makes a greater difference when the generalization results of the subjects are concerned, with those given higher variability stimuli being more capable of generalizing their learning. The explicit production training was yet to help perception improvement, suggesting that perception learning appears to be a prerequisite of production; but the fact that the production accuracy rates could rise after the production training may suggest that production can precede perception. Moreover, providing the subjects with an addition of a production training on top of a perceptual training is not necessarily beneficial to the perceptual modification and development, albeit that the production accuracy can be enhanced with particular perceptual training and production training. All these findings imply complications and interactions underlying different factors. The following chapter will discuss and justify the findings in both the perceptual and production tests with a view to answering all the research questions.

CHAPTER 5

DISCUSSION

5.0 INTRODUCTION

On the basis of the findings and analyses presented in the previous chapter, this chapter aims at discussing and elucidating the research findings. The discussions will be arranged in the order of the research questions stated in Chapter 2, with interpretations on the perceptual aspect presented first before the production. Both the effectiveness of HVPT and LVPT on improving the perception and production of the three English vowel pairs, /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ of Hong Kong Cantonese ESL learners, will be evaluated in Section 5.1 at the outset by looking into and comparing the scores and statistical robustness of the results, as a response to both Research Questions 1 and 2. Following this will be Section 5.2, in answer to Research Question 3, which turns the attention to the efficacy of the explicit production training in training both the perception and production of the three target vowel pairs of the subjects. In Section 5.3, the role of perceptual training intensity, as stated in Research Question 4, will be discussed before comparing and contrasting the usefulness of the different training types adopted in the present study in Section 5.4 so that Research Question 5 can be answered. Since the flow of understanding and organization will be strengthened if the effects of learning generalization to both the perception and production aspects – as stated in Research Question 6 – can be incorporated into each of the above sections that evaluate the various different paradigms, a sub-section will be devoted to discussing the effects of generalization under Sections 5.1 to 5.4. These sub-sections will delve into whether the training effects can be generalized to perceiving new words by familiar or new speakers as well as producing the vowel contrasts more accurately at the sentence level. Given this, I will still discuss the learning in generalization and particular observations that are intriguing in Section 5.5, as a supplementary remark and round-up for answering RQ6. All the above sections are concerned about the learning in general, i.e. averaged across the performance of all three target vowel pairs; Section 5.6 will discuss the performance of the subjects across the three target vowel pairs specifically with a view to discovering any learning pattern that is derived from the nature of the vowels and finally answer Research Question 7. Section 5.7 will be an all-inclusive section which offers the theoretical implications of the current study, linking previous research findings and theories with the present discoveries. I will end this chapter by summarizing the findings in Section 5.8, before stating the limitations and contributions of this study in next chapter.

5.1 EVALUATION OF PERCEPTUAL TRAINING APPROACHES – RESEARCH QUESTIONS 1 & 2 (&6)

The present study evaluates the effectiveness of several training approaches in the modification of perception and production of the three target English vowel contrasts among our target groups. The two perceptual training approaches, High Variability Phonetic Training (HVPT) and Low Variability Phonetic Training (LVPT) will be compared first. The subjects' performance in the two training approaches will also be compared to those in the control group. Besides the data in the pretests and the posttests, the results in Tests of Generalization will also be discussed to give a more complete picture of the subjects' performance. For convenience, the two main research questions, and the one about the effects of generalization, are repeated below.

RQ1. Are the two perceptually-based phonetic training approaches, HVPT and LVPT, effective in improving
a) the perception and
b) the production of the English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/?

RQ2. Which training approach, HVPT or LVPT, is more effective in improving
a) the perception and
b) the production of vowels /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/, and how different are they?

RQ6. Can any learning effect be generalized to
a) the perception of new words produced by both familiar and new speakers, or
b) to the production in a more naturalistic environment?

5.1.1 PERCEPTUAL DOMAIN

5.1.1.1 PRETEST VS. POSTTEST

Promising results were obtained when comparing the identification scores from the pretest to the posttest of the subjects trained under HVPT, LVPT and the control group. Both HVPT and LVPT improved the identification of the vowel pairs (averaged across the three target pairs) significantly from the pretest to the posttest, but the control group did not. The improvement observed for the HVPT subjects was 21.95% (18.50% for /ɪ/-/iː/; 30.43% for /e/-/æ/; and 16.93% for /ʊ/-/uː/) whereas that of the LVPT subjects was 23.79% (16.30% for /ɪ/-/iː/; 31.81% for /e/-/æ/; and 23.28% for /ʊ/-/uː/), which is slightly higher than that obtained by the HVPT subjects. The control group had insignificant differences. Further statistical analysis showed that the two trained groups did significantly better than the control group; yet, no significant differences were found between HVPT and LVPT, meaning that the relative benefit brought by the differences in the variability of the stimuli cannot be determined at this stage. Still, all these results evidenced that both HVPT and LVPT were beneficial in improving the subjects' ability to perceive the difficult vowel contrasts and the elucidations are presented in the sections under 5.1.1.1.1, whereas the explanation of no significant differences between HVPT and LVPT will be offered in 5.1.1.1.2. The following table serves as a summary for a better recapitalization of the results obtained in perceptual pretest and posttest in HVPT and LVPT groups:

Table 5.1

Summary of the Effects of Perceptual Training Approaches in the Perceptual Pretest vs. Posttest

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Perceptual Training Type (HVPT vs. LVPT vs. Control)	✓	HVPT > Control LVPT > Control
2. Test (Pretest vs. Posttest)	✓	Posttest > Pretest
3. Vowel (front vs. low vs. back)	✓	front > low back > low
4. Test × Perceptual Training Type	✓	<u>Pretest:</u> HVPT vs. LVPT vs. Control (n.s.) <u>Posttest:</u> HVPT > Control LVPT > Control ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ <u>HVPT:</u> Posttest > Pretest <u>LVPT:</u> Posttest > Pretest <u>Control:</u> Posttest vs. Pretest (n.s.)
5. Perceptual Training Type ×	✓	<u>HVPT:</u>

Vowel		front vs. low vs. back (n.s.) LVPT: front > low front > back Control: front > low back > low
6. Test × Vowel	✓	Pretest: front > low front > back back > low Posttest: front vs. low vs. back (n.s.) ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ Front vowel pair: Posttest > Pretest Low vowel pair: Posttest > Pretest Back vowel pair: Posttest > Pretest
7. Perceptual Training Type × Test × Vowel	✓	Front vowel pair: HVPT: Posttest > Pretest LVPT: Posttest > Pretest Control : Posttest vs. Pretest (n.s.) Low vowel pair: HVPT: Posttest > Pretest LVPT: Posttest > Pretest Control : Posttest vs. Pretest (n.s.) Back vowel pair: HVPT: Posttest > Pretest LVPT: Posttest > Pretest Control : Posttest vs. Pretest (n.s.)

5.1.1.1.1 GENERAL SUCCESS OF THE HVPT AND THE LVPT GROUPS IN THE PERCEPTUAL LEARNING

Chapter 2 has delineated that both HVPT (e.g. Bradlow et al., 1997; Hardison, 2003; Wang, 2002) and LVPT (e.g. McCandliss et al., 2002; Strange & Dittmann, 1984) have been shown in previous studies as useful paradigms in modifying the phonological categories of some non-native contrasts, though they differ in the degree from which the subjects can benefit. The results in the present study, in general, parallel these previous findings and it is speculated that the success that the two paradigms can bring to the subjects is attributed to the following reasons.

Firstly, a consistent adoption of the same task has caused the effectiveness in the two approaches. The overall design of HVPT and LVPT adopted in this research replicated the previous studies, with only modifications in the number of choices in the tests, i.e. a four-alternative forced-choice (4AFC) identification test was adopted in the present work. Even though in training, a two-alternative forced-choice (2AFC) identification test was used, the two still remained as the same type of listening tasks which helps promote the re-categorization of more near native-like phonological patterns. As reported in some related studies in this aspect (e.g. Lambacher et al., 2005; Shiffrin & Schneider, 1977), maintaining a consistent mapping of the phonological categorization through referencing between the training stimuli and the responses in the same type of listening tasks throughout the testing and training phrases has been shown to promote the general success in both training methods.

The second explanation for the general success of both HVPT and LVPT paradigms may be attributed to the nature of the training tasks used. Directing the subjects' attention to the subtle differences between the minimal pairs by using isolated words in both paradigms may have allowed the subjects to be exposed to the critical attributes of the stimuli that distinguish one sound from another. Since the 2AFC paradigm was used in the training tasks in which only two choices were given, the subjects' attention and consciousness to the target sound segments would be held, making them easier to concentrate on the contrast at hand. I have also blocked the training stimuli under vowels such that the subjects were only exposed to one single target vowel pair in one block before moving to a new block. It is speculated that under such a training condition in which the attentional resources were available, the subjects could detect the differences of the vowels in a more focused manner, leading to successful perception of the target vowel pairs.

The sole adoption of identification tasks in lieu of discrimination tasks in both HVPT and LVPT approaches may also contribute to the general success of the two trained groups. Identification tasks aim to focus on the subjects' capability to map an appropriate linguistic label to a phone as all identifications are absolute as they are made based on each separate stimulus' characteristics, but not on eliminating other stimuli that were not their targets (Beddor & Gottfried, 1995; Logan & Pruitt, 1995). As shown in some previous studies (e.g. Flege, 1995a; Jamieson & Morosan, 1986), training using identification tasks may be more beneficial to the perceptual learning of the subjects, since the subjects were only required to focus on only one stimulus at a time which then helps the subjects tap into memory for the specific perceptual category. This can draw their attention to that specific segment and allow more focused incorporation of within-category variability and the assignment of a particular linguistic label to a segment. In other words, it forces the subjects to form a more native-like phonetic category by supplementing the learning (Iverson et al., 2011; Jamieson & Morosan, 1986; Lambacher et al., 2005; Logan & Pruitt, 1995). Also, it has been reported that subjects' motivation in doing identification tasks was higher than that of discrimination, which led to more successful perceptual learning (Flege, 1995a). The present research does not aim at re-investigating the effectiveness of an identification task versus a discrimination task, but the success of the present training approaches may be partly ascribed to the utilization of an identification task.

Two types of feedback were adopted in the present study to promote the perceptual learning of the subjects, and hence the effectiveness of HVPT and LVPT in general. They are session-by-session cumulative feedback and trial-by-trial immediate feedback. A cumulative score after each training session was given to the subjects. This kind of feedback which spans through a session promotes the motivation of the subjects to continue the training. Although it is a more indirect way of helping the subjects to improve their perception of sounds, it has already provided them with a general picture of their on-going performance and progress in each session to enhance their motivation (Logan & Pruitt, 1995). Administering immediate feedback after each answer submission during training also allows the subjects to learn the correct segment immediately when they were still focused on the sound segment of a particular word. It is a more direct way of learning than the session-by-session feedback. Offering trial-by-trial immediate feedback in the experiment was meant to accelerate the perceptual learning process and the general success of the two trained groups can be attributed to this (Lambacher et al., 2005; Logan & Pruitt, 1995). It is also a crucial element for a successful phonetic training approach since the use of immediate feedback facilitates the learning process by enabling the subjects to substantially concentrate on the salient acoustic features and cues of the target segments (Logan et al., 1991; Pisoni, 1977). Through trial and error, together with immediate feedback, the subjects will not leave their attention to the sound segments unattended but are consistently offered a chance to uptake the feedback. Although the feedback was designed by simply showing "correct" or "incorrect" to the subjects, this has already provided the subjects with adequate chances to learn the correct answer by listening to the stimulus again, hence augmenting the positive effects of the learning.

5.1.1.1.2 No Significant Differences between HVPT and LVPT

It is logical to believe that after any perceptual training, the subjects' perception performance could improve to a certain extent. Yet, one interesting observation in the present experiment that is worth more discussion goes to the situation that there was no significant differences between the performance of the subjects trained under HVPT and LVPT, and that the LVPT groups were even demonstrating slightly better learning than the HVPT groups. Although previous studies seldom compared the two paradigms, the general belief that high stimulus variability (as utilized in HVPT) promotes robust perception learning appears to be not supported by the present findings when only the pretest and posttest scores are concerned. This is also not in line with my previous studies (Wong, 2010; 2012) which showed that the subjects trained under HVPT outperformed those trained under LVPT.

One possible explanation for the null effect of training difference is related to the number of training sessions provided to the subjects. A training study which investigated the role of attention in learning of phonetic categories conducted by Guion and Pederson (2007) intentionally used relatively few (a total of two) and short (120 tokens per a self-paced session which lasted for 30-40 minutes) training sessions because they believed that longer periods of training would change the size of the effect.

Thomson (2007), who compared the training effects of three vowel training methods with the target vowel being manipulated in different ways, also found no effect of phonetic training conditions. He ascribed this finding also to the length of training period (a total of eight sessions with 200 items per session). It is possible that the subjects in both HVPT and LVPT in the present investigation had overcome the difficulties that they initially encountered during the identification when they were given such a high magnitude of training (20 sessions with 40 tokens per vowel pair in each session, each lasted for 15-20 minutes). The differences between the two training paradigms might have been detected if it were the early stages because with more training, all the participants became successful in discerning the difference between the vowel contrasts, regardless of stimulus variability. My previous works (Wong, ibid.) only used 10 sessions (60 tokens per session) and together with a comparison to Thomson (2007), a total amount of around 600 tokens in less than 10 sessions might perhaps be the minimal amount of training that can leads to successful training. This claim certainly demands further research.

Another explanation is about the use of talker in the pretest and the posttest stimuli. The stimuli used in the two tests were produced by the same talker who contributed all the training tokens in LVPT. However, this talker only produced one-sixth of the tokens in HVPT since a total of six speakers were recruited for making the HVPT stimuli. When the subjects trained under LVPT were exposed to 800 instances per vowel produced by the same talker, the talker-familiarity effect became more pervasive. Besides the classically expected positive results (e.g. Lambacher et al., 2005) demonstrated by the HVPT subjects due to the learning-facilitating nature of stimulus variability (this will be discussed in Section 5.1.3 when a totally different result was observed in the generalization tests), previous perceptual training studies (e.g. Mullennix et al., 1989; Pisoni & Lively, 1995) discovered that a greater degree of trial-by-trial variability in tokens, i.e. HVPT, was a more difficult task than identifying words produced by a single talker, i.e. LVPT. Talker-specific information together with important acoustic cues which differentiate the vowel contrasts might become more easily acquired by the subjects trained under LVPT than HVPT. The LVPT subjects became more familiarized with the singular voice they heard throughout the tests and training. Given also that the learning observed among the LVPT group was limited to the trained sets but not to novel stimuli (see Section 5.1.1.2), it is speculated that the improvement from the pretest to the posttest may not present genuine perceptual learning, rather, it is only a shift in strategy to attend to task relevant cues, e.g. talker's characteristics, and resulted in similar performance with HVPT. Due to these reasons, it is thus reasonable to find that the differences between the subjects trained under HVPT and LVPT paradigms are not robust, because the degree of learning promoted by high stimulus variability might be more comparable to the familiarity of a single voice attained by the LVPT subjects.

Jongman and Wade (2007) discovered that prototypical examples with low variability, rather than high, may facilitate learning at the initial stage; whereas a recent training study by Perrachione et al. (2011) found out that some low-aptitude learners would face more difficulties in HVPT than LVPT, but their posttest were evaluated by using stimuli produced by a new voice, i.e. same as the Test of Generalization 2 (TG2) adopted in the present study. In fact, a different picture was also shown in the results in TGs in the present study and this will be discussed in detail in the next section. Comparing the groups' performance in TGs may allow an understanding of more genuine learning.

5.1.1.2 Effects of Perceptual Generalization: HVPT over LVPT

In this study, one measure and indicator of the effect of learning was accomplished by comparing the identification scores, i.e. percentage increase of accurate identification from the pretest to the posttest, obtained by subjects trained under HVPT and LVPT in the pretest and the posttest. However, in real life, learners are exposed to new voices and words in their ambient language. Whether the perceptual learning can be transferred to novel stimuli produced by novel speakers may shed some light on a more genuine learning experience by the subjects. This section will discuss the learning of the subjects from the pretest to the Tests of Generalization and the difference in subjects' learning under the two paradigms.

The difference scores obtained by subtracting the pretest scores from TG scores in each group showed that HVPT was a more effective approach than LVPT. The generalization effects were only observed in HVPT. Different from the null effects of training types observed by comparing simply the results in the pretest and the posttest, the high degree of success implied by the difference scores of TGs showed that perceptual stimuli produced by multiple talkers were more useful in modifying the perceptual identification of the subjects. This finding is in line with earlier studies (Bradlow et al., 1997, Bradlow et al., 1999; Wang, 2002; Wong, 2010, 2012, 2013a, b; Yamada, 1993) demonstrating the importance of adopting stimuli with highly variable contexts produced by multiple speakers, i.e. HVPT, to excel training in non-native contrasts. LVPT appears to be an easier training platform for the subjects to learn the identification, probably due to the talker effect, but it has been found to be ineffective in the modification of listeners' phonetic perception pattern (Strange & Dittmann, 1986). The results in the present study also confirmed previous findings.

In the present study, the difference scores (obtained by subtracting pretest scores from TG scores) obtained in the three TGs by HVPT (TG1: 15.25%; TG2: 15.25%; TG3: 15.83%) were all in general higher than those of LVPT (TG1: 3.5%; TG2: 9.89%; TG3: 11.06%), hinting that exposure to multiple speakers and phonetic contexts during training enhanced perceptual learning and the effect is externally valid. Before proceeding to give possible reasons for the success in using stimuli produced by various speakers and contexts to generally improve the perception of the three vowel pairs, a summary of the results obtained in the three TGs is given for a review:

Table 5.2

Summary of the Effects of Perceptual Training Approaches in the TGs

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Perceptual Training Type (HVPT vs. LVPT vs. Control)	✓	HVPT > Control HVPT > LVPT LVPT > Control
2. TG (TG1 vs. TG2 vs. TG3)	✓	TG2 > TG1 TG3 > TG1
3. Vowel (front vs. low vs. back)	✓	low > front
4. TG × Perceptual Training Type	×	
5. Perceptual Training Type × Vowel	×	
6. TG × Vowel	×	

7. Perceptual Training Type × TG × Vowel	×	

The first explanation for HVPT being superior to LVPT is due to the use of highly variable stimuli in HVPT which promoted the subjects' selective attention to specific acoustic cues. One of the main goals to use highly variable stimuli in the present study was to avoid idiosyncrasies brought by stimuli produced just by one single speaker, since any discrepancies present may deter the realization of the acoustic cues necessary for identifying the three vowel pairs. Replicating previous studies' findings (e.g. Lengeris, 2009; Lively et al., 1993; Logan et al., 1991; Wong, 2010, 2012), the positive results observed from the transfer in generalization with HVPT being a more effective approach than LVPT gave evidence that stimuli in HVPT are more advantageous over those in LVPT. This finding is intriguing in the way that the genuine learning through different paradigms in perception appears to show a difference only when it came to generalization; the comparison between the pretest and posttest scores showed simply no robust differences between HVPT and LVPT, which has been justified in the previous section.

It is suggested that the use of highly variable stimuli in HVPT promotes selective attention to the criterial acoustic cues which are to be observed by the subjects from the large pool of stimuli with wider variability. As reported in some earlier studies (e.g. Jusczyk, 1989; Logan et al., 1991; Nosofsky, 1986, 1987; Strange, 1986), the role of selective attention is vital in perceptual learning and the development of phonetic categories, no matter whether the subjects are infants or adult learners. The possibility of altering of the use of acoustic cues is also consistent with some previous findings

(Francis, Baldwin & Nusbaum, 2000; Francis & Nusbaum, 2002; Francis, Nusbaum & Fenn, 2007; Holt & Lotto, 2006). Exposure to highly variable stimuli is necessary for the subjects to form robust phonetic representations by learning which acoustic cues are relevant to a specific sound and to develop more language-specific phonetic categories. The subjects had to strive to ignore the variations introduced by the use of different voices and phonetic environments, and then selectively concentrate on the dimensions important for distinguishing the categories within the given wide range of acoustic dimensions and weightings. Once the L2 categories have started to strengthen and become more stable, the learners begin to be able to benefit from the stimuli with higher variability. It is also deemed that highly variable stimuli encourage the subjects to use a more abstract and higher-level of phonological encoding in the short-term memory (e.g. Flege, 2003; Højen & Flege, 2006; Iverson et al., 2011). HVPT has thus been thought to generalize better than LVPT, which is supported by the positive results given in the present experiment.

Based on the above accounts, if highly variable stimuli promote better perceptual learning because the attentional weights for different phonetic dimensions have changed, then it is also speculated that it gives evidence to the exemplar-based approach to speech perception. Exemplar-based models of speech perception have recently been applied to speech perception and processing (e.g. Goldinger, 1990, 1996, 1997; Goldinger et al. 1991; Johnson, 1997; Pierrehumbert, 2001; Pisoni, 1990, 1992, 1997; Pisoni et al., 1985). The core idea underlying the exemplar-based models is that mental representations consist of memory traces of specific tokens. A perceptual category is hence defined as a collection of all experienced instances of the category, i.e. detailed information of any speech signal which entails phonetic and linguistic information will become part of the stored and labeled tokens, or *exemplars*. Each category is represented in the memory by a space of remembered exemplars and the exemplar space is continuously updated upon new speech events. Therefore, exposing the subjects to a wider range of natural stimuli which are more variable in terms of speaker and context variability would be beneficial to enrich their experience, hence forming exemplars throughout the training. The exemplar perceptual space of the subjects will be stretched wider along dimensions where two vowels in a contrast differ and are shrunk along dimensions that do not show distinctions for the two vowels. It follows that the subjects trained under HVPT can hence draw on a wider exemplar space than the LVPT group when they listen to novel words and voices. The results of the present experiment may suggest that this change has occurred.

However, even with all the above accounts, cautions have to be paid as there is still no solid evidence from this experiment that the reweighting of cues occurred which consequently brought about corresponding change in the best exemplars space because this research had not directly measured the change of cue weightings. The suggestions given were only inferred based on the positive results observed from the training study and future research must assess directly the mapping of best exemplars and the change in categorization during training so as to confirm the claims.

Following the line of reasoning, if no concrete evidence is available at this stage concerning the occurrence of reweighting of cues, then what can be inferred to explain the better performance in the subjects trained under HVPT than LVPT in the generalization tests, given that HVPT and LVPT only differs in stimulus variability? Akin to similar training research works (e.g. Iverson & Evans, 2007; Iverson et al., 2009; Iverson et al., 2011), the data obtained in the present experiment suggests that stimulus variability may simply play a role in training the subjects to apply their L2 categories to real speech, instead of leading to changes in best exemplars, which is unknown in the present investigation. Some studies (e.g. Hattori & Iverson, 2009; Heeren & Schouten, 2008; Iverson et al., 2005; Iverson et al., 2009) even showed that subjects could improve their identification accuracy without altering the acoustic cues they used. Provided with stimuli with higher variability, the subjects were trained in the process of how to apply the category knowledge in natural variable speech, but without changing the knowledge itself, i.e. the use of acoustic cues, the introspective notion of what constitutes a best exemplar. While all learners should have prior knowledge of the categories, even though they have difficulty in discerning the differences, HVPT has provided them with a training of the process to put category knowledge in practice in real, variable speech. They are given a wide range of unexpected situations in the training when they have to apply the knowledge from time to time and from trial to trial. This, in part, facilitates and straightens the application process to become more automatic and efficient, thus leading to the more successful and genuine learning (indicated by TGs) in HVPT than LVPT.

Besides directing the subjects' attention to the acoustic cues and making the application of category knowledge smooth, stimuli with high variability may lead to more success in the perceptual learning of the subjects because it simulates real-life experience where the subjects encounter the non-native contrasts produced by different speakers and from a variety of words. Highly variable stimuli provide the greatest opportunity for the subjects to accommodate variations in daily experience, while not ignoring the importance of delivering a concentrated and intensive training. The HVPT approach is more effective than LVPT due to the fact that employing a wider range of speakers in the training stimuli forces the subjects to develop more stimulus-general representations and hence identify the contrast more successfully than the subjects trained under one speaker. Previous studies (e.g. Bohn & Flege, 1990; McKain et al., 1981; Polka, 1995; Werker, 1994; Yamada, 1995) also suggested that even some exceptionally-difficult phonetic categories can be re-learnt and be attuned if the subjects are exposed to the language used in a daily basis. All in all, given sufficient time in training, stimulus variability used in the HVPT group promotes more perceptual improvement than the LVPT group since it gives them the greatest opportunity to induce the general phonetic category from the characteristics of the ensemble of stimuli.

5.1.2 Production Domain (Pretest vs. Posttest vs. Test of Contextualization)

The above discussion has offered possible reasons why HVPT and LVPT were effective in modifying the perception of the target vowel contrasts. Transfer from the perceptual learning to production is also worth investigating by comparing the production performance of the subjects before and after the training. The subjects trained under HVPT and LVPT all improved from the pretest to posttest: HVPT improved their averaged production accuracy by 23.92% (27.95% for /ɪ/-/iː/; 27.50% for /e/-/æ/; and 16.53% for /ʊ/-/uː/), significantly higher than the improvement of 14.44% shown in LVPT (15.76% for /ɪ/-/iː/; 13.62% for /e/-/æ/; and 13.43% for /ʊ/-/uː/). The control had a slight drop of 1.5%. Both the HVPT and LVPT subjects outperformed the control group with significance, while the HVPT group demonstrated more robust learning than the LVPT. While in the TC, the HVPT group outperformed both the LVPT and control groups by 15.70% and 22.49% respectively with significance, but that the LVPT group performing 7.43% better than the control group was not robust. Two summary tables of the production results are provided for recapitulation:

Table 5.3

Summary of the Effects of Perceptual Training Approaches in the Production Pretest vs. Posttest

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Perceptual Training Type (HVPT vs. LVPT vs. Control)	✓	HVPT > Control HVPT > LVPT LVPT > Control
2. Test (Pretest vs. Posttest)	✓	Posttest > Pretest
3. Vowel (front vs. low vs. back)	✓	front > back low > back
4. Test × Perceptual Training Type	✓	Pretest: HVPT vs. LVPT vs. Control (n.s.) Posttest: HVPT > Control HVPT > LVPT LVPT > Control ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ HVPT: Posttest > Pretest LVPT: Posttest > Pretest Control: Posttest vs. Pretest (n.s.)
5. Perceptual Training Type × Vowel	×	
6. Test × Vowel	×	
7. Perceptual Training Type × Test × Vowel	×	

Table 5.4

Summary of the Effects of Perceptual Training Approaches in the TC

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Perceptual Training Type (HVPT vs. LVPT vs. Control)	✓	HVPT > Control HVPT > LVPT
2. Vowel (front vs. low vs. back)	×	
3. Perceptual Training Type × Vowel	×	

Both trained groups' significant improvement in perception was accompanied by significantly improved productions, even to the sentence level, meaning that perceptual training approaches had brought about a considerable amount of success in production as a result of the perceptual training. HVPT is again shown to be a more effective training approach than LVPT, which promotes transfer of perceptual learning to production. The fact that only the HVPT group but not the LVPT outperformed the control group with significance in the TC further supports that generalization occurs more readily among subjects who were trained under highly variable stimuli. The reasons of which have been delineated in detail in previous sections.

A closer read of the percentage of improvement in the production domain (pretest vs. posttest), and comparing and contrasting the amount of learning in both perception and production of different groups of subjects, in lieu of interpreting the production performance alone, give a rather intriguing learning pattern that renders further investigation, which will be discussed in Section 5.5 when perceptual-production vector plots representing the amount of learning in both domains of each individual across groups will be compared. To this end, simply by reading the figures, one can still see that the HVPT group's percentage increase in the production domain slightly exceeded the improvement in perception; while the LVPT group had the other way round. In the pretest, almost all the subjects had better performance in the perception domain than in production; after training, although the general picture held true, some subjects particularly those who were trained under HVPT started to have more accurate production than perception. Some were even reaching the ceiling in production. The results demonstrated by the HVPT subjects do not replicate similar research studies (e.g. Bradlow et al., 1997; Wang, 2000) which showed more perceptual improvement than that of production. It may be attributed to the tendency of producing one of the counterparts of two contrasts by using a most similar existing Cantonese L1 category, i.e. /i/ in /ɪ/-/iː/ pair, /ɛ/ in /e/-/æ/ pair, and /u/ in /ʊ/-/uː/ pair, resulting in higher production scores in general.

The observation concerning transfer of perceptual learning to the aspects of production appears to link to the theoretical accounts of the motor theory (Liberman, 1991; Liberman & Mattingly, 1985) which states that changes or improvement in the perception domain lead to changes in the abstract representation of intended gestures, thus resulting in improvement in production. Under perceptual training, the innate and human-specific module that mediates between speech perception and production will change the internal phonetic representation of motor or vocal tract movements perceived and hence promote production. An alternative possibility, along the line of a direct-realist approach, would suppose that perceptual oriented training leads to the subjects' learning of how to articulate the gestures directly because they have become more attuned to the invariant gestural features of the vowel contrasts which guides to more accurate and gesturally-defined production. The changes in perception after perception learning should hence simultaneously result in changes in the production.

Nevertheless, I hasten to point out that the present study does not aim to offer competing predictions or evidence to differentially support these theoretical accounts. They only offer speculations to the possible underlying mechanisms that are responsible for the transfer of learning. Future research is demanded for testing the theories. Still, the transfer of perceptual learning to production – shown in both the pretest results and the TC – in the absence of explicit production instruction appears to suggest that both processes are closely related to each other. Whether the two processes share a common underlying representation can be tested when the perception results in the group trained under production training are also considered. This will be discussed in Section 5.4.

5.1.3 SUMMARY

Research Questions 1 and 2, and part of Research Question 6, have been answered affirmatively. From the above interpretations, we can conclude that both HVPT and LVPT approaches promote both perceptual and production improvement, though to different extents. As explicated, HVPT is a more effective approach than LVPT, which echoes with many previous studies (e.g. Bradlow et al., 1997; Lambacher et al., 2005; Lively et al., 1993; Lively et al., 1994; Logan et al., 1991) that the use of highly variable stimuli facilitates with the perception and production learning of the subjects. More generalization effects could be found in HVPT than LVPT as well. Taken together, it seems that the findings have reinforced the notion that HVPT is an efficacious approach in modifying even difficult sound contrasts. It will be worth proceeding to the next research question, the effects of perceptual training intensity – a factor that has been overlooked in the literature.

5.2 THE EFFECTS OF PERCEPTUAL TRAINING INTENSITY – RESEARCH QUESTION 3 (&6)

One factor that has been overlooked in previous similar L2 non-native contrast training studies is the intensity of the training paradigm. This section will be devoted to justifying the results from both the posttest and the generalization tests obtained in the study by subjects trained under HVPT or LVPT, which were subdivided into groups under the parameter, perceptual training intensity: some were trained under 10 sessions a day (intensive) while the other two sessions a day (standard). This investigation can cast more light on the speech learning mechanisms and offer implications for L2 classrooms. However, as illustrated in Chapter 4, perceptual training intensity appears to show no effect at all in the learning. So as to simplify the presentation and discussion, the null effect in the perception, production and generalization will be explicated together. Before the discussion, the two relevant research questions are readdressed here for reference:

RQ3. What are the training effects of different perceptual training intensities in
a) the perception and
b) the production of the English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/?

RQ6. Can any learning effect be generalized to

a) the perception of new words produced by both familiar and new speakers, or
b) to the production in a more naturalistic environment?

5.2.1 PERCEPTUAL AND PRODUCTION ASPECTS & THE GENERALIZATION EFFECTS

Results have shown that subjects trained under a perceptual paradigm with different intensity levels, be it standard or intensive, displayed no significant differences in their perceptual, productive and generalization performance in general. The following summary tables give a clear picture of the results obtained when comparing the groups with different intensity levels in different tests:

Table 5.5

Summary of the Effects of Perceptual Training Intensity (HVPT & LVPT) in the Perception Pretest vs. Posttest

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Perceptual Training Intensity (Intensive vs. Standard)	✗	
2. Test (Pretest vs. Posttest)	✓	Posttest > Pretest
3. Vowel (front vs. low vs. back)	✗	
4. Test × Perceptual Training Intensity	✗	
5. Perceptual Training Intensity × Vowel	✗	
6. Test × Vowel	✓	Pretest: front > low front > back back > low Posttest: front vs. low vs. back (n.s.) ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ Front vowel pair: Posttest > Pretest Low vowel pair: Posttest > Pretest Back vowel pair: Posttest > Pretest
7. Perceptual Training Intensity × Test × Vowel	✗	

Table 5.6

Summary of the Effects of Perceptual Training Intensity (HVPT) in the Perception TGs

Factors / Interactions	**Significance?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Perceptual Training Intensity (Intensive vs. Standard)	✗	
2. TG (TG1 vs. TG2 vs. TG3)	✗	
3. Vowel (front vs. low vs. back)	✗	
4. TG × Perceptual Training Intensity	✗	
5. Perceptual Training Intensity × Vowel	✗	
6. TG × Vowel	✗	
7. Perceptual Training Intensity × TG × Vowel	✗	

Table 5.7

Summary of the Effects of Perceptual Training Intensity (LVPT) in the Perception TGs

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Perceptual Training Intensity (Intensive vs. Standard)	✗	
2. TG (TG1 vs. TG2 vs. TG3)	✓	TG3 > TG1 TG2 > TG1
3. Vowel (front vs. low vs. back)	✗	
4. TG × Perceptual Training Intensity	✗	
5. Perceptual Training Intensity × Vowel	✗	
6. TG × Vowel	✗	
7. Perceptual Training Intensity × TG × Vowel	✗	

Table 5.8

Summary of the Effects of Perceptual Training Intensity (HVPT) in the Production Pretest vs. Posttest

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Perceptual Training Intensity (Intensive vs. Standard)	✗	
2. Test (Pretest vs. Posttest)	✓	Posttest > Pretest
3. Vowel (front vs. low vs. back)	✓	front > back low > back
4. Test × Perceptual Training Intensity	✗	
5. Perceptual Training Intensity × Vowel	✗	
6. Test × Vowel	✗	
7. Perceptual Training Intensity × Test × Vowel	✗	

Table 5.9

Summary of the Effects of Perceptual Training Intensity (LVPT) in the Production Pretest vs. Posttest

Factors / Interactions	Significant?	Details of *significant* main effects or interactions (test of simple effects)
1. Perceptual Training Intensity (Intensive vs. Standard)	×	
2. Test (Pretest vs. Posttest)	✓	Posttest > Pretest
3. Vowel (front vs. low vs. back)	×	
4. Test × Perceptual Training Intensity	×	
5. Perceptual Training Intensity × Vowel	×	
6. Test × Vowel	×	
7. Perceptual Training Intensity × Test × Vowel	×	

Table 5.10

Summary of the Effects of Perceptual Training Intensity (HVPT & LVPT) in TC

Factors / Interactions	Significant?	Details of *significant* main effects or interactions (test of simple effects)
1. Perceptual Training Intensity (Intensive vs. Standard)	×	
2. Vowel (front vs. low vs. back)	×	
3. Perceptual Training Intensity × Vowel	×	

The null effect of training intensity in the present study is rather surprising, as the difference in time pressure, retention of interest, task persistence, cognitive loading, just to mention a few, are all underlying parameters of intensity that were expected to affect the performance of the subjects. To my knowledge, the role of training intensity – taken as perceptual training intensity in this study – has not been investigated in any L2 speech training research. Even studies related to pathological issues and speech-language therapies have only begun to touch upon the notion of treatment intensity in recent decades. Although the methodologies adopted in the clinical studies are not exactly equivalent to that of the present study, they still shed some light on how training intensity can be a focus of research, particularly in training speech problems which are phonological in nature.

A number of studies which focused on the impact of dose frequency as a parameter of intensity reported inconclusive outcomes. Some studies (e.g. Gillam, Loeb, & Friel-Patti, 2001; Gillam, Loeb, Hoffman, Bohman, Champlain, & Thibodeau, 2008; Torgesen, Alexander, Wagner, Rashotte, Voeller, & Conway, 2001;Van Hattum, 1959) indicated that higher intensity led to better outcomes; some (e.g. Bambra & Warren, 1993; Riches, Tomasello, & Conti-Ramsden, 2005; Weston and Harber, 1975) reported less intensive intervention was more effective; while others (e.g. Denton, Cirino, Barth, Romain, Vaughn, Wexler, Francis, & Fletcher, 2011; Fein, Golman, Kone, & McClintock, 1956; Ukrainetx, Ross, & Harm, 2009) found no significant differences between intensity levels. However, the varying results of these studies may be due to the fact that they did not control other intensity variables stated in Warren et al. (2007), such as dose frequency or total intervention duration.

The present study defines training intensity by controlling dose, dose form, total intervention duration as constant while keeping dose frequency, cumulative intervention intensity and training episode distribution as independent variables. This more controlled definition of intensity, however, has only been adopted in very few clinical studies which also showed inconclusive outcomes: Page, Pertile, Torresi, & Hudson (1994) showed no significant differences between intensity levels; while Allen (2013) demonstrated the intensive intervention was more effective. The results in the present study also showed no robust differences between those who were trained under the standard or intensive scheme. Besides the problem of small sample size, there are two more explanations that I will provide for this finding.

5.2.1.1 GENUINE NULL EFFECT OF PERCEPTUAL TRAINING INTENSITY

The present null effects observed may simply because the two intensity levels have no meaningful effect. Once the subjects' phonological systems were stimulated, they began to learn the contrasts. It is probably the number of exemplars that drive them to detect the subtle differences in the target sounds because the exposure to which is fundamental to advancing the reweighting of cues or application of category knowledge of the subjects. This hints that because the number of stimuli provided, regardless of high or low variability, was high enough to bring about significant learning, whether the stimuli were distributed within a short or long period of time makes no difference to their learning. Future research may compare the present training with one that has a shorter period and hence the number of training sessions to see if the effect of intensity will become more evident.

5.2.1.2 ACTIVE INGREDIENTS MATTER MORE

Most clinical research on speech disorder has pointed out that intensive treatment in general brings more significant improvement. However, it may be the active ingredients of the training that matter more (Baker, 2012; Lee, Kaye, & Cherney, 2009). I interpret the meaning of *active ingredients* in training as the type of training utilized (e.g. identification or discrimination task; perception only or production only), the training components used (e.g. perceptual stimuli with high/low variability, computer program, types of feedbacks), how the training is delivered (e.g. through normal class teaching, a self-paced computer program) in L2 training studies. This interpretation is based on those clinical studies whose meanings are about the best forms of dose (Baker, 2012) such as the acts of the subjects (e.g. production practice) and inputs that lead to successful training. Intensity alone is insufficient to determine the training outcome; rather, it is the active ingredients that are more influential. Following this line of reasoning, it is probably the adoption of perceptual stimuli (regardless of high or low variability), identification tasks, the use of feedbacks, etc. that already become the contributing parameters leading to successful modification of perception or production of the non-native contrasts in investigation, diminishing the effect of training intensity.

5.2.2 SUMMARY

Given that the definition of training intensity is still not unanimous even among clinical researchers (e.g. Allen, 2013; Baker, 2012; Warren et al., 2007) who have rather extensive investigations of the notion of intensity, one cannot give a solid conclusion to the present findings hastily without considering the above three justifications which are new topics that demand follow-up research. However, the results in the present study – though being negative – have offered preliminary insights into a possibility of investigating this factor. This rather unchartered field certainly requires more investigation especially in L2 speech learning research. Meanwhile, if the factor of training intensity displays no real effects in training the perception and production of the three vowel contrasts, it may have given more choices to language teachers that highly intensive training within a short period of time is as effective as those spread over a longer period of time, which is the preferred norm in language classrooms. This will help modify teaching schedule, pedagogical strategies and planning in language lessons.

Research Question 1 to 3 all pertained to the notion of perceptual training. Another focus of the present research was the effect of production learning to both perception and production aspects, and this will be discussed in the next section.

5.3 The Effects of Explicit Production Training – Research Question 4 (&6)

Having the positive results of the effectiveness of the two perceptual training on perception and production of the target vowel contrasts in mind, I now turn to discuss the results obtained from the subjects who were trained under explicit production training. This will explore the learning effects from the other way round: how training in the production can influence learning in perception. The generalization effects are also one focus. To logically present the arguments, I will discuss the effects of production training on production first before perception. Here, the two research questions are reiterated for reference:

RQ4. Is the production training effective in improving
a) the perception and
b) the production of the English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/?

RQ6. Can any learning effect be generalized to
a) the perception of new words produced by both familiar and

new speakers, or

b) to the production in a more naturalistic environment?

5.3.1 PRODUCTION RESULTS (PRETEST VS. POSTTEST VS. TEST OF CONTEXTUALIZATION)

Positive results were obtained from the subjects trained under production paradigm, as they not only improved their production of the three target vowel pairs for 17.42% from the pretest to posttest, but they also performed significantly better than the control group for 35.73%. When in TC, the group with production training also performed better than the control group for 16.16% in general. The results both signified that production training had been an effective program in promoting more accurate production of the vowel contrasts; the learning of which could also be transferred to the production at the sentence-level. This table offers a quick review of the results:

Table 5.11

Summary of the Effects of Production Training in the Production Pretest vs. Posttest

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Production Training (With vs. Without)	✓	With > Without
2. Test (Pretest vs. Posttest)	✓	Posttest > Pretest
3. Vowel (front vs. low vs. back)	✗	
4. Test × Production Training	✓	<u>Pretest:</u> With vs. Without (n.s.) <u>Posttest:</u> With > Without ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ <u>With:</u> Posttest > Pretest <u>Without:</u> Posttest vs. Pretest (n.s.)
5. Production Training × Vowel	✗	
6. Test × Vowel	✗	
7. Production Training × Test × Vowel	✗	

Table 5.12

Summary of the Effects of Production Training in the TC

Factors / Interactions	Significant?	Details of *significant* main effects or interactions (test of simple effects)
1. Production Training (With vs. Without)	✓	With > Without
2. Vowel (front vs. low vs. back)	×	
3. Production Training × Vowel	×	

The promising results are not surprising, as a great number of previous studies (e.g. Archibald, 1998; Alves & Magro, 2011; Arteaga, 2000; Castino 1996; Cenoz & Lecumberri, 1999; Flege, Frieda, Walley, & Randazza, 1998; Gon-zález-Bueno, 1997; Gordon, Darcy, & Ewert, 2012; Hattori & Iverson, 2008; Leather, 1999; Lord, 2005; Major, 1998; Moyer, 1999; Neufeld, 1977,1978, 1979; Simões, 1996) using explicit production training on segmental or suprasegmental contrasts have shown favorable results. The effectiveness of the training can be attributed to the adoption of an input-output-feedback scheme.

The basic premise that enables successful language acquisition to take is the input provided to the language learners (Krashen, 1982). The present pronunciation training, though was short, had given quality inputs to the learners to facilitate production learning. Explicit instructions delivered by me as the first channel of offering visual speech and information were deemed as beneficial to the subjects' learning. Explanations of the articulatory aspect of sounds by using hands, mirrors, and illustrations, as well as deliberately pointing out the differences between each vowel contrast have raised their awareness of the L1-L2 phonological differences.

The input in the production training was perceptual in nature because no production training can be delivered without a learner listening to some inputs. However, this approach differed from the perceptual training paradigm adopted in this study in the way that extra instructions on the articulation of sounds and feedback from the researcher were also offered. Based on previous studies (e.g. Hazan et al., 2005; Hirata & Kelly, 2010; Massaro, Bigler, Chen, Perlman, & Ouni, 2008) which reported that visible speech can contribute to learning new speech distinctions and promoting active learning, I adopted video materials with native speakers' productions of the three target vowel pairs to facilitate the subjects' learning. The use of video was to direct learner's attention to the facial movement of the native speaker so that they could learn the articulations made by model speakers and became aware of their areas of deficiency and the differences between their own pronunciations and the native speaker's. Akahane-Yamada and colleagues (1997) suggested in their study that the mental representation of L2 phonetic categories incorporates both auditory and visual information. Thus, through listening and watching the videos, the integration of both auditory and visual information can be enhanced and it can benefit the leaners' production.

Receiving simply the comprehensible input does not appear to be sufficient for successful production improvements. It is practicing the sounds that became fundamental to advancing more fluent and accurate productions. By producing the minimal pair words in this study, the subjects were allowed to compare their own productions with the input models, forming more accurate and target-like productions consequently. The subjects were asked to imitate the native speaker's productions when watching the video clips as well. This step allows the subjects and the researcher to verify whether they have developed the ability to model a specific sound (Logan & Pruitt, 1995). When they also received feedbacks (to be discussed in the next section) from the researcher, they would become more aware of their ill-forms and reflect within themselves before making another attempt. The progress of attempting to produce a more accurate output simply offered the subjects a platform to test their own hypothesis concerning how the L2 pronunciation should be like. Each of their output serves as a test sample or vehicle that is to compare to the researcher's feedback, or the native model. Gradually, they would be able to develop the ability to adjust the use of articulators, the auditory system, motor programs, etc. so as to make a satisfying production (Leather & James, 1997). Through this, their productions will become more target-like.

Simply letting the subjects receive inputs and produce outputs cannot guarantee successful learning in the production aspect. I intended to bring the subjects to focus on the contrasts also by using corrective feedbacks in the form of explicit corrections, recasts and elicitation. These feedbacks acted as a way to interact with the subjects as well as another type of input that allowed them to *notice* the discrepancies between their output and the target productions after they had produced the sounds (Schmidt, 1990). Feedbacks can stimulate the subjects to produce the target forms and develop their ability in monitoring their own productions.

It has been highlighted in Chapter 2 and 3 about the importance of adding production training as one variable in investigation: it aimed to provide evidence to whether learning in the production domain can be transferred to the perception domain and hence give more evidence to the underlying mechanisms of speech learning. This intriguing question will be dealt with in detail in the next section.

5.3.2 Perceptual Results (Pretest vs. Posttest vs. Tests of Generalization)

Generally speaking, the group with production training performed significantly better than the control group for 9.34% in perception identification posttest across all three vowel pairs. It is also only the group with production training that improved in the perception of the three vowel pairs for 6.36% from the pretest to the posttest with significance, whereas the control group did not. Across the three TGs, the production training group also had more significant improvement than the control group for an average of 4.90%, but further analysis showed that the generalization effect was found only in the new speaker condition. The tables present a summary of the above findings for easier reference:

Table 5.13
Summary of the Effects of Production Training in the Perceptual Pretest vs. Posttest

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Production Training (With vs. Without)	✗	
2. Test (Pretest vs. Posttest)	✗	
3. Vowel (front vs. low vs. back)	✓	low > front low > back
4. Test × Production Training	✓	Pretest: With vs. Without (n.s.) Posttest: With > Without ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ With: Posttest > Pretest

		Without : Posttest vs. Pretest (n.s.)
5. Production Training × Vowel	×	
6. Test × Vowel	×	
7. Production Training × Test × Vowel	×	

Table 5.14

Summary of the Effects of Production Training in the TGs

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Production Training (With vs. Without)	✓	With > Without
2. TG (TG1 vs. TG2 vs. TG3)	✓	TG3 > TG1 TG2 > TG1
3. Vowel (front vs. low vs. back)	✓	low > front low > back
4. TG × Production Training	×	
5. Production Training × Vowel	×	
6. TG × Vowel	×	
7. Production Training × TG × Vowel	×	

The partially positive results are, however, interesting in the way that if individual vowel pairs are concerned, only the improvement in the low vowel pairs from the pretest to posttest was significant and showed generalization effects. The fact that there exists a difference in the performance of vowel pairs will be fully dealt with when Research Question 7 is answered. However, the present positive results in general, although rather moderate, still suggested that production training had some effects on the perception of the vowel contrasts.

Despite the scarcity of research work on the effects of production training on perception by L2 learners, the existing ones show inconclusive results. Baese-Berk (2010) who investigated the effect of production training on the perception and production of non-native contrasts found that production training influenced perceptual learning less than those who were trained in the perception alone. Some (e.g.

Aliaga-Garcia, 2010; Gómez Lacabex, 2009; Hirata, 2004; Leather, 1999) found positive transfer of learning from production to perception, and the learning from perception training to both domains were to similar extent. Catford and Pisoni (1970) found that both perception and production training influenced the two modalities positively, but production training had more positive effects on production than perception learning. Some (e.g. Hattori, 2009; Hattori & Iverson, 2008) found that production training simply did not promote perception learning. However, the moderate gain in the perception modality from production training in the present study may have to be understood with cautions. One may interpret that transfer of production learning to the perception domain means that the learners begun to successfully exercise their abilities to map other's output to their own phonetic space, hence resulting in the positive results. It may be fairer to claim that, given only the moderate results, it is the implicit perception components which could not be completely isolated from the production training that influenced the perception learning. Yet, the quantity of perception elements in the production training provided may not suffice a strong perception learning effect. Meanwhile, since the generalization effect was found only in the new speaker condition, it is further speculated that it is the available, although limited, perceptual components with two speakers' information (the native speaker in the video clips and the researcher during the imitation task) that enriched the learners' generalization ability in the new speaker condition.

Another possibility is that in production training, the subjects not only received instructional materials about the articulation details, but they also listened to the native speaker's productions in an audio-visual form. During their imitation, besides having chances to verify if they had begun developing the ability to model a specific sound, they might have the link between the perception and visual information enhanced, as well as having the perception modeling ability in development (Akahane-Yamada et al., 1997; Logan & Pruitt, 1995; McGurk & MacDonald, 1976). The moderate finding suggested that the ability was developing, but it might not be sufficient to overcome other impacts. This view is in line with previous findings that imitation tasks cannot fully avoid linguistic influence (Alivuotila, Hakokari, Savela, Happonen, & Aaltonen, 2007; Nielsen, 2007). Thus, it is still not solely the act of producing the sounds that led to change in perception, but also the visual component used in the video clips.

Even with the present data, one still cannot conclude whether the perception and production processes are independent or closely related to each other. Allowing a longer period of production training as well as minimizing the perceptual components in production training may perhaps offer more insights into this matter.

5.3.3 SUMMARY

Findings in the present study have offered evidence to support the efficacy of an explicit production training paradigm in improving learners' production performance, but the effects on perception were fairly moderate. Still, the result is not conclusive and future research with more controlled variables is demanded.

5.4 THE EFFECT OF TRAINING DIVERSITY & DESIGN – RESEARCH QUESTION 5 (&6)

Comparing the effectiveness of the various training types adopted in the present research, including HVPT, LVPT, production training as well as a combination of HVPT/LVPT with production training[17], is also one focus of the present investigation. It casts more light on the usefulness of different training methods and provides evidence for future adoption of similar paradigms in L2 classrooms. This section will be devoted to answering the two research questions and mainly offering a summary for the effectiveness of the training paradigms adopted for comparison in the present study. A discussion of all the protocols adopted will also offer an integral platform for investigating the link between speech perception and production, which will be presented after the following two research questions are discussed:

RQ5. Are there are any differences in both
a) the perception and
b) the production performances of the English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ when the participants are i) trained only under either HVPT or LVPT; ii) trained only under the explicit production training or iii) trained under both the perception approach (either HVPT or LVPT) and the production training, and how different are they?

RQ6. Can any learning effect be generalized to
a) the perception of new words produced by both familiar and new speakers, or
b) to the production in a more naturalistic environment?

[17] Mnemonic abbreviations for the five types of training: HVPT only = H; HVPT with production = HP; LVPT only = L; LVPT with production = LP ; production only = P.

5.4.1 Perceptual Aspect (Pretest vs. Posttest vs. Generalization)

As fully explicated in the above research questions, it is evident that both HVPT and LVPT were equally useful in training a basic set of stimuli with vowel pairs which posed perceptual and production problems to the target participant; whereas production training only contributed slight yet significant perceptual learning. Further analysis showed that H, HP, L and LP types were all performing significantly better than the P type. No robust difference was found between the performance of subjects trained under H/L type and HP/LP type, meaning that an addition of production training did not bring extra benefits to the subjects. When it came to the generalization tests, the learning pattern was not the same. Besides knowing that HVPT was more effective than LVPT and the production training, a comparison of all training types revealed that the HP type also had more generalization learning than L, LP and P types. The performance of HP and H types were, however, not statistically different. This table gives a more structured representation of the above results for easier comparison:

Table 5.15

Summary of the Effects of Training Type in the Perceptual Pretest vs. Posttest

Factors / Interactions	Significant?	Details of *significant* main effects or interactions (test of simple effects)
1. Training Type (H vs. HP vs. L vs. LP vs. P)	✓	H, HP, L, LP > P
2. Test (Pretest vs. Posttest)	✓	Posttest > Pretest
3. Vowel (front vs. low vs. back)	✓	low > front low > back
4. Test × Training Type	✓	Pretest: H vs. HP vs. L vs. LP vs. P (n.s.) Posttest: H, HP, L, LP, P ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ H, HP, L, LP, P: Posttest > Pretest
5. Training Type × Vowel	✗	
6. Test × Vowel	✓	Pretest:

		front > low front > back back > low Posttest: front vs. low vs. back (n.s.) ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ Front vowel pair: Posttest > Pretest Low vowel pair: Posttest > Pretest Back vowel pair: Posttest > Pretest
7. Training Type × Test × Vowel	✗	

Table 5.16

Summary of the Effects of Training Type in the TGs

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Training Type (H vs. HP vs. L vs. LP vs. P)	✓	H > L, LP, P HP > L, LP, P
2. TG (TG1 vs. TG2 vs. TG3)	✓	TG2 > TG1 TG2 > TG3
3. Vowel (front vs. low vs. back)	✓	low > front low > back
4. TG × Training Type	✓	TG1: H, HP > L, LP, P TG2, TG3: H, HP > P ~~~~~~~~~~~~~~~~~ H, HP: TG1 vs. TG2 vs. TG3 (n.s.) L, LP: TG3 > TG1 TG2 >TG1 P: TG3 > TG1
5. Training Type × Vowel	✗	
6. TG × Vowel	✓	TG1, TG2, TG3: low > front low > back ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ front, low: TG3 > TG1
7. Training Type × TG × Vowel	✗	

As mentioned in Chapter 2, there were very few studies (e.g. Aliaga-García & Mora, 2009; Baese-Berk, 2010; Tsushima & Hamada, 2005; Wong, 2013b) investigating the effects of perception training plus production training on training the perception of some contrasts and the results were inconclusive. Baese-Berk (2010) even found that production training on top of a perception training disrupted perceptual learning. The present results simply contradicted Baese-Berk's results, but it might be due to the design of experiment. It was administered by having the subjects completed both perception and production training simultaneously: instead of doing a traditional identification task in the perception test, subjects were instructed to imitate the perceptual stimulus once they heard that. The present experiment ran the sessions of the two training protocols alternatively, which was not as cognitive or linguistic demanding as the task adopted in Baese-Berk's (2010). The subjects in my study can focus their attention on the specific task required in each training session, hence resulting in different results.

One may argue that it is simply logical to think that *more is better*: offering an addition of production training to perceptual training benefited subjects because they received more learning. However, similar to the findings in some pathological studies (e.g. Warren et al., 2007) which noted that *more is not necessarily better*, the present study showed that an addition of production training on HVPT or LVPT, i.e. HP/LP, had neither interference nor benefits to perception learning. Results in the TGs showing that only the H and HP types, but not L or even LP types, that had generalization effects implied that it is rather the quality or nature of training that matters more. Using highly variable perceptual stimuli in training appears to be the necessary and sufficient condition for promoting more accurate identification of the non-native contrasts, while the learning of which can also be generalized to new conditions.

5.4.2 PRODUCTION ASPECT (PRETEST VS. POSTTEST VS. CONTEXTUALIZATION)

Notwithstanding the particular effectiveness of HVPT (H or HP types, since an addition of production training did not offer extra benefits to the perceptual learning of the subjects) shown in perceptual learning and generalization to novel voices and speakers of non-native contrasts, the learning pattern across training types was different in the production aspect. Although each training type helped modify the production patterns of the three vowel pairs to some extent, results showed that the performance did not differ from any other type with any significant difference. The only exception was the HP type being the most efficacious one and that it outperformed all other types with robust differences. All training types improved from the pretest to posttest with significance, but only the HP type performed better than H, L, LP, and P types with robust differences. While in TC, it was only the HP type which had better contextualization performance than L, LP and P types. This summary table again allows a brief review of the above results:

Table 5.17

Summary of the Effects of Training Type in the Production Pretest vs. Posttest

Factors / Interactions	**Significant?**	**Details of *significant* main effects or interactions (test of simple effects)**
1. Training Type (H vs. HP vs. L vs. LP vs. P)	✓	HP > H, L, LP, P
2. Test (Pretest vs. Posttest)	✓	Posttest > Pretest
3. Vowel (front vs. low vs. back)	✓	low > front low > back
4. Test × Training Type	✓	<u>Pretest:</u> H vs. HP vs. L vs. LP vs. P (n.s.) <u>Posttest:</u> HP > H, L, LP, P ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ <u>H, HP, L, LP, P:</u> Posttest > Pretest
5. Training Type × Vowel	✗	
6. Test × Vowel	✓	<u>Pretest:</u>

		front > back low > back <u>Posttest:</u> front > back low > back ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ <u>H, HP, L, LP, P:</u> Posttest > Pretest
7. Training Type × Test × Vowel	✗	

Table 5.18

Summary of the Effects of Training Type in the TC

Factors / Interactions	Significant?	Details of *significant* main effects or interactions (test of simple effects)
1. Training Type (H vs. HP vs. L vs. LP vs. P)	✓	HP > L, LP, P
2. Vowel (front vs. low vs. back)	✓	front vs. low vs. back (n.s.)
3. Training Type × Vowel	✗	

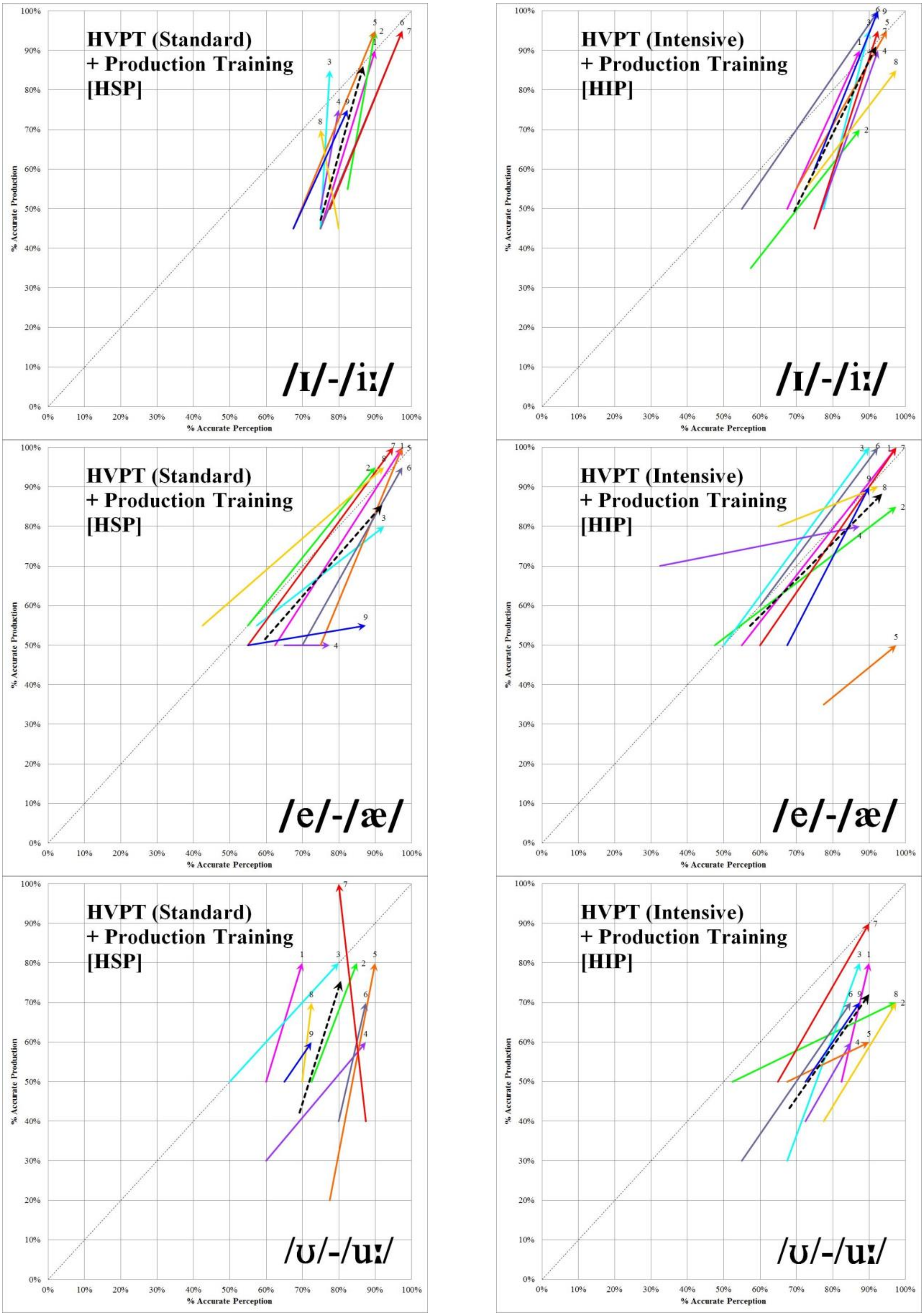
The results in the production aspect revealed that although all training paradigm were useful in bolstering the production accuracy, though to different degrees, it was the HVPT with production training approach that was most outstanding. Different from perceptual learning in which the sole adoption of highly variable perceptual stimuli sufficed, learning in the production domain appeared to require not only explicit and direct production training, but also some perceptual training – training in only either one domain could not give optimal outcomes. Production training alone also facilitated more in the production aspect than in the perception; however, only when it is accompanied by a perceptual training with highly variable stimuli that becomes more indicative of improvement. This combination appears to boost the training efficacy to a higher degree.

5.4.3 The Relationship between Perception and Production

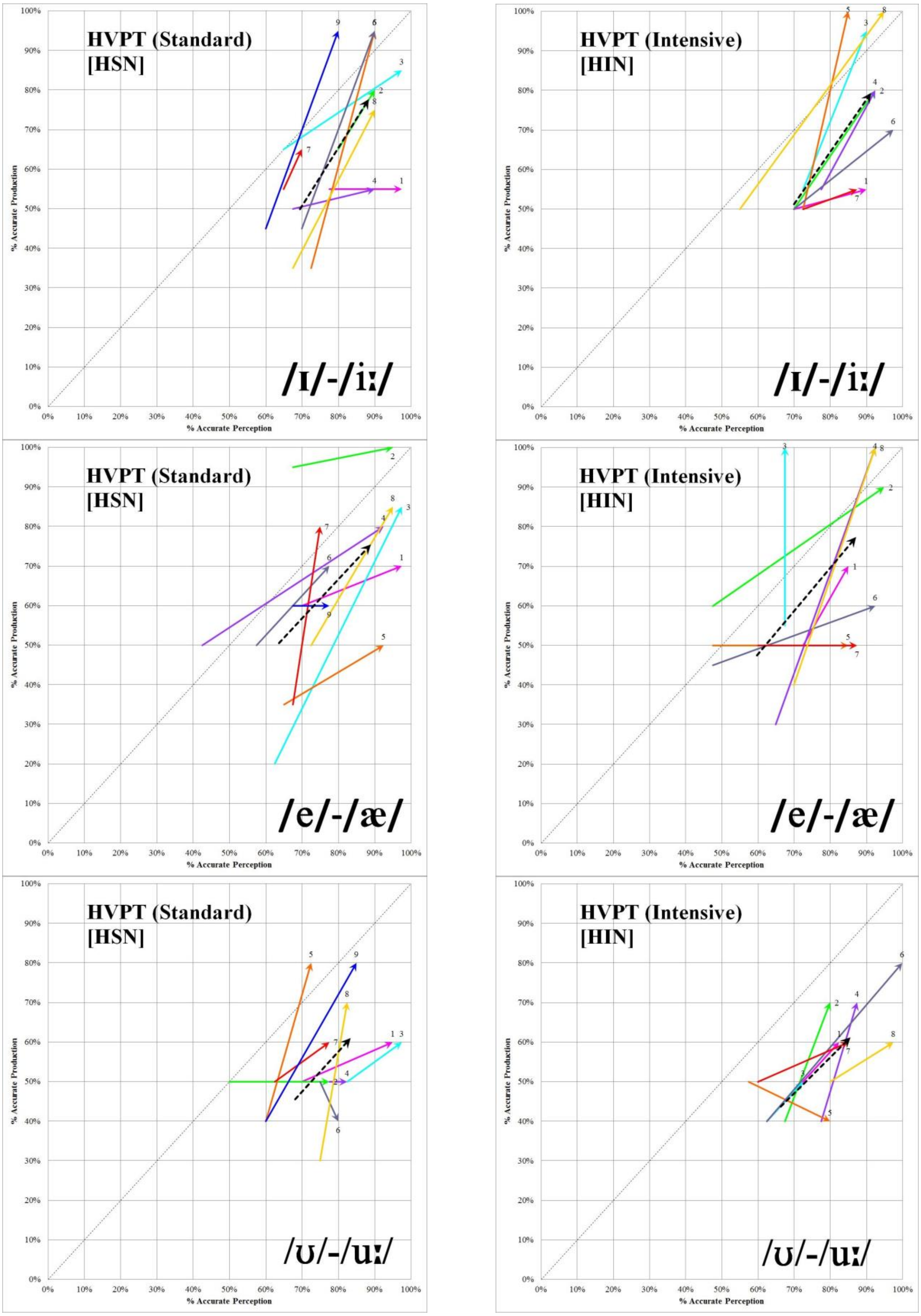
The present study did not aim to examine the link between speech perception and production; still, the present findings can shed more light on this issue by investigating the learning in both modalities. Since all the research questions concerning the effectiveness of the training paradigms have been discussed in the above sections, it is an appropriate point to stop and understand the data from another perspective. This will be achieved by comparing and contrasting the amount of learning in perception and production of all ten groups of subjects from the pretest to posttest, grouped under each training type, by using a representation of perceptual-production space presenting individual performances (adopted from Bradlow et al., 1997). All vector plots of the five training types, subcategorized under intensity levels, would be presented on the next several pages. The results of the control group were also given for a baseline comparison.

The perceptual-production space of each group shows the amount of learning in both domains. The x-axis represents the subjects' accuracies in the identification of each vowel pair (shown in three separate graphs); the y-axis represents the percentage of accurate production of the target vowels. Each vector indicates an individual subject's performance in both domains, with the direction of arrow indicating the change from the pretest to posttest. The bold, dotted arrow shows the mean percentage of the group while the grey diagonal shows the ideal direct proportional change which perfectly correlates the change in perception and production (slope = 1).

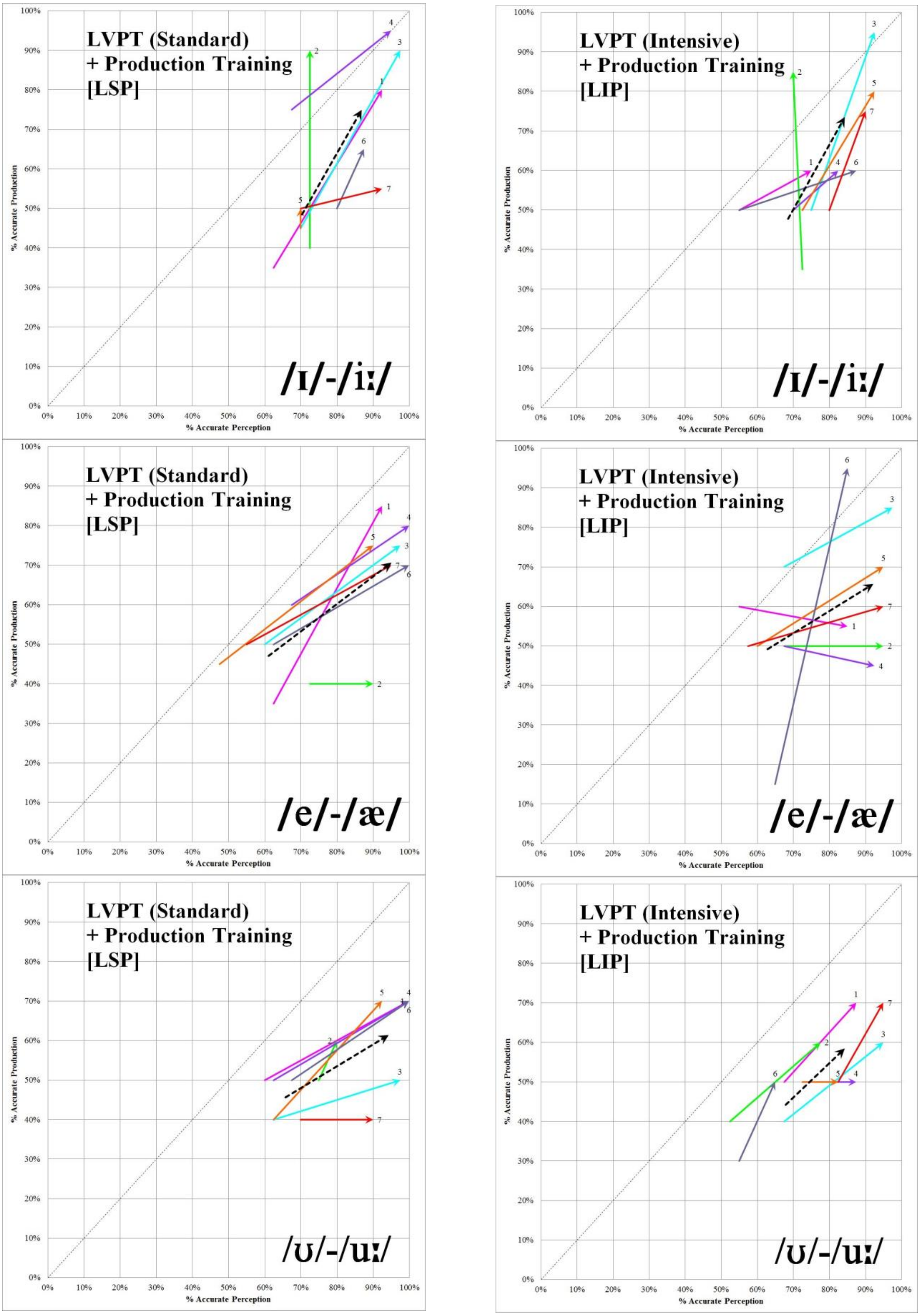
Previous studies (e.g. Bradlow et al., 1997; Hazan, et al., 2005; Lengeris, 2009;



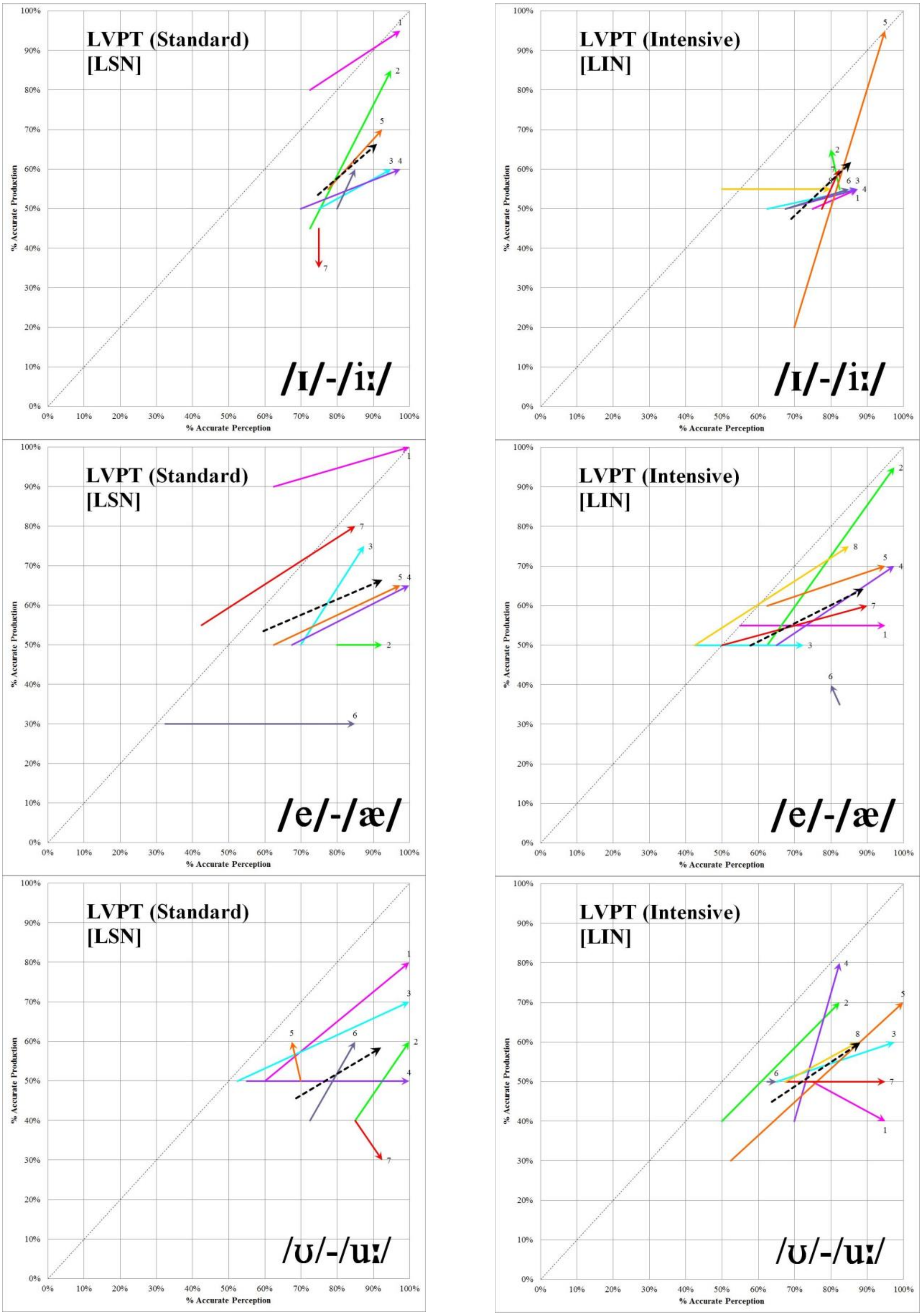
Figures 5.1a, b Vector plots of individual subjects' perceptual identification accuracies (x axis) and target productions (y axis) from the pretest to the posttest in HVPT (intensive/standard) + production training type [HP]. A numbered vector is used to indicate each individual's performance. The bold and dotted arrow represents the group mean, while the diagonal indicates the hypothetical and ideal location for a perfect correlation between speech perception and production



Figures 5.1c, d Vector plots of individual subjects' perceptual identification accuracies (x axis) and target productions (y axis) from the pretest to the posttest in HVPT (intensive/standard) type [H]. A numbered vector is used to indicate each individual's performance. The bold and dotted arrow represents the group mean, while the diagonal indicates the hypothetical and ideal location for a perfect correlation between speech perception and production



Figures 5.1e, f Vector plots of individual subjects' perceptual identification accuracies (x axis) and target productions (y axis) from the pretest to the posttest in LVPT (intensive/standard) + production training type [LP]. A numbered vector is used to indicate each individual's performance. The bold and dotted arrow represents the group mean, while the diagonal indicates the hypothetical and ideal location for a perfect correlation between speech perception and production



Figures 5.1g, h Vector plots of individual subjects' perceptual identification accuracies (x axis) and target productions (y axis) from the pretest to the posttest in LVPT (intensive/standard) type [L]. A numbered vector is used to indicate each individual's performance. The bold and dotted arrow represents the group mean, while the diagonal indicates the hypothetical and ideal location for a perfect correlation between speech perception and production

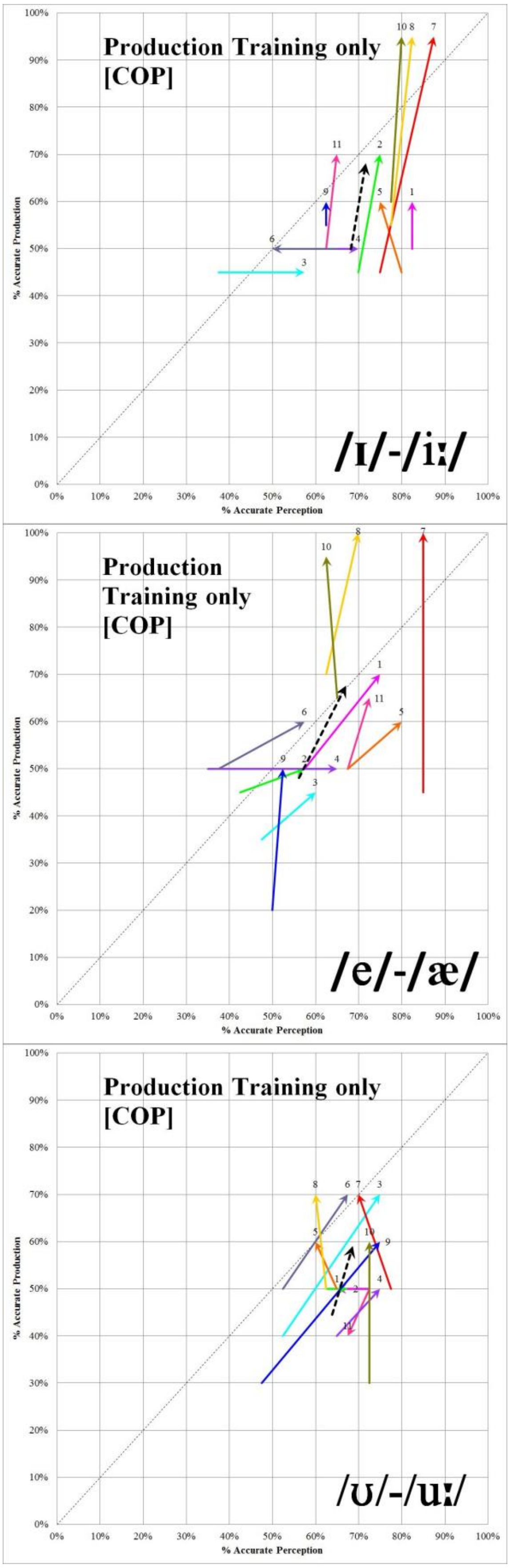


Figure 5.1i Vector plots of individual subjects' perceptual identification accuracies (x axis) and target productions (y axis) from the pretest to the posttest in Production Training type [P]. A numbered vector is used to indicate each individual's performance. The bold and dotted arrow represents the group mean, while the diagonal indicates the hypothetical and ideal location for a perfect correlation between speech perception and production

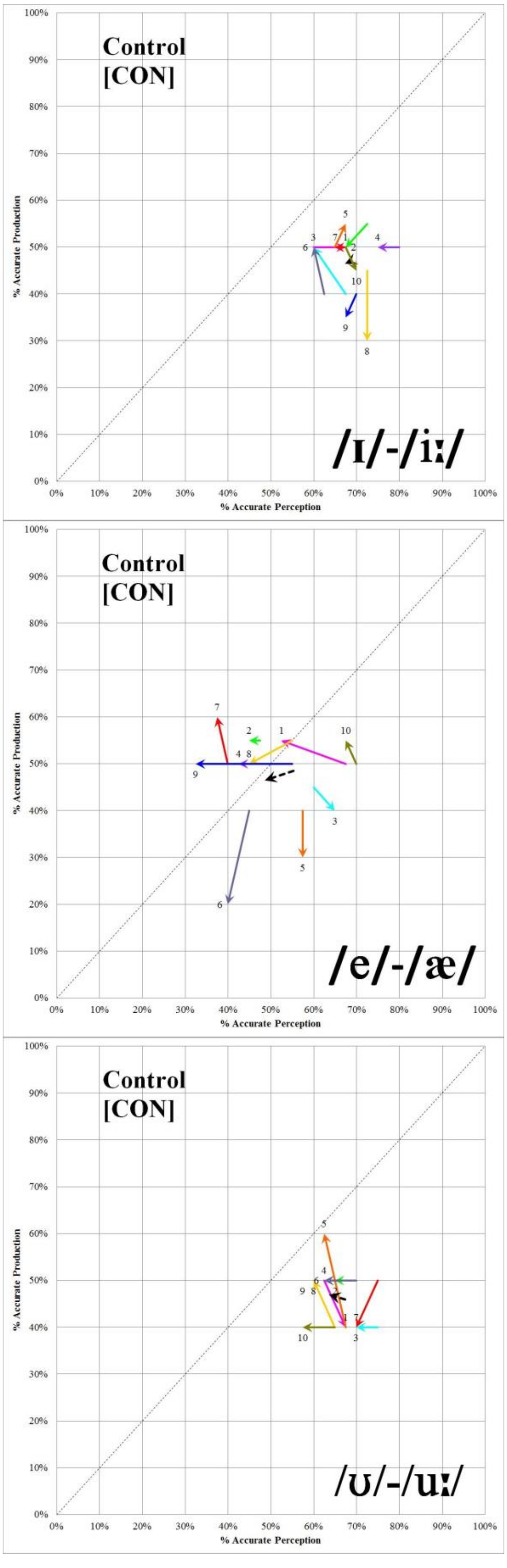


Figure 5.1j Vector plots of individual subjects' perceptual identification accuracies (x axis) and target productions (y axis) from the pretest to the posttest in the control group. A numbered vector is used to indicate each individual's performance. The bold and dotted arrow represents the group mean, while the diagonal indicates the hypothetical and ideal location for a perfect correlation between speech perception and production

Wang, 2000) have demonstrated large individual differences in cross-sectional studies examining the perception and production of non-native contrasts. Although the current study does not aim to explore the idiosyncrasies present in the effectiveness of the training approaches by looking at individual differences, it is still worth looking into individuals' performance as a group to observe the general change in the performance so as to collectively compare how effective the approaches were.

Figures 5.1 a-j display the amount of learning in perception and production of all the subjects across different training paradigms. In general, all groups that received training, regardless of its types, demonstrated fairly noticeable improvement in both domains, as indicated by the long vectors. The performance of the control group subjects has no pattern at all. Most of them had only minor changes (shown by the relatively short vectors) in their performance. Three observations are particularly intriguing and will be presented below.

Firstly, even though all trained groups in general showed a positive slope in their vectors and that the HP (HSP, HIP) and H (HSN, HIN) types had relatively longer vectors than the LP (LSP, LIP), L (LSN, LIN) and P (COP) types, a very wide range of individual differences is still observed, particularly in the LP, L and P types. Some subjects had extremely large improvement and became outstanding in the graphs (e.g. subject 6 in LIP for the low vowel pair; subject 5 in LIN for the front vowel pair) whereas some lagged behind and did worse in the perception (e.g. subject 7 in HSP for the back vowel pair; subject 2 in LIP for the front vowel pair) or production (e.g. subject 6 in HSN for the back vowel pair; subject 7 in LSN for the back vowel pair) posttest.

Some subjects' performance in perception (e.g. subject 2 in LSP for the front vowel pair; subject 7 in LSN for the front vowel pair) or production (e.g. subject 6 in LSN for the low vowel pair; subject 3 in COP for the front vowel pair) remained the same even after training. It is particularly evident in the P type in which many subjects had an almost vertical vector which indicates no or very little improvement in the perceptual domain. All these individual differences hinted at two points: the learning of the two domains proceeded at different rates and the rates varied by individuals.

Even though individual differences were observed and only the performance in the pretest and posttest was compared, the general, more consistent success and relatively fewer individual differences among the subjects in the H and HP types already suggests that perceptual and production learning were substantial, provided that the subjects were given sufficient and highly variable stimuli. Almost all the subjects could gain from the paradigms. Meanwhile, if only the performance of the H type is concerned, similarly positive slopes and relatively long length of the vectors also provided clues that there existed a possible link between speech perception and production since a considerable amount of perceptual learning did transfer to production. In contrast, even in statistical analysis no significant differences were observed in the perceptual pretest and posttest for the H and L types, the more scattered vector distribution of the subjects in the L type due to a wide range of individual differences suggests that perceptual learning with single-environment stimuli was neither a necessary nor a sufficient condition for perception and production enhancement. The almost-vertical vectors observed in a number of subjects in the P type confirm that training in the production domain may not contribute much to perception.

The second observation is that a wide range of individual differences constitutes a pattern in terms of vowel pair as well: both the front and low vowel pairs demonstrated less individual differences than the back vowel pair. From statistical analyses, it is shown that the factor vowel has always been robust, meaning that the degree of learning of the three vowel pairs was not equivalent. The more observable individual differences in the back vowel pair also suggests that this vowel pair is the most difficult one to learn. The reasons will be delineated in detail when the last research question is answered in the next section.

Finally, it is clear that the trained subjects' perception performance (x axis) exceeded the production performance (y axis) in general in the pretest, and it still held true in the posttest as shown by almost all vectors laying in the right sector of graph. There are also many long vectors with slopes greater than 1 across different training types, indicating more learning in the production than perception. It did not replicate previous research (e.g. Bradlow et al., 1997; Wang, 2000), however, which showed more perceptual improvement than production.

5.4.4 SUMMARY

This section has served as a wrap-up that compares the effectiveness of all training types. The HP type was shown to be the best paradigm in promoting both perception and production mastery, and the H type was also more beneficial than the others. Comparing the degree of perception-production learning by using the vector plots also gives an overall picture of individual performance. It is also suggested that future training paradigm design should pay attention to individual differences and design materials that are suitable for different learners.

5.5 Generalization Effects of Training – Research Question 6

The effects of learning in generalization in the perception and production aspects have been discussed under each main research question in the above sections. This part will be devoted to summarizing the findings of generalization effects and discussing several points further to deepen the understanding of the effects of generalization, particularly by comparing the difference between the scores in the pretest and each TG (TG1: new words by a new speaker; TG2: new words by a familiar speaker; TG3: familiar words produced by a new speaker) adopted in perceptual generalization. A recapitulation of the related research question is as follows:

RQ6. Can any learning effect be generalized to
a) the perception of new words produced by both familiar and new speakers, or
b) to the production in a more naturalistic environment?

5.5.1 Perceptual Aspect

By comparing the identification scores of each TG with the pretest, one can have a first glimpse of the degree of improvement in generalization that each group attained. Of the four factors I investigated – effects of perceptual training, perceptual training intensity, production training and training types, only some showed the generalization effects. Only the HVPT groups (with or without an addition of production training) demonstrated generalization effects to both new word and new speakers condition, shown by the insignificant differences between TG1, TG2 and TG3. The LVPT groups' (with or without an addition of production training) TG3 difference scores (TG3 minus pretest) were always significantly higher than those in TG1, and the improvement in TG2 (TG2 minus pretest) was also significantly higher than TG1, indicating that familiar words and speakers were still easier for them to identify. Production training could also only promote generalization in the new speaker but not in new word condition. The perceptual training intensity again showed no effects on generalization. All these findings indicated that if the subjects are given perceptual stimuli with higher variability, they demonstrated more genuine perceptual learning.

5.5.1.1 Generalization to New Speakers

Findings in the present study showed that only HVPT and production training could benefit the subjects through generalizing perceptual learning to new voices. This ability to generalize learning to a new voice was in line with previous findings (e.g. Bradlow et al., 1997; Pisoni & Lively, 1995; Lively et al., 1994; Wang, 2002; Wong, 2010, 2012) that providing learners with perceptual stimuli with higher variability facilitated perceptual learning. These highly-variable stimuli help the subjects to discover the critical acoustic differences between talkers. Even the two voices in the production training may have a contribution to the transfer of learning to a new voice, although the effect was rather moderate.

5.5.1.2 GENERALIZATION TO NEW WORDS

Consistent with previous research findings (e.g. Broersma, 2005; McClaskey, Pisoni, & Carrell, 1983; Rochet, 1995; Wang, 2002; Wong, 2010, 2012), the present result revealed that learning in one context (i.e. training with the basic set with a set of phonetic environments) can sometimes be transferred to new contexts (i.e. novel tokens in TG1 and TG2). In both generalization tests in which new words were involved, the new contexts were set as having the target vowels before fricatives such as /ð/ and /ʃ/, affricates, nasals, the lateral and some consonant clusters. Some were even disyllabic words. While in the training, all target vowels were placed in front of the plosives /p/, /t/, /k/, /b/, /d/, /g/, with a few exceptions of fricative /s/, making a basic set for training. The finding concerning transfer to new contexts appears to contradict a claim raised in Speech Learning Model (Flege, 1995) about the position-sensitive nature of L2 speech learning. Flege suggested that L1 and L2 sounds are perceptually related at a position-sensitive allophonic level, thus the phonetic environments in which a vowel is positioned may influence the learning of itself. Training to identify some non-native vowel pairs by using stimuli with variable but controlled stimuli could still benefit the learning. However, since individual differences among the subjects were not considered and given the individual difference observed among them (see Appendix I for individual performance), it is not fair to conclude that the current findings totally contradict Flege’s claim; rather, a general effect of position-sensitivity may be present because there were subjects who still found difficulties in transferring learning to the new contexts.

5.5.2 Production Aspect

5.5.2.1 Contextualization Effects

The effects of contextualization of the production of the three vowel pairs were very similar to the generalization pattern at the perceptual level: it is the HVPT approach that brought more contextualization success to subjects than those who were trained under other paradigms. Production training was also bringing positive transfer of production learning in the word level to the sentence level. The findings suggested that the use of HVPT benefited the subjects more in production learning as their accuracy in production was not only retained at the word level, but was extended to the sentence level. Direct and explicit production training also had its place in the facilitation of learning.

5.5.3 Summary

Detailed results concerning the effects of generalization in both the perceptual and production aspects have been discussed previously in each section above. This section only serves as a summary of findings. Positive generalization effects to novel words and speakers were found among subjects trained under HVPT but not in LVPT; an addition of production training did not add extra credits to the paradigm. Yet, subjects who were trained solely in the production domain still benefited moderately in terms of generalization of learning to new voices. The results of generalization, as opposed to the posttest results, provide more solid and genuine support to the efficacy of adopting highly variable perceptual stimuli to training L2 learners in the identification of difficult vowel contrasts. Also, the fact that learning can be extended to sentence-level is promising as well.

5.6 THE EFFECT OF VOWELS – RESEARCH QUESTION 7

The last research question to be answered is about the learning pattern observed when comparing both the perceptual and production performance of the subjects in each vowel pair. Disparities are shown in the degree of learning across the three vowel pairs, signifying that the effectiveness of the training approaches may have more effects on learning particular vowel pairs but not the others. Or, it simply means that some vowel pairs are inherently more learnable. The differences in the learning patterns are going to be discussed through answering the following Research Question:

RQ7. Are there any differences in the ease of
a) perceptual identification or
b) production of the three English vowel pairs /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/?

5.6.1 DIFFERENCES IN THE VOWEL LEARNING PATTERN: THE WEIGHT OF SPECTRAL CUES VS. DURATIONAL CUES IN CANTONESE

Recall that the discussions of Research Questions 1 to 6 were explicated in terms of the subjects' performance of the three vowel pairs in general, but not individually. Surveying through differences in the perceptual performance of each vowel pair gives an impression that all the subjects in general had more difficulties in perceiving the back vowel pair /ʊ/-/uː/. From the pretest to posttest, the perception improvement of both the front vowel pair /ɪ/-/iː/ and the low vowel pair /e/-/æ/ were usually more significantly identified than the back vowel pair when the performance of different paradigms were compared. An even more interesting pattern of the perceptual performance was observed when comparing the production training subjects' pretest and TGs results: the low vowel pair became consistently and significantly more accurately identified than both the front and back vowel pairs. Readers can refer back to previous summary tables for a more thorough understanding of the effect of vowel and its interactions with other factors.

The production pattern, when the pretest and posttest data were considered, is similar to the perceptual pattern as stated: the subjects in general had greater production accuracy in the front and low vowel pairs, whereas the back vowel pair remained as the most difficult contrast even after training. The fact that the production accuracy rates of the back vowel pair stayed at 63.01% (averaged across all the groups that received training) in the posttest simply indicated that this pair remained difficult when compared to the other two: all the trained groups' posttest production of the front vowel pair had an average of 75.49% in the posttest, having an increase of 26.48% from the pretest. Their performance of the low vowel pair was 73.54% in the posttest, having an increase of 23.78% from the pretest. Acoustic analysis also showed that the F1 and F2 values of the back vowel pairs produced by both groups, though have begun to show signs of separating, still remained more overlapped than the other two pairs even after training. However, the 18.52% of increase in the accuracy of the back vowel production still suggests that training can benefit the subjects in more accurate production.

The finding that the front and the low vowel pairs were consistently better identified as well as more accurately produced may not be accounted for by simply comparing the differences in L1 and L2 vowel systems. Even though Hong Kong Cantonese has phonemic /i/, /ɛ/ and /u/ as well as allophonic [ɪ], [e] and [ʊ] when in closed syllables and the present study did not directly measure the assimilation patterns of the three vowel pairs, the heavy reliance of choosing /iː/, /e/ and /uː/ when disambiguating the front, low and back vowel pairs respectively in both perception and production pretests hinted that the L2 learners might tend to assimilate the English vowel pairs /ɪ/-/iː/ into Cantonese /i/; English /e/-/æ/ into Cantonese /ɛ/ and English /ʊ/-/uː/ into Cantonese /u/ which were the closest to the English vowel pairs. The assimilation pattern (i.e. Single-Category), according to the Perceptual Assimilation Model (PAM) by Best (1995), was similar when it was in the pretest. Yet, the observation that the low vowel being the most learnable, followed by the front vowel pair and lastly the back vowel pair, may be attributed to the spectral cues, rather than the durational cues, that Cantonese speakers rely on.

The two English vowel pairs, /ɪ/-/iː/ and /ʊ/-/uː/ distinguish themselves within the two counterparts not just only in the vowel quality, but also the durational differences. The subjects in the present study appeared to take advantage of durational differences, which was claimed by Bohn (1995) as a universal principle, to discern vowel pairs. However, due to the shortening phenomenon of the tokens with long vowels /iː/ and /uː/ in front of a voiceless sound, which then became relatively shorter and similar to the short vowels /ɪ/ and /ʊ/, getting a consistent learning by relying on durational cues solely became difficult because one-fourth of the tokens had the long vowels placed in front of a voiceless consonant in the perception training, creating confusion to the subjects. The fact that the subjects in general produced all /iː/ and /uː/ with longer vowel lengths than the target means that they did not consider the effect of the phonetic environments and this confirms the claim that the subjects, even though trying to rely on durational cues to discern the vowels, may have been unexpectedly disrupted by the token distribution, hence resulting in relatively lower results than the low vowel pair. However, due to the low functional load[18] of duration in Cantonese (/a/-/ɐ/ was the only contrastive phoneme, other vowels have no short/long phonemic counterparts) and the production performance shown by the F1-F2 space plots in which the subjects have begun to separate the vowels in terms of spectral differences but not just durational differences, it is hinted that even though durational cues may have only mild effects in the subjects, but spectral cues may have a greater role in their learning.

In fact, Zhang, Peng and Wang (2011) discovered that Cantonese speakers were affected more by vowel quality cues than durational cues when they perceived Cantonese vowels. Their experiment found that vowel quality cues were used as linguistic cues that are language-specific in Cantonese. Duration alone may not be an efficient cue in categorization; rather, vowel quality cues overrode the effect of duration. Previous studies (e.g. Altenberg & Vago, 1983; Bohn, 1995; Cebrian, 2002; Escudero, 2001, 2005; Flege & Bohn, 1989; Flge, Bohn, & Jung, 1997; Kondaurova & Francis,

[18] Functional load "is a measure of the work which two phonemes (or a distinctive feature) do in keeping utterances apart" (King, 1967, p.831). It measures the number of lexical items in a language that make distinctions by certain features. It is usually measured by calculating the number of minimal pairs of the feature. Phonological contrasts that have high functional load will pose more difficulties for a listener to guess the segment of feature if the speakers do not make the differences in the contrast.

2004, 2006; Minnik-Fox & Maeda, 1999; Morrison, 2006; Munro, 1993; Wang, 2002; Wang & Munro, 1999) investigating the primary cues that L2 learners with different L1 backgrounds relied on when discriminating some vowels were mostly vowel duration. Hong Kong Cantonese speakers may therefore be a group of speakers who are more exceptional in the way that they could transfer the more reliance on spectral cues in their L1 to the learning of L2. The better results found in the perception and production accuracy in the low vowel pair /e/-/æ/ have lent support to this idea.

Even though some previous studies (e.g. Hillenbrand & Clark, 2000; Morrison, 2006; Wang, 2002) reported that the low vowel pair /e/-/æ/ is more confusing and thus posing more learning resistant to L2 learners, the present data suggests that this is not the case for the population in the present study. Instead, it is the vowel pair which had the highest accuracy rates in both perception and production data in general. I speculated that the success in perception of the mid vowel pair is possibly due to the fact that the vowels /e/ and /æ/ display more salient and seeable differences in terms of their articulation. They display more articulatory differences than the other two vowel pairs. Their differences are not only easier to be observed than the other two vowel pairs during the training, but also easier to be heard as the difference in the mouth openness, which contributed important quality cues to the perceiver, is relatively larger than the other two vowel pairs. The perceptual finding that only this low vowel pair improved in the production test has given some support to this claim. Also, even though /e/ and /æ/ also have differences in the vowel length and were affected similarly under the token distribution as stated above, the fact the these two vowels differ more in the articulatory aspect may then become more salient than the front and back vowel pair, hence resulting in better perceptual and productive improvement. It is thus speculated that, the subjects in the present study could make use of the quality cues that they rely primarily on in their L1 when learning the three target vowel pairs because Cantonese is a specific language which is different from other languages where durational differences were usually their primary cues. Although the low vowel pair was found to be the most difficult in previous studies, the finding that this, being the most difficult pair, received the most important and significant improvement also highlights the pedagogical values of the training paradigms. The L1 background of the learners is also very fundamental when designing suitable training paradigms.

5.6.2 SUMMARY

A preliminary understanding of the L2 vowel learning pattern as shown by the three target vowel pairs, /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/, has indicated that the back vowel pair, at the aggregate level, is in general and relatively more difficult to learn. However, the present finding is not conclusive; it does not mean that this pair is not learnable; it is only posing relatively more difficulty to the subjects in the present investigation, as shown by their perceptual and production improvement. The fact that a lot of individual differences among subjects are observed also points out that the learning pattern varies across learners and should not be overlooked. Moreover, the subjects' significant improvement in the perception of the three vowel contrasts appears to support the Speech Learning Model that perceptual categories can be established once the differences between the two vowels are noticed and perceived. Laboratory training paradigms were useful, though to different extents, in modifying the perceptual categories and (by association) production mastery.

5.7 Theoretical Implications

In previous sections, I provided a detailed explication of the results obtained and answered the seven research questions raised in the present study. Before ending the chapter, I will turn to discuss the findings in the theoretical aspect by relating them to previous empirical research and speech learning models that are crucial for a more global understanding of L2 speech learning.

5.7.1 Acquisition of Second Language Vowels in General

Even though the learners cannot achieve native-like L2 perception or production accuracy, it is proposed here that the learning observed in the three contrasts may have constituted a developmental sequence of L2 vowel learning. At the initial stage, L2 learners had immense difficulties in distinguishing, both in perception and production, two non-native contrasts and this was the *no-contrast stage*. During this stage they may assimilate the L2 categories into existing L1 categories, depending on the similarities and differences between the L1 and L2. Later on, L2 learners began learning to use different cues to identify the contrasts during training. They relied on some cues that are primary in the L1 but secondary in the L2 to discriminate the contrasts. This *developing stage* may take the longest time and how the learning was administered may influence the effectiveness and the degree of success of learning. Given training stimuli with higher variability, learners could strive to find out the specific cues that are more relevant in discerning the vowels. Lastly, the learners could develop the ability to utilize the primary cues exclusively when native speakers are perceiving the contrasts, and this is the *native-like stage*. This stage may take over years to achieve. The present study is however not a longitudinal one and the developmental sequence was only inferred from the cross-sectional patterns observed in the data, although it has been claimed as plausible by some researchers (e.g. Escudero, 2006; Polka & Werker, 1994; Weker, Gilbert, Humphrey, & Tees, 1981; Wode, 1976). The order of the development sequence was also not proved, and future research can focus on longitudinal and long-term development so as to confirm the order of learning of L2 vowels.

5.7.2 IMPLICATIONS FOR L2 SPEECH LEARNING MODELS

Although the present study does not aim to test any perception or production models or theories, the current results still lend some support to the L2 speech learning models that were discussed in Chapter 2.

The Speech Learning Model (SLM; Flege, 1991, 1995b, 1999, 2002, 2003) proposes that when the differences between L1 and L2 sounds or between two L2 sounds are perceived, L2 phonological categories can be established. Also, L2 vowels which are more unlike any L1 vowel will be an advantage for the learners to establish robust categories than L2 vowels that are similar to any in L1. Thus, according to the SLM, since the three vowel pairs have the corresponding Cantonese counterparts, the perception and identification of the vowel pairs was difficult. Moreover, the participants showed significant improvement after perceptual training implied that the perceptual categories, although not yet completely established, had started to be established among them when they started to be able to perceive the differences between the target vowel contrasts. This is further supported by the results in acoustic analysis that the subjects started to shift from attending to duration to relying on the crucial spectral cues for distinguishing the vowel contrasts, hence modifying the phonetic categories.

Similar to the SLM, the Perceptual Assimilation Model (PAM; Best, 1994, 1995; Best et al., 1988; Best et al., 2001; Best & Strange, 1992; Best & Tyler, 2007) accounts for cross-language perception by postulating that L2 learners classify L2 sound contrasts into different phonetic categories according to the similarity of the two sounds and the perceived difference between the L1 and L2 sounds. It also posits four assimilation patterns for two L2 sounds to the learners' L1 system: Two Category Assimilation (TC), Category Goodness Assimilation (CG), Single Category Assimilation (SC) and Non-Assimilation (NA). This model predicts that learners' successful performance in distinguishing different L2 sounds will be in the order: TC > CG > SC. The hypotheses of the PAM have been given in Chapter 2. The Second Language Linguistic Perception model (L2LP) (Escudero, 2005) incorporated the approach towards the assimilation pattern in the PAM and posited similarly some scenarios in which the learners will encounter when they perceive the target language sounds: a *new scenario*, a *similar scenario* and a *subset scenario* (detailed descriptions of these scenario are given in Chapter 2). The model proposes that the *new scenario* is the most difficult situation.

The current findings suggest that the subjects did not have a clear category distinction for the three vowel pair contrasts. The subjects' accurate identification at the perception pretest was 70.06% for /ɪ/, 69.85% for /iː/; 59.87% for /e/, 57.55% for /æ/; 65.18% for /ʊ/ and 69.65% for /uː/.The pattern of identification of all three vowel pairs may be due to an SC assimilation pattern according to the PAM, or it is the *new scenario* in the L2LP. The subjects may have assimilated the two categories in the front, low and back contrasts to a corresponding Cantonese counterpart, /i/, /ɛ/ and /u/ respectively. The SC contrast is predicted to be difficult since the two sounds are similar to the L1 category; whereas the new scenario is difficult in the way that the L2 learner will assimilate the two L2 categories into one single L1 category. It has been shown in some previous research (e.g. Hung, 2000; Wang 2012; Wong, 2010, 2012, 2013a, 2013b) and the present pretest results that the perceptual confusion of the three vowel pairs existed among Cantonese-speaking subjects. Although the aim of the present study was not to confirm the assimilation pattern by the models and that the exact cues that the subjects relied on have not been tested, the current interpretation is only deemed to be speculative; future research may assess the assimilation pattern by measuring the mapping of L2 segments onto the L1 system.

5.8 Summary of the Chapter

All the research questions in this study have been responded and elucidated through surveying statistical results and the methodological components. With reference to the seven research questions, the results in the perception and production tests have shown some replications that support previous research findings whereas some novel findings which appear to be different from previous studies were also found. The findings are all illuminating, though inconclusive, and have shed some light on how these training paradigms can be useful in language classrooms and how L2 speech perception and production can be investigated.

Based on the above accounts, the next chapter will focus on the discussion of the pedagogical and research implications. Together with the illustration of limitations and contributions of the studies, this thesis will end by shedding light onto future research possibilities in the next chapter.

CHAPTER 6

PEDAGOGICAL IMPLICATIONS, LIMITATIONS & FUTURE DIRECTIONS

6.0 INTRODUCTION

The promising outcomes obtained in the current study as presented in the previous two chapters have provided insightful ideas on the possible modification of perception and production of non-native vowel contrast, even for some difficult cases for Hong Kong Cantonese ESL learners, and appear to contribute to L2 speech teaching and learning in particular. This final chapter will therefore conclude the whole thesis by revisiting the research questions and key findings, stating the contributions and establishing future directions with reference to the findings and limitations of the results obtained. At the outset of this chapter, a general summary of the research will be given. In Section 6.2, contributions of the current research will be discussed, followed by pointing out the limitations of the study in Section 6.3. Potential pedagogical implications will be briefly described in Section 6.4. Finally, this chapter and the whole thesis will come to a close in Section 6.5 in which conceivable future research agendas in the field of speech perception and production as well as L2 phonetic training will be offered.

6.1 An Overview of the Current Study

The major goal of the present study was to compare and contrast the efficacy of several training paradigms in the modification of the perception and production of three English vowel pairs, /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/, among Cantonese ESL learners. Two purely perceptual training approaches, High Variability Phonetic Training (HVPT) and Low Variability Phonetic Training (LVPT) which differ in the adoption of training stimuli variability, were compared with an explicit production training paradigm. The addition of the production training on top of either one of the perceptual training was also a target of comparison. The generalizability of the training effects in both the perception and production aspects was also an interest in this research. Examining the effects of the perceptual training intensity and the nature of vowels on the subjects' learning also added more insights to discover underlying factors which influence L2 speech learning.

As discussed in detail in the previous chapter, the results of the present work were in general positive and compelling. Both HVPT and LVPT approaches were effective in modifying the subjects' perception and production of the three English target vowel pairs. For perceptual learning, the subjects can generally gain from both training approaches, as indicated by the significant increase of identification scores from the pretest to the posttest; yet no significant differences were found between HVPT and LVPT. Production training also benefited perceptual learning in general and particularly just on the low vowel pair, but it was a less effective approach vis-à-vis HVPT or LVPT. The production training supplementing the perceptual training paradigm appeared to show no extra help to be more accurate in discerning the vowel contrasts, and neither was perceptual training intensity showing any effect on the subjects' perceptual performance.

A broader picture can be seen only when the results in the Tests of Generalization (TG), which tested whether perceptual learning can be generalized to new words or new speakers, are also taken into account. Results in the three TGs revealed that subjects following HVPT generalized significantly better than LVPT, although both subjects receiving either HVPT or LVPT demonstrated significant generalization. Production training also contributed to generalization, and again significantly on the low vowel pair; however, it is the least favourable paradigm to adopt since it displays the least learning transferred to generalization. A consistent pattern of HVPT (and those who were given an addition of production training) being a more useful paradigm across the performance in the TGs than LVPT (and those who were given an addition of production training) as well as those who received solely the production training indicates the importance of stimulus variability in successful training of the perception of non-native contrasts which can be generalized to new words and new voices. Although the group receiving intensive HVPT training had significantly better performance in the TGs than the group with standard HVPT training, no definite conclusion can be drawn at this stage since the effect of intensity was slight and was just found in the group receiving HVPT; those receiving LVPT showed no difference in their generalization performance. All in all, perception training, particularly HVPT, contributes directly to perception learning and generalization; production training appears to have some moderate influence. The perceptual training intensity was not playing a prominent role either. Supplementing production training to HVPT/LVPT did not appear to be superior to groups receiving simply HVPT/LVPT.

Although previous research (e.g. Wang, 2002) testing whether perceptual learning in vowel contrasts can be transferred to production showed inconsistent results, the present training study reported that improvement in perception of the subjects trained under HVPT and LVPT was accompanied by the significant increase of the percentage of target production of the three vowel pairs. The HVPT subjects also outperformed the LVPT group in the posttest. Still, perceptual training intensity was not at all influential to the production performance of the subjects' production performance in the posttest. Direct production learning could be observed for those who received production training as well; but only the group trained under HVPT with production training stood out amongst the other training types. A similar picture was observed when the results in Test of Contextualization were considered. Production accuracy at the sentence level of the HVPT group was significantly better than LVPT and the control; but LVPT was not superior to the control. Production training contributed directly to the production accuracy at the sentence level, whereas the group trained under HVPT and production training (HP type) performed the best. Again, perceptual training intensity displayed no influence on the production accuracy at the sentence level.

From both the perception and production data, the perception and production of the front vowel pair /ɪ/-/iː/ and particularly the low /e/-/æ/ vowel pair were more malleable and susceptible to change since improvement was more evident; whereas the back vowel pair /ʊ/-/uː/, although improvement in the perception and production of which have been seen, it remains as a difficulty – being far short of native-like ability.

The present research replicated the results of a number of previous studies (e.g. Bradlow et al., 1997; Lambacher et al., 2005; Perrachione et al., 2011; Wong, 2010; 2012, 2013) that perceptual training can exert a significant effect on perceptual learning and some degree of transfer to production of non-native contrast which poses great difficulties among the subjects. Even though native-like performance could not be attained after training (and was not the expectation), the subjects' robust improvement and high generalization effects in both perception and production of the vowel contrast suggest phonetic approaches which include identification tasks and immediate feedback are effective. Most importantly, high stimulus variability in the training was attested in the present study to be a beneficial element that leads to a significant level of success of non-native contrast acquisition. The findings also imply that there exists a relationship between perception and production. Learning solely in the perceptual aspect can be transferred to the production domain even no explicit production instructions were given. However, production training in which a certain amount of perception instruction is unavoidable, only contributed very slight effect on the successfully modification of the perception of the three target vowel pairs. It appears to support the claim that it is perception which precedes production; the slight gain in the perception accuracy by the subjects trained under production training might be attributed to the very limited amount of perceptual exposure to the native stimuli that was given during the production training. This account, however, will remain inconclusive until further research can provide empirical evidence, although this is not in line with the results of the very scarce research works (e.g. Leather, 1997) showing that production can be transferred to perception.

6.2 Contributions of the Current Study

Relatively few studies (e.g. Iverson & Evans, 2007; 2009; Lambacher et al., 2005; Lengeris, 2009; Nishi & Kewley-Port, 2007b, 2008; Wang, 2002; Wong, 2010; 2012) in the perceptual training literature have aimed to explore the perception and/or the production of L2 non-native vowel contrasts. Most perceptual training studies (e.g. Bradlow et al., 1997; Bradlow et al., 1999; Lively et al., 1993; Lively et al., 1994; Logan et al., 1991) have focused on non-native consonantal contrasts, among which studies on the English /ɹ/-/l/ contrast by Japanese speakers were the most extensively investigated. Even for those who investigated vowel contrasts, they mainly tested Japanese (Lambacher et al., 2005; Nishi & Kewley-Port, 2007b) or Korean speakers (Nishi & Kewley-Port, 2008). The current study has offered insightful empirical data for more understanding in L2 vowel training. Moreover, with the exception of Wang (2002) who trained Mandarin and Cantonese speakers' perception of English vowel contrasts and several studies conducted by me (Wong, 2010, 2012, 2013a, 2013b), no other vowel training studies examined the perception and production of the English vowel contrasts /ɪ/-/iː/, /e/-/æ/ and /ʊ/-/uː/ among Cantonese ESL learners. The present research certainly offers enlightening findings that enrich relevant perceptual training studies with subjects having Cantonese as their L1.

A substantial amount of research investigating L2 speech learning by Cantonese ESL learners focused on the production domain. Even for those which are concerned about perceptual problems that the learners encountered, the works were mainly stating the phenomenon or giving suggestion of how pronunciation teaching could help the learners but without providing empirical support. The present study, on one hand, verifies the observations stated in previous studies about the perception and/or production problems of Hong Kong Cantonese ESL learner, on the other, provides promising ways to remedy the confusion patterns. The endeavors that the researcher has put into comparing and testing the efficiency of different training paradigms and design have contributed to the field of L2 phonological teaching and learning in particular.

Previous non-native contrast training studies have deployed generalization tests to measure the external validity of the learning. Almost all these generalization tests gauged only the subjects' perceptual aspects, without any consideration in the production domain, such as producing new words with the target words, or whether the production learning can be retained at the sentence level during passage reading or in casual speech. The adoption of the Test of Contextualization (TC) in the current study thus gives more insights on how transfer of learning to the production aspect can be elaborated and be more practically utilized in natural speech. TC as a way to test the generalization effects in the production domain has been implemented for the same reason in my previous works (Wong, 2010, 2012, 2013a). However, feedback from the subjects in those studies reflected that the passage had too many rhymes due to the sole focus on the target vowel pair /e/-/æ/, making it too much of a tongue-twister and therefore difficult to read naturally and free of error or repairs without extra efforts devoted. This error-inducing drawback originated by the nature of the passage has been avoided in the present study. By using a reading comprehension passage for Grade 5 students that incorporated all the words with the three target vowel pairs more naturally (see Appendix F for the whole passage), I attempted to avoid drawing too much cognitive load from the subjects during TC. Consequently, the adoption of TC in this study improved the previous methodology and offered more insights and possibilities to test to generalization in the production domain.

Another contribution of the present study was the exploration of efficient training design by contrasting directly HVPT and LVPT paradigms, as well as comparing other phonetic training approaches and the effects of a combination of them. Previous works (e.g. Bradlow et al., 1997; Lambacher et al., 2005; Wang, 2002) may have examined the role of stimulus variability in learning new perceptual categories, but they only claimed that HVPT was effective in both perceptual learning and production improvement of some segments, without making a compulsory comparison to its counterpart, LVPT. That high stimulus variability is important to the perceptual and productive improvement in non-native contrasts is usually assumed or deduced in these studies by simply referring to the positive results obtained or previous research that examined only the LVPT. Except my previous works (Wong, 2010, 2012, 2013a, 2013b) and a more recent work by Perrachione et al. (2011) which attempted to compare HVPT directly with LVPT, the need of comparison between the two paradigms, or two or more paradigms as the present study does, has been overlooked. It is highly suggested that future training studies adopting HVPT should incorporate a comparison with LVPT as a control, which serves as a threshold before drawing a more solid conclusion on whether stimulus variability plays a role in successful perceptual training like the present study and by how much. The present study which attempted to compare more than two training designs certainly adds more value to the present training study literature.

6.3 Pedagogical Implications

The primary focus of the present study was pedagogical in nature: it aimed to compare and contrast several phonetic training paradigms on the effectiveness in modifying some perennial perception and/or production problems in L2 contrasts among Hong Kong Cantonese ESL learners. Based on the results obtained from the experiment, implications that are central to facilitating L2 teaching and learning will be presented in this section.

Useful pedagogical information was obtained from the positive results in the study which indicated that there exists the possibility of successfully training ESL learners to perceiving and producing some non-native contrasts more accurately. It appears that no matter which type of training was delivered, learning occurred. What matters more was which type of training would be more beneficial to the learners. The present investigation suggests that offering training stimuli with higher variability will benefit the subjects more in terms of perceptual learning, production performance as well as generalization in both domains. While maintaining the viability of application and simplicity of procedures of the training was also of high importance for teachers, HVPT is practical enough to be adopted easily in schools or learning centers since it is not a complicated design. This paradigm can even be setup online for learners who are willing to receive training at their own pace and time. Efficient as it is, HVPT, as a purely perceptual training paradigm, not only improves learner's perception abilities but also production of even the most difficult contrasts. If the high variability of training stimuli is still difficult to obtain, the simpler paradigm LVPT or merely explicit production training may still help a learner to some extent. Perrachione et al. (2011) explicated in their study that learners with low perceptual aptitude may benefit more from LVPT than HVPT (HVPT may even be detrimental to their learning). Although this study has not compared directly learners with different proficiency levels or perceptual abilities, the compelling results still suggested the benefits of perceptual training paradigms.

What may be even more attractive to L2 teachers and learners concerning the training methodology was the observation that perceptual training intensity, be it intensive or standard, demonstrated no significant differences in affecting the learning in discerning or producing the contrasts accurately. The major problem lies in whether training was given and in what form. This result may be promising to L2 teachers and learners: given the same amount of exposure, time and frequency in training may not be the most influencing factor on successful learning of the non-native contrast. Once HVPT or LVPT was offered, learners could benefit from the training, regardless of whether the training was received within a short or long period of time. One thing that should be noted, however, is that the amount of training sessions might be important to ascertain successful learning. Although the current study did not investigate this factor, the present results show that if a learner is given a daily exposure to 80 training stimuli per contrast a day for 10 days or training with 400 training stimuli per contrast a day for two days, he can still learn the given contrast successfully. Based on the results of the present study, it is suggested that the amount of training stimulus received for successful perception and production learning is 800 stimuli per contrast. Teachers and learners can pay attention to this and be more optimistic that these L2 segmental confusions can be solved with training without being constrained by time.

Interesting as it may seem, that production training per se did not contribute much in perception learning (although production learning was rather evident) may imply some important pedagogical information. The results show that production training only benefited the subjects' perception of the low vowel pair in both perception posttest and in generalization tests, but not for the other two vowel pairs. Yet, for the production aspect, the subjects improved in all three vowel pairs. The observations have been justified in Chapter 5 that the differences in learning were attributed to the articulation differences of the three vowel pairs, the phonetic environments of the tokens and the acquisition pattern in general. The application of the training paradigm is therefore important, because production training, which is prevalent in common L2 classrooms, may only have very limited amount of practicality to the amelioration of the perception and production of confusing vowel contrasts. Offering the right sort of task to subjects to assure successful learning on a particular vowel contrast may therefore be an important caution that L2 teachers have to take. Obviously, if subjects can solidly improve in the perceptual aspect, i.e. in a basic set and to new stimuli, it can almost be guaranteed that production learning occurs. It further hints that HVPT paradigm may suffice for the learning, though an addition of production training on HVPT (but not on LVPT, as shown in the results) may further benefit the subjects. However, if only the production aspect is concerned, production training is still a very useful and easy-to-adopt paradigm in language classroom. Teachers only have to make sure to allow explicit instructions, feedbacks, the use of audio-visual materials to facilitate students' learning. An addition of using audio-recording software to trace the students' performance may also be one useful means for better production learning (McCrocklin, 2012).

Lastly, raising the intelligibility of the L2 speaker, but not ridding the detectable accent, should be the first priority in a language classroom or in the design of instructional proposals. Previous research (e.g. Derwing & Munro, 1997; Munro & Derwing, 1999) showing that only a moderate correlation exists between accent and intelligibility suggests that pedagogical means should not be motivated by the hope to modify the accent of an L2 speaker. Even when a speech is strongly accented, it can still be highly intelligible to native and non-native listeners of the language (Munro & Derwing, 1995). Although the present study demonstrates positive results of modifying the perception and production of three English vowel pairs among the ESL learners, to reach native-like competence and accent certainly remains a long-term ultimate goal that requires time, effort, experience or even elements that have not explored in the field.

Given the tension in the time of instruction in a language classroom, time to spend on vowel perception or production training of the similar kind should only be allotted accordingly to contrasts that may result in serious communication breakdown due to a lack of intelligibility. The vowel pairs chosen in the current study, particularly /ɪ/-/iː/ and /e/-/æ/, were found to have caused miscommunication problems between an L2 speaker (producer) with the native speaker (perceiver) in general (Brown, 1991; Jenkins, 2000; Sewell, 2009), and are thus worth training. Nevertheless, the picture of whether a contrast should be trained is not simply about intelligibility. Brown (1995) argued that different minimal pairs have different functional load, hence they carry different communicative weight. Munro and Derwing's findings (2006) also indicated that errors in phonemes with high functional load affected comprehensibility more than those with low functional load. According to the statistics by Higgins (2013), it appears that both /ɪ/-/iː/ (number of minimal pairs: 466) and /e/-/æ/ (number of minimal pairs: 305) have relatively higher functional load than /ʊ/-/uː/ (number of minimal pairs: 18). It is thus important to train the front and low vowel pairs to maintain the intelligibility and resolve the difficulties which arose due to the high functional load. For the back vowel pair, however, communication difficulty that may be caused by it, based on the functional load, may be relatively remote. The consistent difficulty that is observed among the subjects in developing more native-like perception and production of the back vowel pair might also be due to the extremely scarce instances that can be found for /ʊ/-/uː/. Thus, priority may have to be given to the two pairs or others during training so as to maximize the practicality in learning and optimizing intelligibility.

6.4 LIMITATIONS OF THE CURRENT STUDY

Despite all the efforts that have been made to ascertain the quality and validity of the present research, there are still some limitations that are worth paying attention to so that corresponding modifications can be given to contribute to future research.

First, the sample size in each group was not large enough for a simultaneous comparison of all the factors such as using 5-way ANOVAs, and thus, they were investigated one by one. Even though the problem may not be that serious due to similar results obtained in the omnibus 5-way and 3-way ANOVAs meaning that each variable still has its effects on the subjects when all the variables are considered at a time, the interpretation was made less straightforward and interactions due to the confounding nature of one factor with others were not looked at. Conducting several experiments with fewer variables but more subjects within a group may avoid the problems that may arise in such multiple comparisons. Likewise, the unequal number of subjects in each group cannot offer the most ideal case for a well-controlled and balanced laboratory training experiment either, and thus lowered the fairness of evaluation of the training effects. The above situations, though unavoidable in this study, are attributed to the fact that some subjects had to withdraw from the study due to personal reasons and clash of schedule, as mentioned in Chapter 3. Recruiting more subjects and having a greater control of the variables and grouping would be more desirable. The absence of a counterbalanced design of tests might also have elicited false responses due to other factors. The order of treatment in this repeated measures design should be considered to avoid any pitfalls that might have arose due to other factors or the order effect *per se*.

Nonetheless, a complete counterbalancing design is impractical for the present experimental design as it had multiple conditions. Also, the number of participants in the present study would not suffice for such a design. Some incomplete counterbalancing design such as the Latin Square design might be a viable compromise to reduce the carryover effects and strike a balance between validity and practicality.

Another limitation with regard to the subjects was that some individual differences in performance across tasks were observed within subjects, even if they have similar profiles. As well-documented in previous L2 training studies (e.g. Bradlow et al., 1997; Bradlow et al., 1999; Hazan et al., 2005; Lengeris, 2007), highly variable performance were prevalent before and after training. Although the major focus of the present study is not explaining or investigating individual variability, this observation should not be overlooked as it suggests the existence of idiosyncrasies in the efficacy of the training paradigms since all the outcomes and findings were based on the averaged means in a group. Still, general consistencies among subjects were observed across tasks, supporting the present results.

The production data was based heavily on the transcriptions completed by me and two more researchers to ensure a genuine reflection of the subjects' performance. The Cronbach's alpha was high for the different transcription scripts and over 90% of intra- and inter-rater reliability was observed. Still, given that all the transcribers are non-native speakers of English, it is possible that high intra- and inter-rater reliability was due to the fact that both transcribers were consistently influenced by their Cantonese phonology, which may make the transcriptions deem invalid. Nevertheless, a number of productions (two subjects' productions from each of the 10 groups) were evaluated by a native English researcher, confirming the accuracy of the transcriptions. The fact that the transcriptions and the acoustic analysis align with each other further supported that that the present transcriptions were convincing. A number of differences spotted between the results in the transcription with the acoustic analysis were only due to the fact that the researcher, the inter-rater and the native speaking judge, given limited resources, could only transcribe the data with an assumption that the participants were producing English vowels. The transcribers could not identify whether a vowel /i/ was a production of a Cantonese /i/ or and an English /iː/. The acoustic analysis showed hints that some productions of the participants were closer to the Cantonese vowel rather than the English one, even though most of them aligned with the transcription results. Future studies should include experiments that can compare the productions by monolingual speakers, bilingual speakers and L2 speakers and check against their similarities and differences. Also, future research should recruit native speakers to make a more global preference rating on all the productions.

Finally, no task was given to test the long-term retention of perceptual or productive learning in the present study. This was due to reasons of practicality since retention testing would have required the subjects to complete another battery of tests after a period of time. They were already under the pressure of an imminent public examination for the university admission, and were no longer available after completing the first batch of tests. Nonetheless, the ultimate aim of researching the effectiveness of any phonetic training paradigm was for the long-term benefits of language teachers, and learners in particular. It would be useless if a training of this kind can offer immediate but only short-term and surface help in modifying the perception and production of the difficult segmental contrasts. No solid conclusion can be drawn at this stage as to whether the training paradigm is effective in the long run, although previous studies which showed generalization also demonstrated retention in learning, which is also expected in the present investigation given the robustness of training and positive results obtained (e.g. generalization to new word and new talker; generalization to sentence level).

6.5 Future Research Directions and Implications: A Research Agenda

Building solidly upon the limitations and observations raised in the above sections, future studies can extend or rectify the present research design with a view to contributing further to the research field.

Besides investigating only the three English vowel pairs, future studies may complement this by training subjects on other segmental features (e.g. consonant clusters or other vowel contrasts) and suprasegmental (e.g. stress, rhythm, intonation) ones which also pose difficulties to the learners. Particularly for Hong Kong Cantonese ESL learners, training in the suprasegmental aspect may be very worthwhile for them because it is usually the suprasegmental elements that have greater impact on the intelligibility, as shown in some empirical findings (Field, 2005; Hahn, 2004; McNerny & Mendelsohn, 1992). Gender, ages or L2 experience, etc. will also be worth investigating as it will expand the literature and further verify the external validity of the training programs. This can also shed more light on the learning characteristics of different populations. Recruiting subjects with other L1s and L2s or language backgrounds can be enlightening as these types of cross-language perception or training studies can provide more information about L2 perception and production in general. It is further hoped that once more empirical evidence can be obtained, the relationship between the perception and production can be known and further consolidate previous theoretical premise of how the two domains are connected and interact with each other.

No long term retention task was adopted in the present study to gauge whether perceptual and/or productive learning can be maintained through a period of time. This methodological step is important to the training as well as L2 teaching and learning field as it provides information on whether a particular paradigm can bring permanent or simply long-term learning effect to the learners. As discussed in the last chapter, even though HVPT appears to be an efficient paradigm, it only makes the subjects' category process become more efficient. It only shows short-term learning effect but may not give a full simulation for the changes that take place in perception during longer term L2 language learning. Previous studies (e.g. Escudero & Boersma, 2004; Iverson & Evans, 2007) have shown that long-term exposure to phonetic categories will change the cues they use. This certainly requires further longitudinal studies to gain more insight into L2 learning mechanisms. Moreover, even though a generalization test in the production aspect – Text of Contextualization – has been adopted in the present study, it will give more information if the task can give data that is even more natural and conversational. Tasks like extemporaneous picture description task or interview can strengthen the claim of whether or to what extent learning can be transferred.

In terms of methodology, future studies can assess the learning of the subjects by adopting and comparing different assessment tasks such as identification test, category discrimination test, cue weighting test, etc. which all tap into different aspects of L2 vowel perception and production. Previous studies (e.g. Iverson et al., 2011) have shown that the learning degree revealed in these tasks may differ from each other because different tasks may measure different L2 learning processes. This will certainly shed more light on the effectiveness of the training paradigms. Likewise, the observation that production training only slightly improved perception but, conversely, perception training greatly helped production, plus the finding that a perception and production training (modality trained alternatively, see Figure 3.2 for reference) paradigm showed weaker learning in perception than perception-alone approach, all suggest that a further investigation is needed before claiming the relationship between perception and production. One way to examine this view is by scheduling training in perception before production and comparing the results of training in production first before perception. The order of training may give additional information on the relationship of the two modalities.

One major focus of the present study was to investigate the efficacy of HVPT, a purely perception training technique, in modifying non-native contrasts. It will be interesting to understand or modify the notion of “high” – the “H” in HVPT – to enrich the field of perceptual training studies. Follow-up studies may simply use a single talker in both the LVPT and HVPT conditions but with consistent productions in the LVPT condition and a range of productions (e.g. variable F0, contour, amplitude, speaking rates, emotional content, etc.) in the HVPT condition to evaluate the effect of newly designed highly-variable stimuli. Meanwhile, while the enlightening part of the present study was to compare the effect of different training designs on perceptual and production learning, future studies can further examine this issue by testing in other contexts such as stimuli under different listening conditions (e.g. quiet vs. noise vs. white noise), with different prototypicality (e.g. prototypes vs. non-prototypical sounds), with different attention weight (e.g. stimuli that are sound-attending vs. meaning-attending), with manipulated acoustic cues, or using nonce words (e.g. real vs. nonce), with different visual representation (e.g. orthographical vs. pictorial), with different stimulus manipulations (e.g. electronically-altered vs. synthetic vs. produced by children), providing subjects with different amount of exposure to materials or set size (e.g. subset vs. full set of stimuli), at the sentence level, etc. These factors can be addressed in future studies to access the training effectiveness and robustness in translating the learning to different conditions and further exploring the mechanisms underlying speech perception and production.

Although the aim of the present study was not testing any speech learning models, the assimilation patterns inferred by analyzing the data are still worth noting. The genuine nature of the assimilation pattern should be more carefully and directly assessed in future cross-linguistic studies such as studies conducted by Morrison (2006), Munro et al. (2000) or Thomson (2007). This can be achieved by using perception tests on monolingual speakers of one language who will be asked to label the segments of the second language under investigation in that study in terms of the monolingual speaker's native language. For this study, it will be involving monolingual Cantonese speakers labeling English vowels in terms of Cantonese vowels.

Last but not least, future efforts should be directed at individual variability among the subjects both before and after training. The improvement in different groups in the present study only gave an evaluation of the different training paradigm at the aggregate level without considering individual performances. From this study, there are clearly some learners who benefit from training more than others. Some previous studies (e.g. Perrachione et al., 2011) have also shown that individual's performance at the pretest is a good predictor of learning outcomes, and thus more efficacious training paradigms can be developed and variability found in different training studies can be eliminated when individual differences of subjects can be taken into account.

APPENDIX A
SURVEY FORM & CONSENT FORM

SURVEY FORM
Language background and language-learning habits

Before participating in this research study, please fill in some important information for further arrangement of groupings.

The information you provide will be of high importance to the researcher and contribute a lot to the success of the study. They will be kept confidential and will only be used for the research only.

Please use **ENGLISH** to fill in the form, and use Chinese/Cantonese only when you find that you cannot express yourself.

I. General Information 基本資料:

Name 姓名: (English) ________________ (中文) __________________

Age 年齡: _______ Grade 年級: ___________ Nationality 國籍: _____________

Place of Birth 出生地: _______________

Hearing deficit 聽覺障礙: Yes / No If "yes," please specify 如有，請註明: ______

Speaking deficit 說話障礙: Yes / No If "yes," please specify 如有，請註明: ______

Contact email (for research purpose only) 聯絡電郵地址: ___________________________

II. Language Background 語言背景:

- Medium of Instruction at school 教學語言: EMI / CMI
- Mother tongue/First language 母語 / 第一語言 (e.g. Cantonese): _________________
- Second language 第二語言: _________________
- Third or more languages that you speak 第三或其他能操之語言: ____

- ***Staying*** in an English-speaking country 曾於以英語為主要語言之地方***暫居***:

    Yes / No

    If "yes," please specify 如有，請註明:

    (place) ____________________ (for how long) __________________

- ***Living*** in an English-speaking country 曾於以英語為主要語言之地方***居住***:

    Yes / No

    If "yes," please specify 如有，請註明:

    (place) ____________________ (for how long) __________________

- ***Studying*** in an English-speaking country 曾於以英語為主要語言之地方***留學***:

    Yes / No

    If "yes," please specify 如有，請註明:

    (place) ____________________ (for how long) __________________

- Age when you started learning English 學習英語的年齡:

    __________ (______ years)

- Learning English OUTISIDE school 在校外學習英語的機會:

Types of program 課程類別	**From 由**	**To 至**
e.g. English tutorial courses for HKCEE	April 2009	Present

- Contact of English OUTSIDE school 在校外接觸英語的渠道:

    (e.g. watch English movies regularly, speak to native speakers regularly, etc.)

Ways of contact 接觸英語的方法	**Frequency 頻率**
e.g. Watching English TV dramas	Every night for 3 hours

CONSENT FORM
Speech Perception and Production

The purpose of this research is to investigate the relationship between speech perception and production. Research participants will complete a series of tests and training sessions and fill in a written survey form related to their language background and habits. This research is important to the understanding of how and to what extent speech perception can affect the second language production of particular vowel sounds among Cantonese speakers of English.

At any time in the study, research participants may refuse to participate, discontinue their involvement at any time, and skip any questions or decline to participate in any portion of the study that may make them feel uncomfortable. This will be done without jeopardizing their confidentiality. There are no foreseeable risks involved in this experiment.

Only the participating investigators will have access to the data. All data and audio clips obtained from the interviews will be destroyed after the study had completed. Participants are encouraged to ask the researcher, Janice Wong, questions about the research project at any time. Her contact is listed below:

Wong Wing Sze, Janice
PhD Research Student
Department of English
The Chinese University of Hong Kong
306, Fung King Hei Building
Shatin, N.T., Hong Kong
Phone: xxx
E-mail: xxx

~~~~~~

I verify that this study has been explained to me and that I voluntarily agree to participate. I understand that if I have any hesitation, I reserve the right to discontinue my participation in the project at any time and may request that all information I have provided be destroyed.

____________________________ ________________________

Participant Signature Date

Thank you for your participation!
~~~~~~

Time available to participate in the experiment 實驗時間選擇:

＊請按您的時間選參與下列實驗（請在格加✓，可選多個，最後會隨機分配入一組）＊
如您能夠出席較多次數的實驗，請盡量不要選 Group5 或 6 T.T
您的幫忙對我的畢業論文非常非常重要，懇請您盡量幫手！^^

Please tick here!

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Please tick here!						
步驟一	Production Pre-test **(5-10 mins 內完成)** 錄音 60 個英文字					
步驟二	Perception Pre-test **(10-15 mins 內完成)** 聽 120 個英文字(二選一題目)					
步驟三	分兩日 Perception training ~聽 120 個字(二選一題目) ~一日聽 10 次 ~每次 10-15 mins 內完成 ~小息、lunch 放學分開做		分十日 Perception training ~聽 120 個字(二選一題目) ~一日聽 2 次 ~每次 10-15 mins 內完成 ~小息、lunch 及放學分開做		沒有	沒有
步驟四	分四次 Production training ~看發音影片 ~即場學發音 ~15mins 完成 ~可小息、lunch 及放學分開做	沒有	分四次 Production training ~看發音影片 ~即場學發音 ~15mins 完成 ~可小息、lunch 及放學分開做	沒有	分四次 Production training ~看發音影片 ~即場學發音 ~15mins 完成 ~可小息、lunch 及放學分開做	沒有
步驟五	1) Production Post-Test 01 ~錄音 60 個英文字 **(5-10 mins 內完成)** 2) Production Post-Test 02 ~錄音一篇 500 字文章 **(10 mins 內完成)**					
步驟六	1) Perception Post-test 01 **(10-15 mins 內完成)** ~聽 120 個英文字 (二選一題目) 2) Perception Post-test 02 **(10-15 mins 內完成)** ~聽 120 個英文字 (二選一題目) 3) Perception Post-test 03 **(10-15 mins 內完成)** ~聽 120 個英文字 (二選一題目) 4) Perception Post-test 04 **(10-15 mins 內完成)** ~聽 120 個英文字 (二選一題目)					
步驟七	每組有 2 至 4 位同學會被邀請參與簡單的 interview 對答 **(10 mins 內完成)**					

Appendix B

Perception Pre/Posttest & TG3 Wordlist

Trials - 10 Questions

	A	B	C	Answer
1	clock	cold	cot	clock
2	flyer	flower	floor	flower
3	box	books	bow	box
4	life	love	lift	love
5	broad	board	book	board
6	cut	duck	chuck	cut
7	rings	rains	rhymes	rings
8	dies	eyes	dries	eyes
9	fan	friend	blend	friend
10	paper	piper	picker	paper

Section A - 40 Questions

	A	B	C	Answer
1	shooed	should	showed	shooed
2	suit	soot	sowed	suit
3	wooed	wood	woe	wood
4	Luke	look	luck	Luke
5	who'd	hood	hold	who'd
6	wooed	wood	woe	wooed
7	suit	soot	sowed	soot
8	Luke	look	luck	look
9	shooed	should	showed	shooed
10	wooed	wood	woe	wooed
11	who'd	hood	hold	hood
12	suit	soot	sowed	suit
13	who'd	hood	hold	hood
14	shooed	should	showed	should
15	Luke	look	luck	Luke
16	wooed	wood	woe	wood
17	suit	soot	sowed	soot
18	shooed	should	showed	should
19	who'd	hood	hold	who'd
20	suit	soot	sowed	suit
21	Luke	look	luck	look
22	wooed	wood	woe	wood
23	shooed	should	showed	shooed
24	Luke	look	luck	look
25	who'd	hood	hold	hood
26	shooed	should	showed	should
27	wooed	wood	woe	wooed
28	suit	soot	sowed	soot
29	Luke	look	luck	Luke
30	who'd	hood	hold	who'd
31	wooed	wood	woe	wood
32	suit	soot	sowed	suit
33	Luke	look	luck	look
34	shooed	should	showed	shooed
35	Luke	look	luck	Luke
36	who'd	hood	hold	hood
37	suit	soot	sowed	soot
38	who'd	hood	hold	who'd
39	wooed	wood	woe	wooed
40	shooed	should	showed	should

Section B - 40 Questions

	A	B	C	Answer
1	bed	bad	bold	bed
2	leg	lag	lake	lag
3	dead	dad	deed	dead
4	vet	vat	vote	vat
5	head	had	hide	had
6	beck	back	bake	beck
7	Med	mad	maid	Med
8	hep	hap	hope	hap
9	fed	fad	feed	fed
10	said	sad	sowed	sad
11	leg	lag	lake	leg
12	beck	back	bake	back
13	bet	bat	boat	bet
14	set	sat	seed	sat
15	head	had	hide	head
16	deb	dab	dub	dab
17	bed	bad	bold	bad
18	guess	gas	guts	guess
19	pep	pap	pipe	pap
20	vet	vat	vote	vet
21	met	mat	meet	mat
22	dead	dad	deed	dad
23	hep	hap	hope	hep
24	let	lat	lot	lat
25	said	sad	sowed	said
26	Med	mad	maid	mad
27	beg	bag	bug	beg
28	set	sat	seed	set
29	led	lad	lord	lad
30	wreck	rack	ruck	wreck
31	fed	fad	feed	fad
32	let	lat	lot	let
33	wreck	rack	ruck	rack
34	deb	dab	dub	deb
35	beg	bag	bug	bag
36	bet	bat	boat	bat
37	pep	pap	pipe	pep
38	guess	gas	guts	gas
39	led	lad	lord	led
40	met	mat	meet	met

Section C - 40 Questions

	A	B	C	Answer
1	mead	mid	maid	mid
2	keep	kip	cap	keep
3	leap	lip	loop	leap
4	beat	bit	boat	bit
5	seed	Sid	sad	Sid
6	reed	rid	red	read
7	leave	live (v.)	love	leave
8	keep	kip	cap	kip
9	lead	lid	load	lid
10	bees	biz	buzz	bees
11	deed	did	dude	did
12	seed	Sid	sad	seed
13	bead	bid	bud	bead
14	piece	piss	pass	piss
15	leave	live (v.)	love	live (v.)
16	beef	biff	buff	biff
17	piece	piss	pass	piece
18	mead	mid	maid	mead
19	deed	did	dude	deed
20	heat	hit	hut	hit
21	reed	rid	red	rid
22	beat	bit	boat	beat
23	leap	lip	loop	lip
24	beef	biff	buff	beef
25	meat	mitt	mute	mit
26	bead	bid	bud	bid
27	seat	sit	sad	seat
28	heat	hit	hut	heat
29	keyed	kid	code	keyed
30	deep	dip	dupe	dip
31	bees	biz	buzz	biz
32	heed	hid	hade	heed
33	meat	mitt	mute	meat
34	deep	dip	dupe	deep
35	reek	Rick	rake	Rick
36	heed	hid	hade	hid
37	lead	lid	load	lead
38	reek	Rick	rake	reek
39	seat	sit	sad	sit
40	keyed	kid	code	kid

APPENDIX C
PRODUCTION PRE/POSTTEST WORDLIST

Production Pre-/Post-test

Instructions:

1. Please READ the following wordlist, see which word you don’t know how to pronounce Please ask me or consult the dictionary before recording.

1. 請細閱本字列，如果你發現有任何字你是不懂得讀的，請先問我，或查字典。

2. Please read all the words and record them. Please stop for 1 or 2 seconds between each word.

2. 請錄下下列詞語。在每個字之間請停頓一至兩秒，不要一口氣或太快讀完。

3. You DO NOT have to say the number, You just have to say the word.

3. 你不需要讀 number，只需要讀那個字就可以了。

1 should
2 bad
3 sad
4 deed
5 hood
6 moist
7 beck
8 wooed
9 mid
10 leap
11 page
12 bet
13 suit
14 set
15 bit
16 Rick
17 choice
18 mat
19 mitt
20 bag
21 fad
22 keep
23 bat
24 Luke
25 vet
26 Ruby
27 soot
28 sit
29 leg
30 light
31 sat
32 beat
33 mad
34 keyed
35 rain
36 seed
37 who'd
38 meat
39 met
40 live (v.)
41 bid
42 beg
43 clock
44 vat
45 Med
46 normal
47 rid
48 fed
49 piss
50 said
51 Sid
52 wood
53 cross
54 lead (v.)
55 back
56 look
57 seat
58 lobby
59 bead
60 shooed

Production Pre-/Post-test (NOT for participants)

Instructions:

1. Please READ the following wordlist, see which word you don't know how to pronounce Please ask me or consult the dictionary before recording.

1. 請細閱本字列，如果你發現有任何字你是不懂得讀的，請先問我，或查字典。

2. Please read all the words and record them. Please stop for 1 or 2 seconds between each word.

2. 請錄下下列詞語。在每個字之間請停頓一至兩秒，不要一口氣或太快讀完。

3. You DO NOT have to say the number, You just have to say the word.

3. 你不需要讀 number，只需要讀那個字就可以了。

1 should	/ʊ/	21 fad	/æ/	41 bid	/ɪ/
2 bad	/æ/	22 keep	/iː/	42 beg	/e/
3 sad	/æ/	23 bat	/æ/	43 clock	
4 deed	/iː/	24 Luke	/uː/	44 vat	/æ/
5 hood	/ʊ/	25 vet	/e/	45 Med	/e/
6 moist		26 Ruby		46 normal	
7 beck	/e/	27 soot	/ʊ/	47 rid	/ɪ/
8 wooed	/uː/	28 sit	/ɪ/	48 fed	/e/
9 mid	/ɪ/	29 leg	/e/	49 piss	/ɪ/
10 leap	/iː/	30 light		50 said	/e/
11 page		31 sat	/æ/	51 Sid	/ɪ/
12 bet	/e/	32 beat	/iː/	52 wood	/ʊ/
13 suit	/uː/	33 mad	/æ/	53 cross	
14 set	/e/	34 keyed	/iː/	54 lead (v.)	/iː/
15 bit	/ɪ/	35 rain		55 back	/æ/
16 Rick	/ɪ/	36 seed	/iː/	56 look	/ʊ/
17 choice		37 who'd	/uː/	57 seat	/iː/
18 mat	/æ/	38 meat	/iː/	58 lobby	
19 mitt	/ɪ/	39 met	/e/	59 bead	/iː/
20 bag	/æ/	40 live (v.)	/ɪ/	60 shooed	/uː/

APPENDIX D

PERCEPTION TRAINING WORDLIST

Section A - 40 Questions

	A	B	Answer
1	shooed	should	shooed
2	suit	soot	suit
3	wooed	wood	wood
4	Luke	look	Luke
5	who'd	hood	who'd
6	wooed	wood	wooed
7	suit	soot	soot
8	Luke	look	look
9	shooed	should	shooed
10	wooed	wood	wooed
11	who'd	hood	hood
12	suit	soot	suit
13	who'd	hood	hood
14	shooed	should	should
15	Luke	look	Luke
16	wooed	wood	wood
17	suit	soot	soot
18	shooed	should	should
19	who'd	hood	who'd
20	suit	soot	suit
21	Luke	look	look
22	wooed	wood	wood
23	shooed	should	shooed
24	Luke	look	look
25	who'd	hood	hood
26	shooed	should	should
27	wooed	wood	wooed
28	suit	soot	soot
29	Luke	look	Luke
30	who'd	hood	who'd
31	wooed	wood	wood
32	suit	soot	suit
33	Luke	look	look
34	shooed	should	shooed
35	Luke	look	Luke
36	who'd	hood	hood
37	suit	soot	soot
38	who'd	hood	who'd
39	wooed	wood	wooed
40	shooed	should	should

Section B - 40 Questions

	A	B	Answer
1	bed	bad	bed
2	leg	lag	lag
3	dead	dad	dead
4	vet	vat	vat
5	head	had	had
6	beck	back	beck
7	Med	mad	Med
8	hep	hap	hap
9	fed	fad	fed
10	said	sad	sad
11	leg	lag	leg
12	beck	back	back
13	bet	bat	bet
14	set	sat	sat
15	head	had	head
16	deb	dab	dab
17	bed	bad	bad
18	guess	gas	guess
19	pep	pap	pap
20	vet	vat	vet
21	met	mat	mat
22	dead	dad	dad
23	hep	hap	hep
24	let	lat	lat
25	said	sad	said
26	Med	mad	mad
27	beg	bag	beg
28	set	sat	set
29	led	lad	lad
30	wreck	rack	wreck
31	fed	fad	fad
32	let	lat	let
33	wreck	rack	rack
34	deb	dab	deb
35	beg	bag	bag
36	bet	bat	bat
37	pep	pap	pep
38	guess	gas	gas
39	led	lad	led
40	met	mat	met

Section C - 40 Questions

	A	B	Answer
1	mead	mid	mid
2	keep	kip	keep
3	leap	lip	leap
4	beat	bit	bit
5	seed	Sid	Sid
6	reed	rid	read
7	leave	live (v.)	leave
8	keep	kip	kip
9	lead	lid	lid
10	bees	biz	bees
11	deed	did	did
12	seed	Sid	seed
13	bead	bid	bead
14	piece	piss	piss
15	leave	live (v.)	live (v.)
16	beef	biff	biff
17	piece	piss	piece
18	mead	mid	mead
19	deed	did	deed
20	heat	hit	hit
21	reed	rid	rid
22	beat	bit	beat
23	leap	lip	lip
24	beef	biff	beef
25	meat	mitt	mit
26	bead	bid	bid
27	seat	sit	seat
28	heat	hit	heat
29	keyed	kid	keyed
30	deep	dip	dip
31	bees	biz	biz
32	heed	hid	heed
33	meat	mitt	meat
34	deep	dip	deep
35	reek	Rick	Rick
36	heed	hid	hid
37	lead	lid	lead
38	reek	Rick	reek
39	seat	sit	sit
40	keyed	kid	kid

APPENDIX E
PRODUCTION TRAINING MATERIALS

Day 1 – Front Vowels – /iː/ and /ɪ/

Learning the vowels @ 10 mins
http://www.youtube.com/watch?v=kjHcM4qwiag

Words to practice @ 5 pairs
1. deed did
2. bead bid
3. piece piss
4. beef biff
5. heat hit

Day 2 – Mid Vowels – /e/ and /æ/

Learning the vowels @ 4mins
http://www.youtube.com/watch?v=iErAkPdYenE

Words to practice @ 5 pairs
1. beck back
2. Med mad
3. hep hap
4. fed fad
5. said sad
6.

Passage to read @ 5 mins

Anson is a cat which lives at the corner of Ken's attic. It is an extremely bad cat as it rarely behaves well. It drops a pan of ham onto Ken's bed. It puts Sally's hats onto Ken's head. It destroys all the pens in Ken's bag. It nearly bends the neck of Ben the bat to death! Though Anson is bad, no one dares to kick it out because it is Ken's expensive pet! Ken spent all his salary in February to buy it!

One day, a brave rat called Pat decides to play tricks on this bad fellow. As Pat knows that Anson loves a valuable gem in Ken's home, Pat puts some jam and lettuce which Anson loves on it. Some spicy black pepper is also added! Guess what? Anson really eats the food when playing with the gem. It chokes and goes mad because of the black pepper, but it is just an accident that it even swallows the gem! Everyone is afraid of what will happen next but they all cannot help laughing. When Anson knows that nobody cares for him, it is sad. At the end, Anson seems to have realized that it had better change and so it apologizes! Pat and the others think that Anson has learnt a lesson, so they also say sorry to Anson and shake hands. Pat has even become Anson's fan as Anson is good at standing on a flashlight with only one of its legs! They are now very good friends!

Day 3 – Back Vowels – /uː/ and /ʊ/

Learning the vowels @ 6mins

http://www.youtube.com/watch?v=BYia6FWOdJc

Words to practice @ 5 mins

1. shooed should
2. suit soot
3. wooed wood
4. Luke look
5. who'd hood

Passage to read @ 5 mins

Jim's town has a zoo. The zoo has a new ride called the Boom Zoom. The Boom Zoom goes up and down. It has a loop, too. To get on the ride you have to go up to the roof.

Jim and his friends got in line for the Boom Zoom. They waited in line until noon. The line didn't move. They didn't have food. Finally, they got to the roof, and it was good. After the ride, Jim didn't feel too good.

My mum got mad. She said my room needed to be cleaned soon. She said the dust was so thick she can't even use a broom. She went to the room where we keep the tools. She must have been very mad. She didn't come back with a duster. She didn't bring a broom. She took the tool by the hoop. Her plan was to put it all in the trash in big scoops.

Day 4 – Review of 3 pairs

A Review@ 10 mins
http://www.uiowa.edu/~acadtech/phonetics/english/frameset.html

Go to right hand corner "monothongs" → click "front" and listen to the vowel sounds, read the animation and videos.
Just focus on /i/, /ɪ/, /ɛ/ (but NOT /e/ !!!), /æ/

Then click "back" and listen to the vowel sounds, read the animation and videos.
Just focus on /u/, /ʊ/

Words to practice @ 10 mins

/uː/	/ʊ/	/e/	/æ/	/iː/	/ɪ/
shooed	should	head	had	seat	sit
suit	soot	guess	gas	keyed	kid
wooed	wood	bet	bat	reek	Rick
Luke	look	bed	bad	heed	hid
who'd	hood	beck	back	deep	dip
		beg	bag	bead	bid
		dead	dad	beat	bit
		fed	fad	keep	kip
		let	lat	lead	lid
		said	sad	bees	biz

APPENDIX F

READING PASSAGE IN TEST OF CONTEXTUALIZATION

Test of Contextualization (TC)

April Fools

If you've ever been pranked on April Fools' Day, you may wonder how this tradition started. Well, you're not alone. No one knows for sure how April Fools' Day began, but the most likely explanation has to do with the calendar.

No, that's not an April Fools' Day joke. People used to celebrate New Years Day on April 1st. Just like today, people would choose to have big parties and cook some good food to celebrate. Over time, the calendar changed and so did the date for New Years. In the Fifteen Hundreds, the new calendar put New Year's Day as January 1st. But because there was no Internet or other means to spread the word, the news traveled slowly by word of mouth. It took a while for everyone to speak to others about the change, but some people are not happy with it. They looked at it as a bad push. Due to this reason, they carried on celebrating New Years on April 1st. These people were given the cool nickname, "April fools."

People following the new calendar played tricks on the "April fools" by sending them on "fool's errands." They had the "April fools" release news and deliver invitations to big New Year's celebrations that weren't going to happen. In France, "April fools" were called "April Fish" because people thought fish were easy to catch since they could be fooled into taking the bait on a hook. Children would tag a paper fish on a person's back to mark them as an "April Fish." When the person discovered the fish, the prankster yelled on the streets, shook his head and said, "April Fish!" How crucial it is to be up to date!

Not everyone is fully convinced that this is actually how the tradition of April Fools' Day began in the previous days – no one could prove it and no book recorded this! People have tried to keep pinpointing the exact date of the first April Fool's Day, but this only led to more pranks. A canny professor from Boston University pranked a reporter by making up a story about a court jester who said he could run the empire better than the king. The jester was made king and held a feast to eat for a day on April 1st. This turned out to be a big April Fools' Day trick because the reporter thought the story was true.

Even though we aren't sure how this tradition began, people still celebrate April Fools' Day by playing tricks on each other. So the next time you prank someone and yell "April Fools!" remember that the day may actually be about the people who didn't want to change their traditions when the new calendar was adopted. Or maybe it's just a day to celebrate the joker in all of us.

Test of Contextualization (TC) (NOT for participants)

April Fools

Coding Scale:

Green: /iː/ (15 words) **Deep Pink: /ɪ/ (15 words)**

Purple: /e/ (15 words) **Red: /æ/ (15 words)**

Boxed: /uː/ (10 words) **Blue: /ʊ/ (10 words)**

If you've ever been pranked[1] on April Fools'[1] Day, you may wonder how this tradition[1] started. Well, you're not alone. No one knows for sure how April Fools' Day began[2], but the most likely explanation has to do with the calendar[3].

No, that's not an April Fools' Day joke. People[1] used[2] to celebrate[1] New[3] Years[2] Day on April 1st. Just like today, people would choose[4] to have big[2] parties and cook[1] some good[2] food[5] to celebrate. Over time, the calendar changed and so did the date for New Years. In the Fifteen[3] Hundreds, the new calendar put[3] New Year's Day as January[4] 1st. But because there was no Internet[3] or other means[4] to spread[2] the word, the news traveled slowly by word of mouth. It took[4] a while for everyone to speak[5] to others about the change, but some people are not happy[5] with it. They looked[5] at it as a bad[6] push[6]. Due to this reason[6], they carried[7] on celebrating New Years on April 1st. These people were given[4] the cool[6] nickname[5], "April fools."

People following the new calendar played tricks[6] on the "April fools" by sending[3] them on "fool's errands[8]." They had the "April fools" release[7] news and deliver[7] invitations[8] to big New Year's celebrations that weren't going to happen[9]. In France, "April fools" were called "April Fish[9]" because people thought fish were easy[8] to catch[10] since[9] they could be fooled into taking the bait on a hook[7]. Children[10] would tag[11] a paper fish on a person's back[12] to mark them as an "April Fish." When the person discovered the fish, the prankster yelled[4] on the streets[10], shook[8] his head[5] and said, "April Fish!" How crucial[7] it is to be up to date!

Not everyone[6] is fully[9] convinced[11] that this is actually[13] how the tradition of April Fools' Day began – no one could prove[8] it and no book[10] recorded this! People have tried to keep[11] pinpointing[12] the exact[14] date of the first April Fool's Day, but this only led[7] to more pranks. A canny[15] professor[8] from Boston University[9] pranked a reporter by making up a story about a court jester[9] who said[10] he could run the empire[11] better[12] than the king[13]. The jester was made king and held[13] a feast[12] to eat[13] for a day on April 1st. This turned out to be a big April Fools' Day trick because the reporter thought the story was true[10].

Even[14] though we aren't sure how this tradition began, people still[14] celebrate April Fools' Day by playing tricks on each[15] other. So the next[14] time you prank someone and yell "April Fools!" remember[15] that the day may actually be about the people who didn't want to change their traditions[15] when the new calendar was adopted. Or maybe it's just a day to celebrate the joker in all of us.

Appendix G

Perception Wordlist in TG1

Section A - 40 Questions

	A	B	C	Answer
1	buhl	bull	bail	buhl
2	lucre	looker	licker	lucre
3	stewed	stood	stead	stood
4	boot	book	bake	book
5	lucre	looker	licker	looker
6	stewed	stood	stead	stewed
7	buhl	bull	bail	buhl
8	stewed	stood	stead	stood
9	boot	book	bake	boot
10	lucre	looker	licker	lucre
11	boot	book	bake	book
12	stewed	stood	stead	stewed
13	buhl	bull	bail	bull
14	lucre	looker	licker	lucre
15	stewed	stood	stead	stewed
16	boot	book	bake	book
17	lucre	looker	licker	looker
18	stewed	stood	stead	stood
19	buhl	bull	bail	bull
20	boot	book	bake	boot
21	lucre	looker	licker	looker
22	boot	book	bake	boot
23	buhl	bull	bail	bull
24	stewed	stood	stead	stewed
25	lucre	looker	licker	looker
26	buhl	bull	bail	buhl
27	stewed	stood	stead	stood
28	boot	book	bake	boot
29	buhl	bull	bail	bull
30	stewed	stood	stead	stewed
31	lucre	looker	licker	lucre
32	boot	book	bake	book
33	buhl	bull	bail	bull
34	boot	book	bake	boot
35	lucre	looker	licker	looker
36	boot	book	bake	book
37	buhl	bull	bail	buhl
38	lucre	looker	licker	lucre
39	stewed	stood	stead	stood
40	buhl	bull	bail	buhl

Section B - 40 Questions

	A	B	C	Answer
1	errant	arrant	current	errant
2	Ester	aster	Easter	Ester
3	bend	band	bind	band
4	effluent	affluent	influence	effluent
5	beryl	barrel	bureau	barrel
6	Ken	can	keen	Ken
7	reddish	radish	rubbish	radish
8	ketches	catches	kichens	catches
9	crept	crapped	cripped	crept
10	cress	crass	Chris	crass
11	expend	expand	explain	expend
12	expense	expanse	explict	expense
13	fleck	flak	flick	flak
14	hem	ham	him	ham
15	neck	knack	nuke	neck
16	mention	mansion	emulsion	mansion
17	mellow	mallow	willows	mellow
18	men	man	mean	man
19	cleanse	clans	clunks	clans
20	bread	brad	breed	bread
21	pellet	palette	pillage	palette
22	pen	pan	pin	pan
23	rend	rand	round	rend
24	kettle	cattle	kitten	cattle
25	shed	shad	showed	shad
26	setter	satyr	sitter	setter
27	slept	slapped	slipped	slept
28	spend	spanned	spinned	spend
29	phonetic	fanatic	domestic	fanatic
30	text	taxed	fixed	text
31	letter	latter	litter	latter
32	temper	tamper	timber	temper
33	Ben	ban	bin	ban
34	gem	jam	Jim	jam
35	tend	tanned	towned	tanned
36	flesh	flash	fish	flesh
37	fetter	fatter	fitter	fetter
38	heck	hack	hock	hack
39	send	sand	sinned	send
40	melody	malady	medly	melody

Section C - 40 Questions

	A	B	C	Answer
1	been	bin	bun	bin
2	cheek	chick	chuck	cheek
3	beet	bit	but	bit
4	beta	bitter	butter	bitter
5	cheap	chip	chap	cheap
6	cheat	chit	chat	chit
7	measles	mizzles	muscles	mizzles
8	feat	fit	fat	feat
9	deem	dim	dam	deem
10	eat	it	ate	eat
11	fees	fizz	fuzz	fizz
12	feet	fit	fat	fit
13	greed	grid	grade	grid
14	green	grin	gran	green
15	heap	hip	hope	hip
16	feel	fill	foul	feel
17	lean	Lynn	loan	Lynn
18	neap	nip	nap	nip
19	scenic	cynic	psychic	scenic
20	peach	pitch	punch	pitch
21	queen	quin	quench	queen
22	reed	rid	raid	reed
23	steeple	stiple	staple	steeple
24	schema	skimmer	scamper	skimmer
25	scheme	skim	scam	skim
26	sleet	slit	slate	slit
27	steal	still	stole	steal
28	teen	tin	tone	teen
29	teeter	titter	toddler	titter
30	weep	whip	warp	whip
31	neat	nit	nate	neat
32	seal	sill	sow	seal
33	bean	bin	ban	bean
34	ease	is	ice	is
35	reach	rich	rage	reach
36	Caesars	scissors	session	scissors
37	gene	gin	Jan	gene
38	sleep	slip	slap	sleep
39	greased	grist	grace	grist
40	peep	pip	pipe	peep

Perception Wordlist in TG2

Section A - 40 Questions

	A	B	C	Answer
1	gooed	good	gold	gooed
2	fool	full	fill	fool
3	pool	pull	pill	pull
4	cooed	could	cold	cooed
5	fool	full	fill	full
6	gooed	good	gold	good
7	pool	pull	pill	pool
8	pool	pull	pill	pull
9	cooed	could	cud	could
10	gooed	good	gold	gooed
11	pool	pull	pill	pool
12	fool	full	fill	full
13	pool	pull	pill	pull
14	gooed	good	gold	gooed
15	cooed	could	cud	could
16	fool	full	fill	full
17	cooed	could	cud	could
18	gooed	good	gold	good
19	cooed	could	cud	cooed
20	fool	full	fill	fool
21	cooed	could	cud	could
22	cooed	could	cud	cooed
23	gooed	good	gold	gooed
24	fool	full	fill	full
25	pool	pull	pill	pull
26	cooed	could	cud	could
27	gooed	good	gold	good
28	cooed	could	cud	cooed
29	fool	full	fill	fool
30	pool	pull	pill	pool
31	fool	full	fill	full
32	gooed	good	gold	good
33	pool	pull	pill	pull
34	cooed	could	cud	cooed
35	gooed	good	gold	good
36	fool	full	fill	fool
37	pool	pull	pill	pool
38	gooed	good	gold	gooed
39	pool	pull	pill	pool
40	fool	full	fill	fool

Section B - 40 Questions

	A	B	C	Answer
1	X	axe	its	axe
2	M	am	him	M
3	better	batter	bitter	batter
4	beget	begat	begin	beget
5	blether	blather	blighter	blather
6	kept	capped	keep	kept
7	rent	rant	raind	rent
8	celery	salary	silly	salary
9	Deb	dab	deep	dab
10	end	and	owned	end
11	fellow	fallow	pillow	fellow
12	tense	tans	tins	tans
13	nets	gnats	needs	gnats
14	medley	madly	middle	madly
15	pedal	paddle	piddle	pedal
16	lend	land	lined	lend
17	Bren	bran	brine	Bren
18	messy	massy	missy	massy
19	overleapt	overlapped	overlooked	overlapped
20	peck	pack	pick	pack
21	vessel	vassal	missle	vassal
22	gen	Jan	jean	Jan
23	rep	rap	rub	rep
24	ret	rat	rut	ret
25	sexes	saxes	sixes	sexes
26	shekel	shackle	shuffle	shackle
27	techs	tacks	ticks	techs
28	telly	tally	trolly	tally
29	thresh	thrash	thrift	thresh
30	trek	track	trick	trek
31	Meg	Mag	mug	Mag
32	Ellie	Ali	uni	Ali
33	shell	shall	shield	shell
34	elegy	allergy	eulogy	elegy
35	belly	bally	Billy	bally
36	mental	mantel	middle	mental
37	sec	sac	suck	sac
38	temp	tamp	timb	temp
39	vex	vacs	fix	vex
40	ember	amber	timber	amber

Section C - 40 Questions

	A	B	C	Answer
1	bleats	blitz	blades	blitz
2	bleep	blip	blue	bleep
3	weaner	winner	wander	winner
4	bream	brim	balm	brim
5	cheep	chip	chap	cheep
6	crease	Chris	craze	crease
7	least	list	last	least
8	peel	pill	Paul	pill
9	eel	ill	all	ill
10	field	filled	fade	field
11	fleet	flit	flight	fleet
12	forefeet	forfeit	forsake	forfeit
13	frees	frizz	fries	frizz
14	greet	grit	great	greet
15	feast	fist	fast	fist
16	leas	Liz	laze	leas
17	peace	piss	pass	piss
18	Neil	nil	nail	nil
19	peat	pit	pat	peat
20	reason	risen	rasin	risen
21	teak	tic	take	teak
22	sneak	snick	snake	sneak
23	seen	sin	sown	sin
24	seep	sip	sap	sip
25	sheen	shin	shun	sheen
26	sheep	ship	shape	sheep
27	steel	still	stale	still
28	tweet	twit	twat	tweet
29	wheat	whit	wait	whit
30	wheeze	whizz	waste	wheeze
31	beef	biff	bath	biff
32	he'll	hill	hail	hill
33	meal	mill	male	meal
34	ream	rim	realm	ream
35	wean	win	won	win
36	Celia	sillier	seller	Celia
37	jeep	gyp	jap	gyp
38	reap	rip	rap	reap
39	weaned	wind (n.)	wand	weaned
40	dean	din	dun	din

APPENDIX H

LANGUAGE BACKGROUND OF SUBJECTS

Username		Gender	L1	Age	Age of Learning	Years of Learning	Staying abroad (per year)	Living abroad (per year)	Studying abroad (per year)
HSP									
1	a01	F	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
2	a02	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
3	a11	F	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
4	a37	F	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
5	b02	F	Cantonese	16.00	2.00	14.00	0.00	0.00	0.00
6	b04	F	Cantonese	17.00	2.00	15.00	0.00	0.00	0.00
7	b38	M	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
8	a25	M	Cantonese	17.00	5.00	12.00	0.00	0.00	0.00
9	c38	M	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
			Mean:	**16.44**	**3.22**	**13.22**	**0.00**	**0.00**	**0.00**
			S.D.:	***0.53***	***0.97***	***1.09***	***0.00***	***0.00***	***0.00***
HIP									
1	a17	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
2	c02	F	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
3	c03	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
4	c04	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
5	c06	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
6	c35	M	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
7	c36	M	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
8	c37	M	Cantonese	16.00	5.00	11.00	0.00	0.00	0.00
9	c27	M	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
			Mean:	**16.11**	**3.33**	**12.78**	**0.00**	**0.00**	**0.00**
			S.D.:	***0.33***	***0.71***	***0.83***	***0.00***	***0.00***	***0.00***
HSN									
1	b01	F	Cantonese	17.00	5.00	12.00	0.00	0.00	0.00
2	b03	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
3	b05	F	Cantonese	14.00	3.00	13.00	0.00	0.00	0.00
4	b06	F	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
5	b07	F	Cantonese	18.00	3.00	15.00	0.00	0.00	0.00
6	b08	F	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
7	b25	M	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
8	b21	M	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
9	c39	M	Cantonese	16.00	6.00	10.00	0.05	0.00	0.00
			Mean:	**16.33**	**3.67**	**12.89**	**0.01**	**0.00**	**0.00**
			S.D.:	***1.12***	***1.12***	***1.45***	***0.02***	***0.00***	***0.00***
HIN									
1	a04	F	Cantonese	17.00	4.00	13.00	0.00	0.00	0.00
2	a10	F	Cantonese	16.00	5.00	11.00	0.00	0.00	0.00
3	b11	F	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00

4	c08	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
5	c09	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
6	c15	F	Cantonese	17.00	4.00	13.00	0.00	0.00	0.00
7	b37	M	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
8	b39	M	Cantonese	17.00	4.00	12.00	0.00	0.00	0.00
			Mean:	**16.38**	**3.75**	**12.50**	**0.00**	**0.00**	**0.00**
			S.D.:	**0.52**	**0.71**	**0.76**	**0.00**	**0.00**	**0.00**
LSP									
1	b16	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
2	b17	F	Cantonese	16.00	3.00	13.00	0.06	0.00	0.00
3	a14	F	Cantonese	17.00	5.00	12.00	0.00	0.00	0.00
4	b20	F	Cantonese	16.00	5.00	11.00	0.00	0.00	0.00
5	b29	M	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
6	b35	M	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
7	a23	M	Cantonese	16.00	5.00	11.00	0.00	0.00	0.00
			Mean:	**16.14**	**4.00**	**12.14**	**0.00**	**0.00**	**0.00**
			S.D.:	***0.38***	***1.00***	***0.90***	***0.00***	***0.00***	***0.00***
LIP									
1	c12	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
2	c18	F	Cantonese	17.00	4.00	13.00	0.00	0.00	0.00
3	c21	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
4	c23	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
5	c24	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
6	b27	M	Cantonese	16.00	7.00	9.00	0.00	0.00	0.00
7	c32	M	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
			Mean:	**16.29**	**3.71**	**12.57**	**0.00**	**0.00**	**0.00**
			S.D.:	***0.49***	***1.50***	***1.62***	***0.00***	***0.00***	***0.00***
LSN									
1	a12	F	Cantonese	17.00	2.00	15.00	0.00	0.00	0.00
2	a15	F	Cantonese	17.00	6.00	10.00	0.00	0.00	0.00
3	a19	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
4	b19	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
5	c22	F	Cantonese	16.00	3.00	13.00	0.04	0.00	0.00
6	c26	M	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
7	c31	M	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
			Mean:	**16.43**	**3.29**	**13.00**	**0.00**	**0.00**	**0.00**
			S.D.:	***0.53***	***1.25***	***1.53***	***0.00***	***0.00***	***0.00***
LIN									
1	a09	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
2	a22	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
3	c13	F	Cantonese	18.00	4.00	14.00	0.04	0.00	0.00
4	c17	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
5	c19	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
6	a03	F	Cantonese	17.00	5.00	12.00	0.04	0.00	0.00
7	a21	M	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
8	a24	M	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00

			Mean:	**16.50**	**3.50**	**13.00**	**0.01**	**0.00**	**0.00**
			S.D.:	***0.76***	***0.76***	***0.76***	***0.02***	***0.00***	***0.00***
COP									
1	c01	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
2	c05	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
3	c10	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
4	c20	F	Cantonese	16.00	3.00	13.00	0.06	0.00	0.00
5	c25	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
6	b12	F	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
7	b14	F	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
8	c34	M	Cantonese	16.00	3.00	13.00	0.00	0.00	0.00
9	a26	M	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
10	a29	M	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
11	a31	M	Cantonese	16.00	5.00	11.00	0.00	0.00	0.00
			Mean:	**16.18**	**3.27**	**12.91**	**0.01**	**0.00**	**0.00**
			S.D.:	***0.40***	***0.64***	***0.83***	***0.02***	***0.00***	***0.00***
CON									
1	a16	F	Cantonese	17.00	9.00	7.00	0.00	0.00	0.00
2	a38	F	Cantonese	17.00	6.00	11.00	0.00	0.00	0.00
3	b18	F	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
4	c07	F	Cantonese	17.00	10.00	7.00	0.00	0.00	0.00
5	a05	F	Cantonese	16.00	6.00	10.00	0.00	0.00	0.00
6	b10	F	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
7	a28	M	Cantonese	17.00	4.00	13.00	0.00	0.00	0.00
8	b24	M	Cantonese	17.00	3.00	14.00	0.00	0.00	0.00
9	c29	M	Cantonese	16.00	4.00	12.00	0.00	0.00	0.00
10	c33	M	Cantonese	16.00	5.00	11.00	0.00	0.00	0.00
			Mean:	**16.60**	**5.40**	**11.10**	**0.00**	**0.00**	**0.00**
			S.D.:	***0.52***	***2.41***	***2.51***	***0.00***	***0.00***	***0.00***

*** *Staying abroad* here means staying at an English-speaking country. The subjects who had stayed abroad were in the U.K., Wales or New Zealand.

APPENDIX I
SUBJECTS' PERCENTAGES IN PERCEPTION PRE/POSTTEST & TGS

	Username	Gender	Pre (/ɪ/)	Pre (/iː/)	Pre (avg)	Post (/ɪ/)	Post (/iː/)	Post (avg)
HSP								
1	a01	F	85.00%	65.00%	75.00%	95.00%	85.00%	90.00%
2	a02	F	85.00%	80.00%	82.50%	85.00%	95.00%	90.00%
3	a11	F	85.00%	65.00%	75.00%	90.00%	65.00%	77.50%
4	a37	F	85.00%	65.00%	75.00%	95.00%	65.00%	80.00%
5	b02	F	65.00%	70.00%	67.50%	90.00%	90.00%	90.00%
6	b04	F	65.00%	85.00%	75.00%	100.00%	95.00%	97.50%
7	b38	M	60.00%	95.00%	77.50%	100.00%	95.00%	97.50%
8	a25	M	70.00%	90.00%	80.00%	70.00%	80.00%	75.00%
9	c38	M	80.00%	55.00%	67.50%	95.00%	70.00%	82.50%
		Mean:	**75.56%**	**74.44%**	**75.00%**	**91.11%**	**82.22%**	**86.67%**
		S.D.:	***10.44%***	***13.57%***	***5.00%***	***9.28%***	***12.77%***	***8.29%***
HIP								
1	a17	F	70.00%	65.00%	67.50%	100.00%	75.00%	87.50%
2	c02	F	35.00%	80.00%	57.50%	80.00%	95.00%	87.50%
3	c03	F	80.00%	75.00%	77.50%	85.00%	95.00%	90.00%
4	c04	F	70.00%	80.00%	75.00%	95.00%	90.00%	92.50%
5	c06	F	75.00%	65.00%	70.00%	95.00%	95.00%	95.00%
6	c35	M	50.00%	60.00%	55.00%	95.00%	90.00%	92.50%
7	c36	M	65.00%	85.00%	75.00%	95.00%	90.00%	92.50%
8	c37	M	50.00%	95.00%	72.50%	100.00%	95.00%	97.50%
9	c27	M	70.00%	80.00%	75.00%	90.00%	95.00%	92.50%
		Mean:	**62.78%**	**76.11%**	**69.44%**	**92.78%**	**91.11%**	**91.94%**
		S.D.:	***14.60%***	***11.12%***	***8.08%***	***6.67%***	***6.51%***	***3.25%***
HSN								
1	b01	F	75.00%	80.00%	77.50%	100.00%	95.00%	97.50%
2	b03	F	75.00%	85.00%	80.00%	95.00%	85.00%	90.00%
3	b05	F	65.00%	65.00%	65.00%	100.00%	95.00%	97.50%
4	b06	F	75.00%	60.00%	67.50%	85.00%	95.00%	90.00%
5	b07	F	75.00%	70.00%	72.50%	80.00%	100.00%	90.00%
6	b08	F	70.00%	70.00%	70.00%	90.00%	90.00%	90.00%
7	b25	M	60.00%	70.00%	65.00%	60.00%	80.00%	70.00%
8	b21	M	70.00%	65.00%	67.50%	90.00%	90.00%	90.00%
9	c39	M	45.00%	75.00%	60.00%	75.00%	85.00%	80.00%
		Mean:	**67.78%**	**71.11%**	**69.44%**	**86.11%**	**90.56%**	**88.33%**
		S.D.:	***10.03%***	***7.82%***	***6.35%***	***12.94%***	***6.35%***	***8.57%***
HIN								
1	a04	F	80.00%	60.00%	70.00%	85.00%	95.00%	90.00%

2	a10	F	65.00%	75.00%	70.00%	95.00%	90.00%	92.50%
3	b11	F	85.00%	60.00%	72.50%	100.00%	80.00%	90.00%
4	c08	F	80.00%	75.00%	77.50%	95.00%	90.00%	92.50%
5	c09	F	75.00%	70.00%	72.50%	75.00%	95.00%	85.00%
6	c15	F	60.00%	80.00%	70.00%	100.00%	95.00%	97.50%
7	b37	M	70.00%	75.00%	72.50%	90.00%	85.00%	87.50%
8	b39	M	60.00%	50.00%	55.00%	100.00%	90.00%	95.00%
		Mean:	**71.88%**	**68.13%**	**70.00%**	**92.50%**	**90.00%**	**91.25%**
		S.D.:	***9.61%***	***10.33%***	***6.55%***	***8.86%***	***5.35%***	***4.01%***
LSP								
1	b16	F	60.00%	65.00%	62.50%	95.00%	90.00%	92.50%
2	b17	F	75.00%	70.00%	72.50%	70.00%	75.00%	72.50%
3	a14	F	80.00%	60.00%	70.00%	100.00%	95.00%	97.50%
4	b20	F	60.00%	75.00%	67.50%	95.00%	95.00%	95.00%
5	b29	M	75.00%	65.00%	70.00%	70.00%	70.00%	70.00%
6	b35	M	75.00%	85.00%	80.00%	90.00%	85.00%	87.50%
7	a23	M	80.00%	60.00%	70.00%	95.00%	90.00%	92.50%
		Mean:	**72.14%**	**68.57%**	**70.36%**	**87.86%**	**85.71%**	**86.79%**
		S.D.:	***8.59%***	***9.00%***	***5.29%***	***12.54%***	***9.76%***	***11.06%***
LIP								
1	c12	F	70.00%	40.00%	55.00%	75.00%	75.00%	75.00%
2	c18	F	80.00%	65.00%	72.50%	75.00%	65.00%	70.00%
3	c21	F	75.00%	75.00%	75.00%	100.00%	85.00%	92.50%
4	c23	F	75.00%	65.00%	70.00%	80.00%	85.00%	82.50%
5	c24	F	65.00%	80.00%	72.50%	90.00%	95.00%	92.50%
6	b27	M	50.00%	60.00%	55.00%	100.00%	75.00%	87.50%
7	c32	M	70.00%	90.00%	80.00%	95.00%	85.00%	90.00%
		Mean:	**69.29%**	**67.86%**	**68.57%**	**87.86%**	**80.71%**	**84.29%**
		S.D.:	***9.76%***	***16.04%***	***9.77%***	***11.13%***	***9.76%***	***8.86%***
LSN								
1	a12	F	50.00%	95.00%	72.50%	100.00%	95.00%	97.50%
2	a15	F	75.00%	70.00%	72.50%	100.00%	90.00%	95.00%
3	a19	F	85.00%	65.00%	75.00%	95.00%	95.00%	95.00%
4	b19	F	75.00%	65.00%	70.00%	100.00%	95.00%	97.50%
5	c22	F	60.00%	95.00%	77.50%	95.00%	90.00%	92.50%
6	c26	M	75.00%	85.00%	80.00%	80.00%	90.00%	85.00%
7	c31	M	85.00%	65.00%	75.00%	80.00%	70.00%	75.00%
		Mean:	**72.14%**	**77.14%**	**74.64%**	**92.86%**	**89.29%**	**91.07%**
		S.D.:	***12.86%***	***14.10%***	***3.36%***	***9.06%***	***8.86%***	***8.27%***
LIN								
1	a09	F	70.00%	80.00%	75.00%	85.00%	90.00%	87.50%
2	a22	F	75.00%	90.00%	82.50%	65.00%	95.00%	80.00%
3	c13	F	65.00%	60.00%	62.50%	85.00%	90.00%	87.50%
4	c17	F	80.00%	55.00%	67.50%	80.00%	95.00%	87.50%

5	c19	F	65.00%	75.00%	70.00%	95.00%	95.00%	95.00%
6	a03	F	60.00%	75.00%	67.50%	85.00%	85.00%	85.00%
7	a21	M	70.00%	85.00%	77.50%	85.00%	80.00%	82.50%
8	a24	M	50.00%	50.00%	50.00%	80.00%	80.00%	80.00%
		Mean:	**66.88%**	**71.25%**	**69.06%**	**82.50%**	**88.75%**	**85.63%**
		S.D.:	***9.23%***	***14.58%***	***9.99%***	***8.45%***	***6.41%***	***4.96%***
COP								
1	c01	F	75.00%	90.00%	82.50%	75.00%	90.00%	82.50%
2	c05	F	80.00%	60.00%	70.00%	90.00%	60.00%	75.00%
3	c10	F	45.00%	30.00%	37.50%	65.00%	50.00%	57.50%
4	c20	F	60.00%	60.00%	60.00%	70.00%	70.00%	70.00%
5	c25	F	80.00%	80.00%	80.00%	85.00%	65.00%	75.00%
6	b12	F	85.00%	45.00%	65.00%	85.00%	15.00%	50.00%
7	b14	F	75.00%	75.00%	75.00%	80.00%	95.00%	87.50%
8	c34	M	80.00%	75.00%	77.50%	80.00%	85.00%	82.50%
9	a26	M	60.00%	65.00%	62.50%	60.00%	65.00%	62.50%
10	a29	M	75.00%	80.00%	77.50%	80.00%	80.00%	80.00%
11	a31	M	60.00%	65.00%	62.50%	55.00%	75.00%	65.00%
		Mean:	**70.45%**	**65.91%**	**68.18%**	**75.00%**	**68.18%**	**71.59%**
		S.D.:	***12.34%***	***17.15%***	***12.85%***	***11.18%***	***22.17%***	***11.74%***
CON								
1	a16	F	65.00%	55.00%	60.00%	80.00%	55.00%	67.50%
2	a38	F	75.00%	70.00%	72.50%	75.00%	60.00%	67.50%
3	b18	F	65.00%	70.00%	67.50%	50.00%	70.00%	60.00%
4	c07	F	75.00%	85.00%	80.00%	70.00%	80.00%	75.00%
5	a05	F	65.00%	65.00%	65.00%	70.00%	65.00%	67.50%
6	b10	F	65.00%	60.00%	62.50%	70.00%	50.00%	60.00%
7	a28	M	75.00%	60.00%	67.50%	65.00%	65.00%	65.00%
8	b24	M	75.00%	70.00%	72.50%	70.00%	75.00%	72.50%
9	c29	M	75.00%	65.00%	70.00%	70.00%	65.00%	67.50%
10	c33	M	75.00%	60.00%	67.50%	75.00%	65.00%	70.00%
		Mean:	**71.00%**	**66.00%**	**68.50%**	**69.50%**	**65.00%**	**67.25%**
		S.D.:	***5.16%***	***8.43%***	***5.68%***	***7.98%***	***8.82%***	***4.78%***

	Username	Gender	TG1 (/ɪ/)	TG1 (/i:/)	TG1 (avg)	TG2 (/ɪ/)	TG2 (/i:/)	TG2 (avg)	TG3 (/ɪ/)	TG3 (/i:/)	TG3 (avg)
HSP											
1	a01	F	75.00%	90.00%	82.50%	65.00%	65.00%	65.00%	85.00%	90.00%	87.50%
2	a02	F	60.00%	80.00%	70.00%	80.00%	85.00%	82.50%	90.00%	85.00%	87.50%
3	a11	F	60.00%	75.00%	67.50%	95.00%	55.00%	75.00%	90.00%	60.00%	75.00%
4	a37	F	95.00%	75.00%	85.00%	75.00%	85.00%	80.00%	75.00%	75.00%	75.00%
5	b02	F	90.00%	90.00%	90.00%	100.00%	95.00%	97.50%	85.00%	75.00%	80.00%
6	b04	F	85.00%	90.00%	87.50%	100.00%	90.00%	95.00%	90.00%	100.00%	95.00%
7	b38	M	80.00%	95.00%	87.50%	90.00%	100.00%	95.00%	80.00%	100.00%	90.00%
8	a25	M	65.00%	85.00%	75.00%	60.00%	70.00%	65.00%	85.00%	70.00%	77.50%
9	c38	M	65.00%	70.00%	67.50%	60.00%	65.00%	62.50%	95.00%	55.00%	75.00%
		Mean:	**75.00%**	**83.33%**	**79.17%**	**80.56%**	**78.89%**	**79.72%**	**86.11%**	**78.89%**	**82.50%**
		S.D.:	***13.23%***	***8.66%***	***9.19%***	***16.48%***	***15.57%***	***13.89%***	***6.01%***	***16.16%***	***7.60%***
HIP											
1	a17	F	70.00%	100.00%	85.00%	85.00%	100.00%	92.50%	90.00%	85.00%	87.50%
2	c02	F	70.00%	90.00%	80.00%	60.00%	55.00%	57.50%	65.00%	60.00%	62.50%
3	c03	F	80.00%	75.00%	77.50%	70.00%	65.00%	67.50%	70.00%	80.00%	75.00%
4	c04	F	55.00%	75.00%	65.00%	75.00%	75.00%	75.00%	75.00%	95.00%	85.00%
5	c06	F	80.00%	80.00%	80.00%	70.00%	85.00%	77.50%	75.00%	100.00%	87.50%
6	c35	M	95.00%	75.00%	85.00%	95.00%	90.00%	92.50%	95.00%	95.00%	95.00%
7	c36	M	80.00%	90.00%	85.00%	75.00%	85.00%	80.00%	95.00%	95.00%	95.00%
8	c37	M	70.00%	95.00%	82.50%	90.00%	95.00%	92.50%	100.00%	100.00%	100.00%
9	c27	M	80.00%	85.00%	82.50%	75.00%	90.00%	82.50%	95.00%	100.00%	97.50%
		Mean:	**75.56%**	**85.00%**	**80.28%**	**77.22%**	**82.22%**	**79.72%**	**84.44%**	**90.00%**	**87.22%**
		S.D.:	***11.02%***	***9.35%***	***6.31%***	***10.93%***	***14.60%***	***12.08%***	***13.10%***	***13.23%***	***12.02%***
HSN											
1	b01	F	90.00%	90.00%	90.00%	95.00%	95.00%	95.00%	100.00%	100.00%	100.00%
2	b03	F	75.00%	95.00%	85.00%	80.00%	85.00%	82.50%	85.00%	70.00%	77.50%
3	b05	F	95.00%	100.00%	97.50%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4	b06	F	85.00%	85.00%	85.00%	95.00%	85.00%	90.00%	95.00%	90.00%	92.50%
5	b07	F	80.00%	80.00%	80.00%	80.00%	85.00%	82.50%	75.00%	70.00%	72.50%
6	b08	F	75.00%	80.00%	77.50%	65.00%	75.00%	70.00%	90.00%	85.00%	87.50%
7	b25	M	100.00%	75.00%	87.50%	75.00%	60.00%	67.50%	70.00%	75.00%	72.50%
8	b21	M	70.00%	90.00%	80.00%	95.00%	70.00%	82.50%	90.00%	50.00%	70.00%
9	c39	M	80.00%	80.00%	80.00%	55.00%	90.00%	72.50%	50.00%	95.00%	72.50%
		Mean:	**83.33%**	**86.11%**	**84.72%**	**82.22%**	**82.78%**	**82.50%**	**83.89%**	**81.67%**	**82.78%**
		S.D.:	***10.00%***	***8.21%***	***6.31%***	***15.43%***	***12.53%***	***11.18%***	***16.35%***	***16.77%***	***12.34%***
HIN											
1	a04	F	90.00%	90.00%	90.00%	80.00%	65.00%	72.50%	85.00%	85.00%	85.00%
2	a10	F	95.00%	90.00%	92.50%	80.00%	75.00%	77.50%	50.00%	75.00%	62.50%

3	b11	F	80.00%	85.00%	82.50%	95.00%	80.00%	87.50%	100.00%	85.00%	92.50%
4	c08	F	90.00%	80.00%	85.00%	85.00%	95.00%	90.00%	95.00%	95.00%	95.00%
5	c09	F	75.00%	90.00%	82.50%	75.00%	75.00%	75.00%	90.00%	70.00%	80.00%
6	c15	F	85.00%	100.00%	92.50%	85.00%	95.00%	90.00%	90.00%	100.00%	95.00%
7	b37	M	80.00%	90.00%	85.00%	75.00%	80.00%	77.50%	75.00%	90.00%	82.50%
8	b39	M	85.00%	90.00%	87.50%	100.00%	80.00%	90.00%	100.00%	90.00%	95.00%
		Mean:	**85.00%**	**89.38%**	**87.19%**	**84.38%**	**80.63%**	**82.50%**	**85.63%**	**86.25%**	**85.94%**
		S.D.:	***6.55%***	***5.63%***	***4.11%***	***9.04%***	***10.16%***	***7.56%***	***16.57%***	***9.91%***	***11.25%***
LSP											
1	b16	F	65.00%	90.00%	77.50%	60.00%	95.00%	77.50%	95.00%	100.00%	97.50%
2	b17	F	65.00%	80.00%	72.50%	40.00%	90.00%	65.00%	60.00%	70.00%	65.00%
3	a14	F	50.00%	85.00%	67.50%	65.00%	80.00%	72.50%	95.00%	80.00%	87.50%
4	b20	F	90.00%	90.00%	90.00%	100.00%	100.00%	100.00%	95.00%	100.00%	97.50%
5	b29	M	60.00%	75.00%	67.50%	80.00%	65.00%	72.50%	80.00%	70.00%	75.00%
6	b35	M	75.00%	65.00%	70.00%	90.00%	70.00%	80.00%	80.00%	100.00%	90.00%
7	a23	M	45.00%	85.00%	65.00%	70.00%	90.00%	80.00%	60.00%	90.00%	75.00%
		Mean:	**64.29%**	**81.43%**	**72.86%**	**72.14%**	**84.29%**	**78.21%**	**80.71%**	**87.14%**	**83.93%**
		S.D.:	***15.12%***	***9.00%***	***8.59%***	***19.97%***	***13.05%***	***10.97%***	***15.66%***	***13.80%***	***12.49%***
LIP											
1	c12	F	55.00%	60.00%	57.50%	60.00%	60.00%	60.00%	80.00%	70.00%	75.00%
2	c18	F	50.00%	75.00%	62.50%	55.00%	65.00%	60.00%	75.00%	60.00%	67.50%
3	c21	F	55.00%	65.00%	60.00%	70.00%	90.00%	80.00%	85.00%	95.00%	90.00%
4	c23	F	55.00%	75.00%	65.00%	55.00%	100.00%	77.50%	80.00%	85.00%	82.50%
5	c24	F	65.00%	75.00%	70.00%	80.00%	90.00%	85.00%	90.00%	100.00%	95.00%
6	b27	M	90.00%	60.00%	75.00%	100.00%	55.00%	77.50%	95.00%	85.00%	90.00%
7	c32	M	70.00%	85.00%	77.50%	75.00%	60.00%	67.50%	80.00%	100.00%	90.00%
		Mean:	**62.86%**	**70.71%**	**66.79%**	**70.71%**	**74.29%**	**72.50%**	**83.57%**	**85.00%**	**84.29%**
		S.D.:	***13.80%***	***9.32%***	***7.60%***	***16.18%***	***18.35%***	***10.00%***	***6.90%***	***15.28%***	***9.87%***
LSN											
1	a12	F	95.00%	95.00%	95.00%	95.00%	100.00%	97.50%	100.00%	95.00%	97.50%
2	a15	F	85.00%	80.00%	82.50%	80.00%	85.00%	82.50%	90.00%	95.00%	92.50%
3	a19	F	60.00%	60.00%	60.00%	75.00%	70.00%	72.50%	75.00%	75.00%	75.00%
4	b19	F	60.00%	85.00%	72.50%	65.00%	80.00%	72.50%	90.00%	70.00%	80.00%
5	c22	F	60.00%	65.00%	62.50%	70.00%	75.00%	72.50%	85.00%	85.00%	85.00%
6	c26	M	60.00%	75.00%	67.50%	65.00%	80.00%	72.50%	75.00%	75.00%	75.00%
7	c31	M	70.00%	60.00%	65.00%	65.00%	80.00%	72.50%	75.00%	60.00%	67.50%
		Mean:	**70.00%**	**74.29%**	**72.14%**	**73.57%**	**81.43%**	**77.50%**	**84.29%**	**79.29%**	**81.79%**
		S.D.:	***14.43%***	***13.36%***	***12.54%***	***11.07%***	***9.45%***	***9.57%***	***9.76%***	***13.05%***	***10.58%***
LIN											
1	a09	F	65.00%	75.00%	70.00%	80.00%	70.00%	75.00%	80.00%	75.00%	77.50%
2	a22	F	60.00%	90.00%	75.00%	85.00%	80.00%	82.50%	65.00%	85.00%	75.00%
3	c13	F	55.00%	85.00%	70.00%	65.00%	60.00%	62.50%	60.00%	90.00%	75.00%

4	c17	F	75.00%	80.00%	77.50%	75.00%	65.00%	70.00%	85.00%	85.00%	85.00%
5	c19	F	75.00%	65.00%	70.00%	70.00%	70.00%	70.00%	90.00%	85.00%	87.50%
6	a03	F	60.00%	65.00%	62.50%	85.00%	55.00%	70.00%	65.00%	80.00%	72.50%
7	a21	M	55.00%	60.00%	57.50%	80.00%	80.00%	80.00%	90.00%	70.00%	80.00%
8	a24	M	70.00%	60.00%	65.00%	75.00%	85.00%	80.00%	75.00%	75.00%	75.00%
		Mean:	**64.38%**	**72.50%**	**68.44%**	**76.88%**	**70.63%**	**73.75%**	**76.25%**	**80.63%**	**78.44%**
		S.D.:	***8.21%***	***11.65%***	***6.54%***	***7.04%***	***10.50%***	***6.81%***	***11.88%***	***6.78%***	***5.33%***
COP											
1	c01	F	75.00%	80.00%	77.50%	80.00%	80.00%	80.00%	70.00%	75.00%	72.50%
2	c05	F	70.00%	60.00%	65.00%	50.00%	40.00%	45.00%	80.00%	60.00%	70.00%
3	c10	F	40.00%	65.00%	52.50%	75.00%	40.00%	57.50%	60.00%	55.00%	57.50%
4	c20	F	70.00%	70.00%	70.00%	60.00%	75.00%	67.50%	75.00%	70.00%	72.50%
5	c25	F	50.00%	65.00%	57.50%	75.00%	75.00%	75.00%	80.00%	85.00%	82.50%
6	b12	F	75.00%	10.00%	42.50%	95.00%	10.00%	52.50%	85.00%	25.00%	55.00%
7	b14	F	65.00%	80.00%	72.50%	80.00%	85.00%	82.50%	85.00%	100.00%	92.50%
8	c34	M	65.00%	85.00%	75.00%	65.00%	75.00%	70.00%	30.00%	100.00%	65.00%
9	a26	M	25.00%	50.00%	37.50%	30.00%	60.00%	45.00%	40.00%	60.00%	50.00%
10	a29	M	60.00%	65.00%	62.50%	75.00%	85.00%	80.00%	90.00%	75.00%	82.50%
11	a31	M	55.00%	65.00%	60.00%	60.00%	70.00%	65.00%	80.00%	75.00%	77.50%
		Mean:	**59.09%**	**63.18%**	**61.14%**	**67.73%**	**63.18%**	**65.45%**	**70.45%**	**70.91%**	**70.68%**
		S.D.:	***15.62%***	***20.28%***	***12.96%***	***17.52%***	***23.69%***	***13.78%***	***19.42%***	***21.31%***	***12.99%***
CON											
1	a16	F	60.00%	60.00%	60.00%	35.00%	65.00%	50.00%	80.00%	45.00%	62.50%
2	a38	F	70.00%	50.00%	60.00%	55.00%	85.00%	70.00%	75.00%	85.00%	80.00%
3	b18	F	70.00%	60.00%	65.00%	60.00%	75.00%	67.50%	75.00%	70.00%	72.50%
4	c07	F	30.00%	75.00%	52.50%	60.00%	75.00%	67.50%	50.00%	75.00%	62.50%
5	a05	F	50.00%	60.00%	55.00%	80.00%	70.00%	75.00%	50.00%	65.00%	57.50%
6	b10	F	50.00%	40.00%	45.00%	60.00%	65.00%	62.50%	65.00%	60.00%	62.50%
7	a28	M	45.00%	80.00%	62.50%	55.00%	70.00%	62.50%	55.00%	60.00%	57.50%
8	b24	M	80.00%	30.00%	55.00%	80.00%	35.00%	57.50%	70.00%	25.00%	47.50%
9	c29	M	60.00%	55.00%	57.50%	65.00%	85.00%	75.00%	75.00%	85.00%	80.00%
10	c33	M	40.00%	55.00%	47.50%	55.00%	70.00%	62.50%	65.00%	80.00%	72.50%
		Mean:	**55.50%**	**56.50%**	**56.00%**	**60.50%**	**69.50%**	**65.00%**	**66.00%**	**65.00%**	**65.50%**
		S.D.:	***15.36%***	***14.73%***	***6.37%***	***13.01%***	***14.03%***	***7.73%***	***11.01%***	***18.86%***	***10.53%***

	Username	Gender	Pre (/e/)	Pre (/æ /)	Pre (avg)	Post (/e/)	Post (/æ /)	Post (avg)
HSP								
1	a01	F	65.00%	60.00%	62.50%	100.00%	95.00%	97.50%
2	a02	F	65.00%	45.00%	55.00%	85.00%	95.00%	90.00%
3	a11	F	65.00%	50.00%	57.50%	90.00%	95.00%	92.50%
4	a37	F	70.00%	60.00%	65.00%	60.00%	95.00%	77.50%
5	b02	F	85.00%	65.00%	75.00%	100.00%	95.00%	97.50%
6	b04	F	80.00%	60.00%	70.00%	100.00%	95.00%	97.50%
7	b38	M	60.00%	50.00%	55.00%	100.00%	90.00%	95.00%
8	a25	M	45.00%	40.00%	42.50%	95.00%	90.00%	92.50%
9	c38	M	60.00%	50.00%	55.00%	85.00%	90.00%	87.50%
		Mean:	**66.11%**	**53.33%**	**59.72%**	**90.56%**	**93.33%**	**91.94%**
		S.D.:	***11.67%***	***8.29%***	***9.64%***	***13.10%***	***2.50%***	***6.47%***
HIP								
1	a17	F	50.00%	60.00%	55.00%	95.00%	100.00%	97.50%
2	c02	F	50.00%	45.00%	47.50%	100.00%	95.00%	97.50%
3	c03	F	55.00%	45.00%	50.00%	85.00%	95.00%	90.00%
4	c04	F	40.00%	25.00%	32.50%	90.00%	85.00%	87.50%
5	c06	F	80.00%	75.00%	77.50%	100.00%	95.00%	97.50%
6	c35	M	75.00%	45.00%	60.00%	95.00%	90.00%	92.50%
7	c36	M	50.00%	70.00%	60.00%	100.00%	95.00%	97.50%
8	c37	M	70.00%	60.00%	65.00%	90.00%	95.00%	92.50%
9	c27	M	80.00%	55.00%	67.50%	100.00%	80.00%	90.00%
		Mean:	**61.11%**	**53.33%**	**57.22%**	**95.00%**	**92.22%**	**93.61%**
		S.D.:	***15.16%***	***15.21%***	***13.02%***	***5.59%***	***6.18%***	***3.97%***
HSN								
1	b01	F	75.00%	65.00%	70.00%	100.00%	95.00%	97.50%
2	b03	F	70.00%	65.00%	67.50%	95.00%	95.00%	95.00%
3	b05	F	65.00%	60.00%	62.50%	100.00%	95.00%	97.50%
4	b06	F	35.00%	50.00%	42.50%	90.00%	95.00%	92.50%
5	b07	F	60.00%	70.00%	65.00%	100.00%	85.00%	92.50%
6	b08	F	50.00%	65.00%	57.50%	80.00%	75.00%	77.50%
7	b25	M	60.00%	75.00%	67.50%	75.00%	75.00%	75.00%
8	b21	M	70.00%	75.00%	72.50%	95.00%	95.00%	95.00%
9	c39	M	60.00%	75.00%	67.50%	75.00%	80.00%	77.50%
		Mean:	**60.56%**	**66.67%**	**63.61%**	**90.00%**	**87.78%**	**88.89%**
		S.D.:	***12.10%***	***8.29%***	***9.02%***	***10.61%***	***9.05%***	***9.36%***
HIN								
1	a04	F	65.00%	80.00%	72.50%	85.00%	85.00%	85.00%
2	a10	F	50.00%	45.00%	47.50%	100.00%	90.00%	95.00%

3	b11	F	70.00%	65.00%	67.50%	65.00%	70.00%	67.50%
4	c08	F	65.00%	65.00%	65.00%	95.00%	90.00%	92.50%
5	c09	F	35.00%	60.00%	47.50%	85.00%	85.00%	85.00%
6	c15	F	60.00%	35.00%	47.50%	95.00%	90.00%	92.50%
7	b37	M	60.00%	60.00%	60.00%	90.00%	85.00%	87.50%
8	b39	M	75.00%	65.00%	70.00%	90.00%	95.00%	92.50%
		Mean:	**60.00%**	**59.38%**	**59.69%**	**88.13%**	**86.25%**	**87.19%**
		S.D.:	***12.54%***	***13.74%***	***10.73%***	***10.67%***	***7.44%***	***8.81%***
LSP								
1	b16	F	65.00%	60.00%	62.50%	95.00%	90.00%	92.50%
2	b17	F	70.00%	75.00%	72.50%	95.00%	85.00%	90.00%
3	a14	F	60.00%	60.00%	60.00%	95.00%	100.00%	97.50%
4	b20	F	75.00%	60.00%	67.50%	100.00%	100.00%	100.00%
5	b29	M	65.00%	30.00%	47.50%	90.00%	90.00%	90.00%
6	b35	M	65.00%	60.00%	62.50%	100.00%	100.00%	100.00%
7	a23	M	50.00%	60.00%	55.00%	95.00%	95.00%	95.00%
		Mean:	**64.29%**	**57.86%**	**61.07%**	**95.71%**	**94.29%**	**95.00%**
		S.D.:	***7.87%***	***13.50%***	***8.15%***	***3.45%***	***6.07%***	***4.33%***
LIP								
1	c12	F	70.00%	40.00%	55.00%	85.00%	85.00%	85.00%
2	c18	F	65.00%	70.00%	67.50%	90.00%	100.00%	95.00%
3	c21	F	70.00%	65.00%	67.50%	100.00%	95.00%	97.50%
4	c23	F	60.00%	75.00%	67.50%	85.00%	100.00%	92.50%
5	c24	F	60.00%	60.00%	60.00%	100.00%	90.00%	95.00%
6	b27	M	55.00%	75.00%	65.00%	70.00%	100.00%	85.00%
7	c32	M	55.00%	60.00%	57.50%	95.00%	95.00%	95.00%
		Mean:	**62.14%**	**63.57%**	**62.86%**	**89.29%**	**95.00%**	**92.14%**
		S.D.:	***6.36%***	***12.15%***	***5.29%***	***10.58%***	***5.77%***	***5.09%***
LSN								
1	a12	F	60.00%	65.00%	62.50%	100.00%	100.00%	100.00%
2	a15	F	70.00%	90.00%	80.00%	100.00%	85.00%	92.50%
3	a19	F	75.00%	65.00%	70.00%	90.00%	85.00%	87.50%
4	b19	F	85.00%	50.00%	67.50%	100.00%	100.00%	100.00%
5	c22	F	55.00%	70.00%	62.50%	100.00%	95.00%	97.50%
6	c26	M	25.00%	40.00%	32.50%	80.00%	90.00%	85.00%
7	c31	M	45.00%	40.00%	42.50%	75.00%	95.00%	85.00%
		Mean:	**59.29%**	**60.00%**	**59.64%**	**92.14%**	**92.86%**	**92.50%**
		S.D.:	***20.09%***	***18.03%***	***16.48%***	***10.75%***	***6.36%***	***6.77%***
LIN								
1	a09	F	45.00%	65.00%	55.00%	95.00%	95.00%	95.00%
2	a22	F	65.00%	60.00%	62.50%	100.00%	95.00%	97.50%
3	c13	F	60.00%	25.00%	42.50%	85.00%	60.00%	72.50%

4	c17	F	60.00%	70.00%	65.00%	100.00%	95.00%	97.50%
5	c19	F	60.00%	65.00%	62.50%	95.00%	95.00%	95.00%
6	a03	F	80.00%	85.00%	82.50%	75.00%	85.00%	80.00%
7	a21	M	45.00%	55.00%	50.00%	95.00%	85.00%	90.00%
8	a24	M	30.00%	55.00%	42.50%	90.00%	80.00%	85.00%
		Mean:	**55.63%**	**60.00%**	**57.81%**	**91.88%**	**86.25%**	**89.06%**
		S.D.:	***15.22%***	***17.11%***	***13.33%***	***8.43%***	***12.17%***	***9.16%***
COP								
1	c01	F	45.00%	70.00%	57.50%	70.00%	80.00%	75.00%
2	c05	F	50.00%	35.00%	42.50%	60.00%	55.00%	57.50%
3	c10	F	40.00%	55.00%	47.50%	50.00%	70.00%	60.00%
4	c20	F	55.00%	15.00%	35.00%	60.00%	70.00%	65.00%
5	c25	F	75.00%	60.00%	67.50%	90.00%	70.00%	80.00%
6	b12	F	50.00%	25.00%	37.50%	80.00%	35.00%	57.50%
7	b14	F	80.00%	90.00%	85.00%	80.00%	90.00%	85.00%
8	c34	M	60.00%	65.00%	62.50%	80.00%	60.00%	70.00%
9	a26	M	50.00%	50.00%	50.00%	50.00%	55.00%	52.50%
10	a29	M	65.00%	65.00%	65.00%	60.00%	65.00%	62.50%
11	a31	M	60.00%	75.00%	67.50%	60.00%	85.00%	72.50%
		Mean:	**57.27%**	**55.00%**	**56.14%**	**67.27%**	**66.82%**	**67.05%**
		S.D.:	***12.32%***	***22.36%***	***15.18%***	***13.48%***	***15.54%***	***10.30%***
CON								
1	a16	F	75.00%	60.00%	67.50%	80.00%	25.00%	52.50%
2	a38	F	65.00%	30.00%	47.50%	60.00%	30.00%	45.00%
3	b18	F	55.00%	65.00%	60.00%	60.00%	70.00%	65.00%
4	c07	F	50.00%	60.00%	55.00%	40.00%	45.00%	42.50%
5	a05	F	50.00%	65.00%	57.50%	60.00%	55.00%	57.50%
6	b10	F	45.00%	45.00%	45.00%	35.00%	45.00%	40.00%
7	a28	M	45.00%	35.00%	40.00%	35.00%	40.00%	37.50%
8	b24	M	60.00%	50.00%	55.00%	45.00%	45.00%	45.00%
9	c29	M	60.00%	50.00%	55.00%	45.00%	20.00%	32.50%
10	c33	M	70.00%	70.00%	70.00%	70.00%	65.00%	67.50%
		Mean:	**57.50%**	**53.00%**	**55.25%**	**53.00%**	**44.00%**	**48.50%**
		S.D.:	***10.34%***	***13.37%***	***9.39%***	***15.31%***	***16.30%***	***11.74%***

Username		Gender	TG1 (/e/)	TG1 (/æ /)	TG1 (avg)	TG2 (/e/)	TG2 (/æ /)	TG2 (avg)	TG3 (/e/)	TG3 (/æ /)	TG3 (avg)
HSP											
1	a01	F	80.00%	95.00%	87.50%	100.00%	80.00%	90.00%	90.00%	100.00%	95.00%
2	a02	F	70.00%	75.00%	72.50%	60.00%	70.00%	65.00%	50.00%	85.00%	67.50%
3	a11	F	75.00%	80.00%	77.50%	75.00%	70.00%	72.50%	70.00%	100.00%	85.00%
4	a37	F	65.00%	75.00%	70.00%	60.00%	95.00%	77.50%	55.00%	90.00%	72.50%
5	b02	F	85.00%	75.00%	80.00%	100.00%	70.00%	85.00%	95.00%	100.00%	97.50%
6	b04	F	90.00%	60.00%	75.00%	95.00%	80.00%	87.50%	100.00%	100.00%	100.00%
7	b38	M	80.00%	60.00%	70.00%	100.00%	85.00%	92.50%	75.00%	100.00%	87.50%
8	a25	M	65.00%	90.00%	77.50%	70.00%	55.00%	62.50%	45.00%	75.00%	60.00%
9	c38	M	70.00%	80.00%	75.00%	60.00%	70.00%	65.00%	75.00%	70.00%	72.50%
		Mean:	**75.56%**	**76.67%**	**76.11%**	**80.00%**	**75.00%**	**77.50%**	**72.78%**	**91.11%**	**81.94%**
		S.D.:	***8.82%***	***11.73%***	***5.46%***	***18.54%***	***11.46%***	***11.73%***	***19.86%***	***11.93%***	***14.35%***
HIP											
1	a17	F	80.00%	75.00%	77.50%	80.00%	90.00%	85.00%	55.00%	100.00%	77.50%
2	c02	F	95.00%	85.00%	90.00%	70.00%	95.00%	82.50%	100.00%	80.00%	90.00%
3	c03	F	70.00%	90.00%	80.00%	80.00%	90.00%	85.00%	80.00%	85.00%	82.50%
4	c04	F	80.00%	95.00%	87.50%	75.00%	80.00%	77.50%	95.00%	70.00%	82.50%
5	c06	F	95.00%	90.00%	92.50%	85.00%	80.00%	82.50%	95.00%	100.00%	97.50%
6	c35	M	90.00%	90.00%	90.00%	80.00%	85.00%	82.50%	95.00%	100.00%	97.50%
7	c36	M	100.00%	90.00%	95.00%	80.00%	95.00%	87.50%	85.00%	85.00%	85.00%
8	c37	M	90.00%	85.00%	87.50%	80.00%	90.00%	85.00%	85.00%	100.00%	92.50%
9	c27	M	90.00%	80.00%	85.00%	95.00%	70.00%	82.50%	95.00%	90.00%	92.50%
		Mean:	**87.78%**	**86.67%**	**87.22%**	**80.56%**	**86.11%**	**83.33%**	**87.22%**	**90.00%**	**88.61%**
		S.D.:	***9.39%***	***6.12%***	***5.65%***	***6.82%***	***8.21%***	***2.80%***	***13.72%***	***10.90%***	***7.08%***
HSN											
1	b01	F	85.00%	75.00%	80.00%	85.00%	70.00%	77.50%	90.00%	95.00%	92.50%
2	b03	F	80.00%	65.00%	72.50%	90.00%	85.00%	87.50%	95.00%	100.00%	97.50%
3	b05	F	85.00%	90.00%	87.50%	100.00%	80.00%	90.00%	90.00%	100.00%	95.00%
4	b06	F	70.00%	55.00%	62.50%	70.00%	90.00%	80.00%	85.00%	90.00%	87.50%
5	b07	F	75.00%	90.00%	82.50%	70.00%	70.00%	70.00%	100.00%	75.00%	87.50%
6	b08	F	55.00%	65.00%	60.00%	65.00%	65.00%	65.00%	40.00%	70.00%	55.00%
7	b25	M	80.00%	60.00%	70.00%	70.00%	65.00%	67.50%	45.00%	70.00%	57.50%
8	b21	M	70.00%	50.00%	60.00%	85.00%	95.00%	90.00%	80.00%	75.00%	77.50%
9	c39	M	50.00%	60.00%	55.00%	45.00%	75.00%	60.00%	25.00%	90.00%	57.50%
		Mean:	**72.22%**	**67.78%**	**70.00%**	**75.56%**	**77.22%**	**76.39%**	**72.22%**	**85.00%**	**78.61%**
		S.D.:	***12.53%***	***14.39%***	***11.46%***	***16.29%***	***10.93%***	***11.33%***	***27.74%***	***12.50%***	***17.42%***
HIN											
1	a04	F	90.00%	75.00%	82.50%	80.00%	75.00%	77.50%	75.00%	80.00%	77.50%
2	a10	F	80.00%	90.00%	85.00%	90.00%	90.00%	90.00%	80.00%	65.00%	72.50%

3	b11	F	85.00%	90.00%	87.50%	80.00%	80.00%	80.00%	75.00%	80.00%	77.50%
4	c08	F	85.00%	80.00%	82.50%	75.00%	85.00%	80.00%	95.00%	85.00%	90.00%
5	c09	F	90.00%	80.00%	85.00%	80.00%	70.00%	75.00%	90.00%	85.00%	87.50%
6	c15	F	75.00%	75.00%	75.00%	85.00%	90.00%	87.50%	80.00%	100.00%	90.00%
7	b37	M	85.00%	85.00%	85.00%	90.00%	90.00%	90.00%	90.00%	80.00%	85.00%
8	b39	M	75.00%	85.00%	80.00%	80.00%	85.00%	82.50%	100.00%	100.00%	100.00%
		Mean:	**83.13%**	**82.50%**	**82.81%**	**82.50%**	**83.13%**	**82.81%**	**85.63%**	**84.38%**	**85.00%**
		S.D.:	***5.94%***	***5.98%***	***3.88%***	***5.35%***	***7.53%***	***5.74%***	***9.43%***	***11.48%***	***8.86%***
LSP											
1	b16	F	55.00%	70.00%	62.50%	80.00%	85.00%	82.50%	95.00%	90.00%	92.50%
2	b17	F	75.00%	60.00%	67.50%	65.00%	80.00%	72.50%	50.00%	65.00%	57.50%
3	a14	F	60.00%	65.00%	62.50%	70.00%	75.00%	72.50%	80.00%	95.00%	87.50%
4	b20	F	90.00%	65.00%	77.50%	85.00%	75.00%	80.00%	85.00%	80.00%	82.50%
5	b29	M	60.00%	75.00%	67.50%	75.00%	75.00%	75.00%	80.00%	90.00%	85.00%
6	b35	M	60.00%	80.00%	70.00%	50.00%	90.00%	70.00%	90.00%	85.00%	87.50%
7	a23	M	55.00%	65.00%	60.00%	45.00%	80.00%	62.50%	80.00%	75.00%	77.50%
		Mean:	**65.00%**	**68.57%**	**66.79%**	**67.14%**	**80.00%**	**73.57%**	**80.00%**	**82.86%**	**81.43%**
		S.D.:	***12.91%***	***6.90%***	***5.90%***	***14.96%***	***5.77%***	***6.59%***	***14.43%***	***10.35%***	***11.53%***
LIP											
1	c12	F	80.00%	55.00%	67.50%	60.00%	70.00%	65.00%	65.00%	55.00%	60.00%
2	c18	F	65.00%	60.00%	62.50%	65.00%	85.00%	75.00%	80.00%	90.00%	85.00%
3	c21	F	75.00%	55.00%	65.00%	80.00%	80.00%	80.00%	85.00%	95.00%	90.00%
4	c23	F	90.00%	65.00%	77.50%	90.00%	75.00%	82.50%	75.00%	75.00%	75.00%
5	c24	F	75.00%	70.00%	72.50%	85.00%	85.00%	85.00%	85.00%	90.00%	87.50%
6	b27	M	45.00%	90.00%	67.50%	30.00%	100.00%	65.00%	55.00%	80.00%	67.50%
7	c32	M	50.00%	60.00%	55.00%	70.00%	65.00%	67.50%	80.00%	55.00%	67.50%
		Mean:	**68.57%**	**65.00%**	**66.79%**	**68.57%**	**80.00%**	**74.29%**	**75.00%**	**77.14%**	**76.07%**
		S.D.:	***16.26%***	***12.25%***	***7.18%***	***20.15%***	***11.55%***	***8.50%***	***11.18%***	***16.55%***	***11.62%***
LSN											
1	a12	F	95.00%	30.00%	62.50%	75.00%	80.00%	77.50%	90.00%	85.00%	87.50%
2	a15	F	80.00%	75.00%	77.50%	95.00%	90.00%	92.50%	90.00%	75.00%	82.50%
3	a19	F	85.00%	50.00%	67.50%	65.00%	75.00%	70.00%	80.00%	80.00%	80.00%
4	b19	F	85.00%	50.00%	67.50%	80.00%	75.00%	77.50%	90.00%	95.00%	92.50%
5	c22	F	80.00%	45.00%	62.50%	80.00%	65.00%	72.50%	95.00%	100.00%	97.50%
6	c26	M	75.00%	45.00%	60.00%	80.00%	65.00%	72.50%	65.00%	60.00%	62.50%
7	c31	M	75.00%	60.00%	67.50%	75.00%	85.00%	80.00%	45.00%	75.00%	60.00%
		Mean:	**82.14%**	**50.71%**	**66.43%**	**78.57%**	**76.43%**	**77.50%**	**79.29%**	**81.43%**	**80.36%**
		S.D.:	***6.99%***	***13.97%***	***5.75%***	***9.00%***	***9.45%***	***7.50%***	***18.13%***	***13.45%***	***14.32%***
LIN											
1	a09	F	80.00%	55.00%	67.50%	100.00%	60.00%	80.00%	100.00%	100.00%	100.00%
2	a22	F	75.00%	45.00%	60.00%	80.00%	95.00%	87.50%	95.00%	100.00%	97.50%
3	c13	F	85.00%	30.00%	57.50%	80.00%	30.00%	55.00%	80.00%	40.00%	60.00%

4	c17	F	65.00%	80.00%	72.50%	45.00%	45.00%	45.00%	70.00%	85.00%	77.50%
5	c19	F	65.00%	35.00%	50.00%	85.00%	90.00%	87.50%	60.00%	55.00%	57.50%
6	a03	F	70.00%	75.00%	72.50%	75.00%	70.00%	72.50%	50.00%	60.00%	55.00%
7	a21	M	75.00%	60.00%	67.50%	90.00%	80.00%	85.00%	85.00%	85.00%	85.00%
8	a24	M	70.00%	70.00%	70.00%	70.00%	80.00%	75.00%	40.00%	65.00%	52.50%
		Mean:	**73.13%**	**56.25%**	**64.69%**	**78.13%**	**68.75%**	**73.44%**	**72.50%**	**73.75%**	**73.13%**
		S.D.:	***7.04%***	***18.47%***	***8.07%***	***16.24%***	***22.48%***	***15.69%***	***21.38%***	***22.00%***	***19.45%***
COP											
1	c01	F	85.00%	50.00%	67.50%	75.00%	60.00%	67.50%	70.00%	85.00%	77.50%
2	c05	F	55.00%	60.00%	57.50%	55.00%	70.00%	62.50%	60.00%	55.00%	57.50%
3	c10	F	65.00%	60.00%	62.50%	30.00%	45.00%	37.50%	75.00%	65.00%	70.00%
4	c20	F	55.00%	65.00%	60.00%	40.00%	55.00%	47.50%	40.00%	55.00%	47.50%
5	c25	F	85.00%	40.00%	62.50%	95.00%	80.00%	87.50%	75.00%	90.00%	82.50%
6	b12	F	70.00%	20.00%	45.00%	70.00%	45.00%	57.50%	80.00%	20.00%	50.00%
7	b14	F	95.00%	50.00%	72.50%	95.00%	100.00%	97.50%	100.00%	95.00%	97.50%
8	c34	M	60.00%	55.00%	57.50%	70.00%	65.00%	67.50%	55.00%	90.00%	72.50%
9	a26	M	55.00%	65.00%	60.00%	15.00%	70.00%	42.50%	55.00%	70.00%	62.50%
10	a29	M	100.00%	45.00%	72.50%	100.00%	80.00%	90.00%	95.00%	100.00%	97.50%
11	a31	M	75.00%	50.00%	62.50%	65.00%	55.00%	60.00%	90.00%	60.00%	75.00%
		Mean:	**72.73%**	**50.91%**	**61.82%**	**64.55%**	**65.91%**	**65.23%**	**72.27%**	**71.36%**	**71.82%**
		S.D.:	***16.49%***	***13.00%***	***7.67%***	***27.61%***	***16.56%***	***19.64%***	***18.62%***	***23.67%***	***16.85%***
CON											
1	a16	F	90.00%	40.00%	65.00%	95.00%	30.00%	62.50%	90.00%	35.00%	62.50%
2	a38	F	60.00%	65.00%	62.50%	50.00%	55.00%	52.50%	50.00%	65.00%	57.50%
3	b18	F	55.00%	65.00%	60.00%	40.00%	55.00%	47.50%	40.00%	55.00%	47.50%
4	c07	F	75.00%	40.00%	57.50%	45.00%	60.00%	52.50%	50.00%	65.00%	57.50%
5	a05	F	55.00%	55.00%	55.00%	75.00%	75.00%	75.00%	50.00%	75.00%	62.50%
6	b10	F	35.00%	90.00%	62.50%	75.00%	65.00%	70.00%	85.00%	90.00%	87.50%
7	a28	M	55.00%	50.00%	52.50%	45.00%	35.00%	40.00%	40.00%	60.00%	50.00%
8	b24	M	80.00%	30.00%	55.00%	90.00%	25.00%	57.50%	80.00%	30.00%	55.00%
9	c29	M	50.00%	50.00%	50.00%	45.00%	45.00%	45.00%	60.00%	45.00%	52.50%
10	c33	M	65.00%	60.00%	62.50%	75.00%	60.00%	67.50%	70.00%	60.00%	65.00%
		Mean:	**62.00%**	**54.50%**	**58.25%**	**63.50%**	**50.50%**	**57.00%**	**61.50%**	**58.00%**	**59.75%**
		S.D.:	***16.02%***	***16.91%***	***5.01%***	***20.69%***	***16.24%***	***11.53%***	***18.57%***	***17.98%***	***11.27%***

	Username	Gender	Pre (/ʊ/)	Pre (/uː/)	Pre (avg)	Post (/ʊ/)	Post (/uː/)	Post (avg)
HSP								
1	a01	F	45.00%	75.00%	60.00%	65.00%	75.00%	70.00%
2	a02	F	55.00%	90.00%	72.50%	85.00%	85.00%	85.00%
3	a11	F	45.00%	55.00%	50.00%	75.00%	85.00%	80.00%
4	a37	F	65.00%	55.00%	60.00%	90.00%	85.00%	87.50%
5	b02	F	85.00%	70.00%	77.50%	100.00%	80.00%	90.00%
6	b04	F	80.00%	80.00%	80.00%	85.00%	90.00%	87.50%
7	b38	M	95.00%	80.00%	87.50%	80.00%	80.00%	80.00%
8	a25	M	60.00%	80.00%	70.00%	65.00%	80.00%	72.50%
9	c38	M	75.00%	55.00%	65.00%	70.00%	75.00%	72.50%
		Mean:	**67.22%**	**71.11%**	**69.17%**	**79.44%**	**81.67%**	**80.56%**
		S.D.:	***17.70%***	***13.18%***	***11.66%***	***11.84%***	***5.00%***	***7.48%***
HIP								
1	a17	F	80.00%	85.00%	82.50%	90.00%	90.00%	90.00%
2	c02	F	55.00%	50.00%	52.50%	95.00%	100.00%	97.50%
3	c03	F	70.00%	65.00%	67.50%	75.00%	100.00%	87.50%
4	c04	F	60.00%	85.00%	72.50%	90.00%	80.00%	85.00%
5	c06	F	65.00%	70.00%	67.50%	90.00%	90.00%	90.00%
6	c35	M	60.00%	50.00%	55.00%	80.00%	90.00%	85.00%
7	c36	M	50.00%	80.00%	65.00%	100.00%	80.00%	90.00%
8	c37	M	80.00%	75.00%	77.50%	95.00%	100.00%	97.50%
9	c27	M	70.00%	75.00%	72.50%	80.00%	95.00%	87.50%
		Mean:	**65.56%**	**70.56%**	**68.06%**	**88.33%**	**91.67%**	**90.00%**
		S.D.:	***10.44%***	***13.33%***	***9.75%***	***8.29%***	***7.91%***	***4.68%***
HSN								
1	b01	F	70.00%	70.00%	70.00%	90.00%	100.00%	95.00%
2	b03	F	55.00%	45.00%	50.00%	85.00%	70.00%	77.50%
3	b05	F	75.00%	90.00%	82.50%	95.00%	100.00%	97.50%
4	b06	F	75.00%	80.00%	77.50%	85.00%	80.00%	82.50%
5	b07	F	55.00%	65.00%	60.00%	60.00%	85.00%	72.50%
6	b08	F	70.00%	80.00%	75.00%	80.00%	80.00%	80.00%
7	b25	M	60.00%	65.00%	62.50%	80.00%	75.00%	77.50%
8	b21	M	75.00%	75.00%	75.00%	75.00%	90.00%	82.50%
9	c39	M	60.00%	60.00%	60.00%	80.00%	90.00%	85.00%
		Mean:	**66.11%**	**70.00%**	**68.06%**	**81.11%**	**85.56%**	**83.33%**
		S.D.:	***8.58%***	***13.23%***	***10.52%***	***9.93%***	***10.44%***	***8.20%***
HIN								
1	a04	F	65.00%	60.00%	62.50%	70.00%	95.00%	82.50%
2	a10	F	75.00%	60.00%	67.50%	90.00%	70.00%	80.00%

3	b11	F	60.00%	65.00%	62.50%	70.00%	75.00%	72.50%
4	c08	F	65.00%	90.00%	77.50%	85.00%	90.00%	87.50%
5	c09	F	60.00%	55.00%	57.50%	80.00%	80.00%	80.00%
6	c15	F	60.00%	65.00%	62.50%	100.00%	100.00%	100.00%
7	b37	M	60.00%	60.00%	60.00%	85.00%	85.00%	85.00%
8	b39	M	85.00%	75.00%	80.00%	100.00%	95.00%	97.50%
		Mean:	**66.25%**	**66.25%**	**66.25%**	**85.00%**	**86.25%**	**85.63%**
		S.D.:	***9.16%***	***11.26%***	***8.24%***	***11.65%***	***10.61%***	***9.23%***
LSP								
1	b16	F	55.00%	65.00%	60.00%	100.00%	100.00%	100.00%
2	b17	F	85.00%	65.00%	75.00%	90.00%	70.00%	80.00%
3	a14	F	65.00%	60.00%	62.50%	95.00%	100.00%	97.50%
4	b20	F	55.00%	70.00%	62.50%	100.00%	100.00%	100.00%
5	b29	M	60.00%	65.00%	62.50%	90.00%	95.00%	92.50%
6	b35	M	50.00%	85.00%	67.50%	100.00%	100.00%	100.00%
7	a23	M	80.00%	60.00%	70.00%	95.00%	85.00%	90.00%
		Mean:	**64.29%**	**67.14%**	**65.71%**	**95.71%**	**92.86%**	**94.29%**
		S.D.:	***13.36%***	***8.59%***	***5.35%***	***4.50%***	***11.50%***	***7.46%***
LIP								
1	c12	F	80.00%	55.00%	67.50%	85.00%	90.00%	87.50%
2	c18	F	40.00%	65.00%	52.50%	65.00%	90.00%	77.50%
3	c21	F	80.00%	55.00%	67.50%	95.00%	95.00%	95.00%
4	c23	F	75.00%	80.00%	77.50%	80.00%	95.00%	87.50%
5	c24	F	65.00%	80.00%	72.50%	95.00%	70.00%	82.50%
6	b27	M	45.00%	65.00%	55.00%	55.00%	75.00%	65.00%
7	c32	M	85.00%	80.00%	82.50%	100.00%	90.00%	95.00%
		Mean:	**67.14%**	**68.57%**	**67.86%**	**82.14%**	**86.43%**	**84.29%**
		S.D.:	***17.99%***	***11.44%***	***11.03%***	***16.80%***	***9.88%***	***10.58%***
LSN								
1	a12	F	60.00%	60.00%	60.00%	100.00%	100.00%	100.00%
2	a15	F	90.00%	80.00%	85.00%	100.00%	100.00%	100.00%
3	a19	F	60.00%	45.00%	52.50%	100.00%	100.00%	100.00%
4	b19	F	50.00%	60.00%	55.00%	100.00%	100.00%	100.00%
5	c22	F	55.00%	85.00%	70.00%	50.00%	85.00%	67.50%
6	c26	M	70.00%	75.00%	72.50%	95.00%	75.00%	85.00%
7	c31	M	85.00%	85.00%	85.00%	100.00%	85.00%	92.50%
		Mean:	**67.14%**	**70.00%**	**68.57%**	**92.14%**	**92.14%**	**92.14%**
		S.D.:	***15.24%***	***15.28%***	***13.37%***	***18.68%***	***10.35%***	***12.28%***
LIN								
1	a09	F	70.00%	80.00%	75.00%	90.00%	100.00%	95.00%
2	a22	F	40.00%	60.00%	50.00%	80.00%	85.00%	82.50%
3	c13	F	50.00%	80.00%	65.00%	95.00%	100.00%	97.50%

4	c17	F	70.00%	70.00%	70.00%	80.00%	85.00%	82.50%
5	c19	F	40.00%	65.00%	52.50%	100.00%	100.00%	100.00%
6	a03	F	50.00%	75.00%	62.50%	55.00%	75.00%	65.00%
7	a21	M	50.00%	85.00%	67.50%	90.00%	100.00%	95.00%
8	a24	M	70.00%	65.00%	67.50%	90.00%	85.00%	87.50%
		Mean:	**55.00%**	**72.50%**	**63.75%**	**85.00%**	**91.25%**	**88.13%**
		S.D.:	***13.09%***	***8.86%***	***8.56%***	***13.89%***	***9.91%***	***11.48%***
COP								
1	c01	F	65.00%	65.00%	72.50%	65.00%	80.00%	65.00%
2	c05	F	75.00%	60.00%	62.50%	60.00%	65.00%	67.50%
3	c10	F	85.00%	65.00%	52.50%	60.00%	45.00%	75.00%
4	c20	F	85.00%	65.00%	65.00%	75.00%	55.00%	75.00%
5	c25	F	60.00%	60.00%	65.00%	65.00%	65.00%	60.00%
6	b12	F	60.00%	75.00%	52.50%	60.00%	45.00%	67.50%
7	b14	F	70.00%	70.00%	77.50%	75.00%	80.00%	70.00%
8	c34	M	60.00%	60.00%	62.50%	65.00%	60.00%	60.00%
9	a26	M	65.00%	85.00%	47.50%	40.00%	55.00%	75.00%
10	a29	M	70.00%	75.00%	72.50%	80.00%	65.00%	72.50%
11	a31	M	60.00%	75.00%	72.50%	65.00%	80.00%	67.50%
		Mean:	**68.64%**	**68.64%**	**63.86%**	**64.55%**	**63.18%**	**68.64%**
		S.D.:	***9.51%***	***8.09%***	***9.71%***	***10.60%***	***12.90%***	***5.52%***
CON								
1	a16	F	75.00%	50.00%	62.50%	85.00%	50.00%	67.50%
2	a38	F	55.00%	85.00%	70.00%	70.00%	60.00%	65.00%
3	b18	F	85.00%	65.00%	75.00%	75.00%	65.00%	70.00%
4	c07	F	65.00%	65.00%	65.00%	70.00%	55.00%	62.50%
5	a05	F	55.00%	80.00%	67.50%	50.00%	75.00%	62.50%
6	b10	F	55.00%	85.00%	70.00%	60.00%	65.00%	62.50%
7	a28	M	70.00%	80.00%	75.00%	70.00%	70.00%	70.00%
8	b24	M	60.00%	70.00%	65.00%	50.00%	70.00%	60.00%
9	c29	M	60.00%	55.00%	57.50%	65.00%	50.00%	57.50%
10	c33	M	55.00%	75.00%	65.00%	50.00%	65.00%	57.50%
		Mean:	**63.50%**	**71.00%**	**67.25%**	**64.50%**	**62.50%**	**63.50%**
		S.D.:	***10.29%***	***12.20%***	***5.46%***	***11.89%***	***8.58%***	***4.59%***

Username		Gender	TG1 (/ʊ/)	TG1 (/u:/)	TG1 (avg)	TG2 (/ʊ/)	TG2 (/u:/)	TG2 (avg)	TG3 (/ʊ/)	TG3 (/u:/)	TG3 (avg)
HSP											
1	a01	F	75.00%	85.00%	80.00%	85.00%	80.00%	82.50%	45.00%	70.00%	57.50%
2	a02	F	85.00%	75.00%	80.00%	75.00%	85.00%	80.00%	60.00%	65.00%	62.50%
3	a11	F	70.00%	85.00%	77.50%	65.00%	55.00%	60.00%	60.00%	85.00%	72.50%
4	a37	F	60.00%	70.00%	65.00%	60.00%	70.00%	65.00%	70.00%	70.00%	70.00%
5	b02	F	75.00%	80.00%	77.50%	75.00%	100.00%	87.50%	60.00%	75.00%	67.50%
6	b04	F	80.00%	75.00%	77.50%	90.00%	90.00%	90.00%	75.00%	85.00%	80.00%
7	b38	M	60.00%	85.00%	72.50%	90.00%	80.00%	85.00%	65.00%	75.00%	70.00%
8	a25	M	90.00%	85.00%	87.50%	40.00%	90.00%	65.00%	55.00%	75.00%	65.00%
9	c38	M	65.00%	70.00%	67.50%	55.00%	80.00%	67.50%	65.00%	70.00%	67.50%
		Mean:	**73.33%**	**78.89%**	**76.11%**	**70.56%**	**81.11%**	**75.83%**	**61.67%**	**74.44%**	**68.06%**
		S.D.:	***10.61%***	***6.51%***	***6.86%***	***17.04%***	***12.94%***	***11.39%***	***8.66%***	***6.82%***	***6.35%***
HIP											
1	a17	F	50.00%	100.00%	75.00%	85.00%	85.00%	85.00%	70.00%	75.00%	72.50%
2	c02	F	65.00%	80.00%	72.50%	90.00%	85.00%	87.50%	70.00%	100.00%	85.00%
3	c03	F	90.00%	90.00%	90.00%	80.00%	90.00%	85.00%	60.00%	55.00%	57.50%
4	c04	F	80.00%	75.00%	77.50%	95.00%	55.00%	75.00%	70.00%	100.00%	85.00%
5	c06	F	85.00%	85.00%	85.00%	65.00%	90.00%	77.50%	75.00%	95.00%	85.00%
6	c35	M	65.00%	90.00%	77.50%	80.00%	90.00%	85.00%	75.00%	100.00%	87.50%
7	c36	M	70.00%	90.00%	80.00%	75.00%	65.00%	70.00%	80.00%	100.00%	90.00%
8	c37	M	95.00%	85.00%	90.00%	100.00%	100.00%	100.00%	70.00%	90.00%	80.00%
9	c27	M	90.00%	90.00%	90.00%	60.00%	40.00%	50.00%	65.00%	70.00%	67.50%
		Mean:	**76.67%**	**87.22%**	**81.94%**	**81.11%**	**77.78%**	**79.44%**	**70.56%**	**87.22%**	**78.89%**
		S.D.:	***15.00%***	***7.12%***	***6.93%***	***13.18%***	***19.86%***	***13.96%***	***5.83%***	***16.60%***	***10.83%***
HSN											
1	b01	F	85.00%	100.00%	92.50%	95.00%	95.00%	95.00%	65.00%	100.00%	82.50%
2	b03	F	85.00%	90.00%	87.50%	100.00%	100.00%	100.00%	90.00%	80.00%	85.00%
3	b05	F	80.00%	75.00%	77.50%	80.00%	75.00%	77.50%	65.00%	85.00%	75.00%
4	b06	F	80.00%	75.00%	77.50%	65.00%	70.00%	67.50%	80.00%	70.00%	75.00%
5	b07	F	90.00%	75.00%	82.50%	90.00%	60.00%	75.00%	75.00%	90.00%	82.50%
6	b08	F	80.00%	80.00%	80.00%	50.00%	75.00%	62.50%	95.00%	70.00%	82.50%
7	b25	M	90.00%	80.00%	85.00%	95.00%	60.00%	77.50%	80.00%	90.00%	85.00%
8	b21	M	95.00%	85.00%	90.00%	100.00%	95.00%	97.50%	70.00%	85.00%	77.50%
9	c39	M	80.00%	80.00%	80.00%	75.00%	85.00%	80.00%	75.00%	80.00%	77.50%
		Mean:	**85.00%**	**82.22%**	**83.61%**	**83.33%**	**79.44%**	**81.39%**	**77.22%**	**83.33%**	**80.28%**
		S.D.:	***5.59%***	***8.33%***	***5.46%***	***17.32%***	***15.09%***	***13.29%***	***10.34%***	***9.68%***	***4.04%***
HIN											
1	a04	F	90.00%	75.00%	82.50%	60.00%	60.00%	60.00%	65.00%	90.00%	77.50%
2	a10	F	90.00%	70.00%	80.00%	90.00%	65.00%	77.50%	85.00%	95.00%	90.00%

3	b11	F	85.00%	80.00%	82.50%	90.00%	100.00%	95.00%	80.00%	75.00%	77.50%
4	c08	F	85.00%	75.00%	80.00%	90.00%	90.00%	90.00%	75.00%	90.00%	82.50%
5	c09	F	90.00%	75.00%	82.50%	90.00%	75.00%	82.50%	75.00%	80.00%	77.50%
6	c15	F	90.00%	80.00%	85.00%	100.00%	95.00%	97.50%	70.00%	95.00%	82.50%
7	b37	M	85.00%	85.00%	85.00%	80.00%	55.00%	67.50%	80.00%	80.00%	80.00%
8	b39	M	65.00%	80.00%	72.50%	100.00%	100.00%	100.00%	65.00%	85.00%	75.00%
		Mean:	**85.00%**	**77.50%**	**81.25%**	**87.50%**	**80.00%**	**83.75%**	**74.38%**	**86.25%**	**80.31%**
		S.D.:	***8.45%***	***4.63%***	***4.01%***	***12.82%***	***18.52%***	***14.58%***	***7.29%***	***7.44%***	***4.71%***
LSP											
1	b16	F	40.00%	90.00%	65.00%	100.00%	90.00%	95.00%	75.00%	100.00%	87.50%
2	b17	F	55.00%	70.00%	62.50%	40.00%	65.00%	52.50%	65.00%	90.00%	77.50%
3	a14	F	55.00%	35.00%	45.00%	70.00%	60.00%	65.00%	65.00%	75.00%	70.00%
4	b20	F	60.00%	85.00%	72.50%	90.00%	100.00%	95.00%	75.00%	100.00%	87.50%
5	b29	M	80.00%	75.00%	77.50%	70.00%	65.00%	67.50%	75.00%	65.00%	70.00%
6	b35	M	70.00%	70.00%	70.00%	100.00%	85.00%	92.50%	65.00%	95.00%	80.00%
7	a23	M	45.00%	65.00%	55.00%	75.00%	95.00%	85.00%	65.00%	50.00%	57.50%
		Mean:	**57.86%**	**70.00%**	**63.93%**	**77.86%**	**80.00%**	**78.93%**	**69.29%**	**82.14%**	**75.71%**
		S.D.:	***13.80%***	***17.80%***	***11.07%***	***21.19%***	***16.33%***	***17.13%***	***5.35%***	***19.33%***	***10.77%***
LIP											
1	c12	F	55.00%	90.00%	72.50%	80.00%	85.00%	82.50%	40.00%	65.00%	52.50%
2	c18	F	65.00%	75.00%	70.00%	50.00%	75.00%	62.50%	40.00%	85.00%	62.50%
3	c21	F	65.00%	60.00%	62.50%	55.00%	25.00%	40.00%	70.00%	75.00%	72.50%
4	c23	F	60.00%	80.00%	70.00%	85.00%	95.00%	90.00%	55.00%	90.00%	72.50%
5	c24	F	45.00%	65.00%	55.00%	80.00%	80.00%	80.00%	60.00%	70.00%	65.00%
6	b27	M	55.00%	80.00%	67.50%	55.00%	55.00%	55.00%	50.00%	90.00%	70.00%
7	c32	M	55.00%	80.00%	67.50%	60.00%	65.00%	62.50%	75.00%	75.00%	75.00%
		Mean:	**57.14%**	**75.71%**	**66.43%**	**66.43%**	**68.57%**	**67.50%**	**55.71%**	**78.57%**	**67.14%**
		S.D.:	***6.99%***	***10.18%***	***5.93%***	***14.64%***	***23.22%***	***17.56%***	***13.67%***	***9.88%***	***7.83%***
LSN											
1	a12	F	80.00%	85.00%	82.50%	100.00%	95.00%	97.50%	90.00%	100.00%	95.00%
2	a15	F	50.00%	80.00%	65.00%	90.00%	95.00%	92.50%	50.00%	70.00%	60.00%
3	a19	F	80.00%	80.00%	80.00%	75.00%	75.00%	75.00%	70.00%	80.00%	75.00%
4	b19	F	75.00%	85.00%	80.00%	85.00%	75.00%	80.00%	65.00%	80.00%	72.50%
5	c22	F	50.00%	90.00%	70.00%	70.00%	95.00%	82.50%	45.00%	80.00%	62.50%
6	c26	M	60.00%	90.00%	75.00%	70.00%	85.00%	77.50%	80.00%	70.00%	75.00%
7	c31	M	75.00%	80.00%	77.50%	90.00%	100.00%	95.00%	50.00%	90.00%	70.00%
		Mean:	**67.14%**	**84.29%**	**75.71%**	**82.86%**	**88.57%**	**85.71%**	**64.29%**	**81.43%**	**72.86%**
		S.D.:	***13.50%***	***4.50%***	***6.24%***	***11.50%***	***10.29%***	***9.10%***	***16.94%***	***10.69%***	***11.40%***
LIN											
1	a09	F	20.00%	80.00%	50.00%	50.00%	60.00%	55.00%	60.00%	100.00%	80.00%
2	a22	F	45.00%	85.00%	65.00%	100.00%	70.00%	85.00%	60.00%	95.00%	77.50%
3	c13	F	60.00%	70.00%	65.00%	45.00%	80.00%	62.50%	45.00%	90.00%	67.50%

4	c17	F	65.00%	80.00%	72.50%	60.00%	85.00%	72.50%	65.00%	75.00%	70.00%
5	c19	F	60.00%	75.00%	67.50%	35.00%	50.00%	42.50%	85.00%	55.00%	70.00%
6	a03	F	50.00%	75.00%	62.50%	55.00%	65.00%	60.00%	55.00%	65.00%	60.00%
7	a21	M	65.00%	80.00%	72.50%	60.00%	90.00%	75.00%	70.00%	75.00%	72.50%
8	a24	M	75.00%	90.00%	82.50%	70.00%	80.00%	75.00%	80.00%	95.00%	87.50%
		Mean:	**55.00%**	**79.38%**	**67.19%**	**59.38%**	**72.50%**	**65.94%**	**65.00%**	**81.25%**	**73.13%**
		S.D.:	***16.90%***	***6.23%***	***9.40%***	***19.54%***	***13.63%***	***13.56%***	***13.09%***	***16.20%***	***8.43%***
COP											
1	c01	F	25.00%	80.00%	52.50%	40.00%	45.00%	42.50%	45.00%	75.00%	60.00%
2	c05	F	55.00%	70.00%	62.50%	65.00%	90.00%	77.50%	55.00%	60.00%	57.50%
3	c10	F	50.00%	45.00%	47.50%	70.00%	55.00%	62.50%	45.00%	50.00%	47.50%
4	c20	F	45.00%	60.00%	52.50%	85.00%	75.00%	80.00%	55.00%	70.00%	62.50%
5	c25	F	30.00%	35.00%	32.50%	60.00%	75.00%	67.50%	65.00%	70.00%	67.50%
6	b12	F	45.00%	45.00%	45.00%	95.00%	0.00%	47.50%	95.00%	5.00%	50.00%
7	b14	F	60.00%	65.00%	62.50%	65.00%	75.00%	70.00%	45.00%	85.00%	65.00%
8	c34	M	40.00%	80.00%	60.00%	75.00%	90.00%	82.50%	65.00%	65.00%	65.00%
9	a26	M	45.00%	70.00%	57.50%	25.00%	75.00%	50.00%	30.00%	75.00%	52.50%
10	a29	M	65.00%	90.00%	77.50%	50.00%	85.00%	67.50%	60.00%	100.00%	80.00%
11	a31	M	40.00%	60.00%	50.00%	55.00%	55.00%	55.00%	60.00%	80.00%	70.00%
		Mean:	**45.45%**	**63.64%**	**54.55%**	**62.27%**	**65.45%**	**63.86%**	**56.36%**	**66.82%**	**61.59%**
		S.D.:	***11.93%***	***16.90%***	***11.61%***	***19.79%***	***26.22%***	***13.62%***	***16.60%***	***24.32%***	***9.50%***
CON											
1	a16	F	45.00%	75.00%	60.00%	70.00%	60.00%	65.00%	50.00%	50.00%	50.00%
2	a38	F	65.00%	60.00%	62.50%	60.00%	50.00%	55.00%	45.00%	65.00%	55.00%
3	b18	F	45.00%	60.00%	52.50%	85.00%	75.00%	80.00%	55.00%	70.00%	62.50%
4	c07	F	65.00%	65.00%	65.00%	75.00%	75.00%	75.00%	35.00%	70.00%	52.50%
5	a05	F	50.00%	65.00%	57.50%	55.00%	60.00%	57.50%	55.00%	80.00%	67.50%
6	b10	F	75.00%	15.00%	45.00%	65.00%	60.00%	62.50%	40.00%	100.00%	70.00%
7	a28	M	50.00%	75.00%	62.50%	60.00%	100.00%	80.00%	80.00%	55.00%	67.50%
8	b24	M	45.00%	75.00%	60.00%	50.00%	30.00%	40.00%	60.00%	55.00%	57.50%
9	c29	M	60.00%	55.00%	57.50%	35.00%	75.00%	55.00%	60.00%	65.00%	62.50%
10	c33	M	10.00%	85.00%	47.50%	20.00%	55.00%	37.50%	60.00%	75.00%	67.50%
		Mean:	**51.00%**	**63.00%**	**57.00%**	**57.50%**	**64.00%**	**60.75%**	**54.00%**	**68.50%**	**61.25%**
		S.D.:	***17.76%***	***19.18%***	***6.65%***	***19.04%***	***18.68%***	***14.96%***	***12.65%***	***14.54%***	***7.10%***

APPENDIX J
RESULTS OF THE PERCEPTION DATA SUBMITTED TO 5-WAY ANOVA (PRETEST VS. POSTTEST)

Between-subjects factor:

Perception Training (HVPT, LVTP, none)

Production Training (with, without)

Perceptual Training Intensity (standard, intensive, none)

Within-subjects factor (repeated-measures):

Test (pretest, posttest)

Vowel (front, mid, back)

Main Effects/Interactions	**df (between-subject)**	**df (within-subject)**	**Mean Square**	**F**	**Sig.**
Test	1	75	34050.425	543.147	.000
Test * Perception Training	1	75	77.710	1.240	.269
Test * Production Training	1	75	329.959	5.263	.025
Test * Perceptual Training Intensity	1	75	42.516	.678	.413
Test * Perception Training * Production Training	1	75	28.825	.460	.500
Test * Perception Training * Perceptual Training Intensity	1	75	452.342	7.215	.009
Test * Production Training * Perceptual Training Intensity	1	75	.000	.000	.998
Test * Perception Training * Production Training * Perceptual Training Intensity	1	75	187.114	2.985	.088
Vowel	2	150	1642.937	20.736	.000
Vowel * Perception Training	2	150	88.082	1.112	.332
Vowel * Production Training	2	150	211.934	2.675	.072
Vowel * Perceptual Training Intensity	2	150	3.513	.044	.957
Vowel * Perception Training * Production Training	2	150	68.202	.861	.425

Vowel * Perception Training * Perceptual Training Intensity	2	150	86.734	1.095	.337
Vowel * Production Training * Perceptual Training Intensity	2	150	5.771	.073	.930
Vowel * Perception Training * Production Training * Perceptual Training Intensity	2	150	38.167	.482	.619
Test * Vowel	2	150	1173.305	20.688	.000
Test * Vowel * Perception Training	2	150	144.544	2.549	.082
Test * Vowel * Production Training	2	150	170.347	3.004	.053
Test * Vowel * Perceptual Training Intensity	2	150	20.770	.366	.694
Test * Vowel * Perception Training * Production Training	2	150	61.846	1.090	.339
Test * Vowel * Perception Training * Perceptual Training Intensity	2	150	27.083	.478	.621
Test * Vowel * Production Training * Perceptual Training Intensity	2	150	25.055	.442	.644
Test * Vowel * Perception Training * Production Training * Perceptual Training Intensity	2	150	27.061	.477	.621
Perception Training	1	75	17.214	.111	.740
Production Training	1	75	325.483	2.100	.151
Perceptual Training Intensity	1	75	177.118	1.143	.288
Perception Training * Production Training	1	75	18.025	.116	.734
Perception Training * Perceptual Training Intensity	1	75	317.477	2.049	.156
Production Training * Perceptual Training Intensity	1	75	71.317	.460	.500
Perception Training * Production Training * Perceptual Training Intensity	1	75	1.513	.010	.922

Results of the Perception Data submitted to 5-way ANOVA (Differences between three TGs & Pretest)

Between-subjects factor:

Perception Training (HVPT, LVTP, none)

Production Training (with, without)

Perceptual Training Intensity (standard, intensive, none)

Within-subjects factor (repeated-measures):

Test (TG1, TG2, TG3)

Vowel (front, mid, back)

Main Effects / Interactions	df (between-subject)	df (within-subject)	Mean Square	F	Sig.
Test	2	150	1997.652	23.095	.000
Test * Perception Training	2	150	971.947	11.237	.000
Test * Production Training	2	150	109.948	1.271	.284
Test * Perceptual Training Intensity	2	150	148.489	1.717	.183
Test * Perception Training * Production Training	2	150	23.556	.272	.762
Test * Perception Training * Perceptual Training Intensity	2	150	15.535	.180	.836
Test * Production Training * Perceptual Training Intensity	2	150	3.171	.037	.964
Test * Perception Training * Production Training * Perceptual Training Intensity	2	150	64.351	.744	.477
Vowel	2	150	7151.193	20.198	.000
Vowel * Perception Training	2	150	58.565	.165	.848
Vowel * Production Training	2	150	568.948	1.607	.204
Vowel * Perceptual Training Intensity	2	150	293.382	.829	.439
Vowel * Perception Training * Production Training	2	150	968.692	2.736	.068
Vowel * Perception Training * Perceptual Training Intensity	2	150	302.667	.855	.427
Vowel * Production Training * Perceptual Training Intensity	2	150	27.919	.079	.924

Vowel * Perception Training * Production Training * Perceptual Training Intensity	2	150	122.563	.346	.708
Test * Vowel	4	300	343.295	5.398	.000
Test * Vowel * Perception Training	4	300	16.634	.262	.902
Test * Vowel * Production Training	4	300	45.239	.711	.585
Test * Vowel * Perceptual Training Intensity	4	300	121.100	1.904	.110
Test * Vowel * Perception Training * Production Training	4	300	30.582	.481	.750
Test * Vowel * Perception Training * Perceptual Training Intensity	4	300	108.278	1.703	.149
Test * Vowel * Production Training * Perceptual Training Intensity	4	300	48.724	.766	.548
Test * Vowel * Perception Training * Production Training * Perceptual Training Intensity	4	300	73.811	1.161	.328
Perception Training	1	75	7024.912	20.454	.000
Production Training	1	75	31.011	.090	.765
Perceptual Training Intensity	1	75	520.494	1.516	.222
Perception Training * Production Training	1	75	18.830	.055	.816
Perception Training * Perceptual Training Intensity	1	75	3526.344	10.268	.002
Production Training * Perceptual Training Intensity	1	75	3.438	.010	.921
Perception Training * Production Training * Perceptual Training Intensity	1	75	335.039	.976	.326

APPENDIX K
SUBJECTS' PERCENTAGES IN PRODUCTION PRE/POSTTEST & TC

	Username	Gender	Pre (/ɪ/)	Pre (/i:/)	Pre (avg)	Post (/ɪ/)	Post (/i:/)	Post (avg)	TC (/ɪ/)	TC (/i:/)	TC (avg)
HSP											
1	a01	F	0.00%	90.00%	45.00%	90.00%	90.00%	90.00%	100.00%	100.00%	100.00%
2	a02	F	20.00%	90.00%	55.00%	100.00%	90.00%	95.00%	93.33%	100.00%	96.67%
3	a11	F	0.00%	90.00%	45.00%	70.00%	100.00%	85.00%	73.33%	80.00%	76.67%
4	a37	F	10.00%	90.00%	50.00%	50.00%	100.00%	75.00%	46.67%	100.00%	73.33%
5	b02	F	10.00%	80.00%	45.00%	100.00%	90.00%	95.00%	86.67%	100.00%	93.33%
6	b04	F	10.00%	80.00%	45.00%	100.00%	90.00%	95.00%	100.00%	100.00%	100.00%
7	b38	M	20.00%	80.00%	50.00%	100.00%	90.00%	95.00%	93.33%	100.00%	96.67%
8	a25	M	0.00%	90.00%	45.00%	50.00%	90.00%	70.00%	13.33%	86.67%	50.00%
9	c38	M	10.00%	80.00%	45.00%	80.00%	70.00%	75.00%	26.67%	80.00%	53.33%
		Mean:	**8.89%**	**85.56%**	**47.22%**	**82.22%**	**90.00%**	**86.11%**	**70.37%**	**94.07%**	**82.22%**
		S.D.:	***7.82%***	***5.27%***	***3.63%***	***21.08%***	***8.66%***	***10.24%***	***33.18%***	***9.09%***	***19.86%***
HIP											
1	a17	F	10.00%	90.00%	50.00%	90.00%	90.00%	90.00%	93.33%	93.33%	93.33%
2	c02	F	10.00%	60.00%	35.00%	80.00%	60.00%	70.00%	46.67%	93.33%	70.00%
3	c03	F	20.00%	80.00%	50.00%	90.00%	100.00%	95.00%	46.67%	100.00%	73.33%
4	c04	F	0.00%	90.00%	45.00%	90.00%	90.00%	90.00%	86.67%	93.33%	90.00%
5	c06	F	30.00%	80.00%	55.00%	100.00%	90.00%	95.00%	53.33%	93.33%	73.33%
6	c35	M	0.00%	100.00%	50.00%	100.00%	100.00%	100.00%	53.33%	100.00%	76.67%
7	c36	M	0.00%	90.00%	45.00%	100.00%	90.00%	95.00%	93.33%	93.33%	93.33%
8	c37	M	10.00%	100.00%	55.00%	70.00%	100.00%	85.00%	80.00%	100.00%	90.00%
9	c27	M	20.00%	100.00%	60.00%	100.00%	100.00%	100.00%	66.67%	100.00%	83.33%
		Mean:	**11.11%**	**87.78%**	**49.44%**	**91.11%**	**91.11%**	**91.11%**	**68.89%**	**96.30%**	**82.59%**
		S.D.:	***10.54%***	***13.02%***	***7.26%***	***10.54%***	***12.69%***	***9.28%***	***19.72%***	***3.51%***	***9.40%***
HSN											
1	b01	F	10.00%	100.00%	55.00%	10.00%	100.00%	55.00%	80.00%	100.00%	90.00%
2	b03	F	30.00%	100.00%	65.00%	60.00%	100.00%	80.00%	93.33%	60.00%	76.67%
3	b05	F	50.00%	80.00%	65.00%	80.00%	90.00%	85.00%	46.67%	93.33%	70.00%
4	b06	F	0.00%	100.00%	50.00%	10.00%	100.00%	55.00%	60.00%	100.00%	80.00%
5	b07	F	10.00%	60.00%	35.00%	90.00%	100.00%	95.00%	46.67%	53.33%	50.00%
6	b08	F	30.00%	60.00%	45.00%	90.00%	100.00%	95.00%	73.33%	53.33%	63.33%
7	b25	M	10.00%	100.00%	55.00%	30.00%	100.00%	65.00%	33.33%	100.00%	66.67%
8	b21	M	10.00%	60.00%	35.00%	50.00%	100.00%	75.00%	93.33%	100.00%	96.67%
9	c39	M	10.00%	80.00%	45.00%	90.00%	100.00%	95.00%	46.67%	100.00%	73.33%
		Mean:	**17.78%**	**82.22%**	**50.00%**	**56.67%**	**98.89%**	**77.78%**	**63.70%**	**84.44%**	**74.07%**
		S.D.:	***15.63%***	***18.56%***	***11.18%***	***33.54%***	***3.33%***	***16.41%***	***22.14%***	***21.86%***	***14.02%***

HIN

1	a04	F	0.00%	100.00%	50.00%	10.00%	100.00%	55.00%	46.67%	100.00%	73.33%
2	a10	F	0.00%	100.00%	50.00%	60.00%	100.00%	80.00%	46.67%	100.00%	73.33%
3	b11	F	60.00%	50.00%	55.00%	90.00%	100.00%	95.00%	86.67%	100.00%	93.33%
4	c08	F	60.00%	50.00%	55.00%	60.00%	100.00%	80.00%	100.00%	86.67%	93.33%
5	c09	F	0.00%	100.00%	50.00%	100.00%	100.00%	100.00%	46.67%	100.00%	73.33%
6	c15	F	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	80.00%	100.00%	90.00%
7	b37	M	0.00%	100.00%	50.00%	10.00%	100.00%	55.00%	46.67%	53.33%	50.00%
8	b39	M	90.00%	10.00%	50.00%	100.00%	100.00%	100.00%	73.33%	100.00%	86.67%
		Mean:	**26.25%**	**76.25%**	**51.25%**	**58.75%**	**######**	**79.38%**	**65.83%**	**92.50%**	**79.17%**
		S.D.:	***37.39%***	***35.03%***	***2.31%***	***36.82%***	***0.00%***	***18.41%***	***21.80%***	***16.50%***	***14.77%***

LSP

1	b16	F	20.00%	50.00%	35.00%	80.00%	80.00%	80.00%	93.33%	80.00%	86.67%
2	b17	F	10.00%	70.00%	40.00%	90.00%	90.00%	90.00%	73.33%	73.33%	73.33%
3	a14	F	10.00%	80.00%	45.00%	90.00%	90.00%	90.00%	60.00%	53.33%	56.67%
4	b20	F	60.00%	90.00%	75.00%	100.00%	90.00%	95.00%	100.00%	86.67%	93.33%
5	b29	M	30.00%	60.00%	45.00%	30.00%	70.00%	50.00%	33.33%	80.00%	56.67%
6	b35	M	0.00%	100.00%	50.00%	30.00%	100.00%	65.00%	93.33%	80.00%	86.67%
7	a23	M	10.00%	90.00%	50.00%	10.00%	100.00%	55.00%	20.00%	60.00%	40.00%
		Mean:	**20.00%**	**77.14%**	**48.57%**	**61.43%**	**88.57%**	**75.00%**	**67.62%**	**73.33%**	**70.48%**
		S.D.:	***20.00%***	***17.99%***	***12.82%***	***36.71%***	***10.69%***	***18.26%***	***31.37%***	***12.17%***	***19.85%***

LIP

1	c12	F	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	33.33%	86.67%	60.00%
2	c18	F	10.00%	60.00%	35.00%	70.00%	100.00%	85.00%	40.00%	86.67%	63.33%
3	c21	F	10.00%	90.00%	50.00%	90.00%	100.00%	95.00%	80.00%	100.00%	90.00%
4	c23	F	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	20.00%	100.00%	60.00%
5	c24	F	0.00%	100.00%	50.00%	60.00%	100.00%	80.00%	20.00%	100.00%	60.00%
6	b27	M	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	33.33%	86.67%	60.00%
7	c32	M	20.00%	80.00%	50.00%	50.00%	100.00%	75.00%	20.00%	73.33%	46.67%
		Mean:	**5.71%**	**90.00%**	**47.86%**	**47.14%**	**######**	**73.57%**	**35.24%**	**90.48%**	**62.86%**
		S.D.:	***7.87%***	***15.28%***	***5.67%***	***28.12%***	***0.00%***	***14.06%***	***21.33%***	***10.08%***	***13.11%***

LSN

1	a12	F	80.00%	80.00%	80.00%	90.00%	100.00%	95.00%	80.00%	86.67%	83.33%
2	a15	F	30.00%	60.00%	45.00%	80.00%	90.00%	85.00%	26.67%	86.67%	56.67%
3	a19	F	0.00%	100.00%	50.00%	60.00%	60.00%	60.00%	66.67%	53.33%	60.00%
4	b19	F	10.00%	90.00%	50.00%	40.00%	80.00%	60.00%	46.67%	73.33%	60.00%
5	c22	F	30.00%	80.00%	55.00%	70.00%	70.00%	70.00%	53.33%	86.67%	70.00%
6	c26	M	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	6.67%	33.33%	20.00%
7	c31	M	10.00%	20.00%	15.00%	10.00%	60.00%	35.00%	33.33%	60.00%	46.67%
		Mean:	**22.86%**	**75.71%**	**49.29%**	**52.86%**	**80.00%**	**66.43%**	**44.76%**	**68.57%**	**56.67%**
		S.D.:	***28.12%***	***28.20%***	***19.02%***	***30.39%***	***17.32%***	***19.52%***	***24.86%***	***20.63%***	***19.81%***

LIN											
1	a09	F	10.00%	90.00%	50.00%	20.00%	90.00%	55.00%	40.00%	33.33%	36.67%
2	a22	F	20.00%	90.00%	55.00%	50.00%	80.00%	65.00%	46.67%	60.00%	53.33%
3	c13	F	0.00%	100.00%	50.00%	10.00%	100.00%	55.00%	26.67%	100.00%	63.33%
4	c17	F	0.00%	100.00%	50.00%	10.00%	100.00%	55.00%	46.67%	100.00%	73.33%
5	c19	F	30.00%	10.00%	20.00%	100.00%	90.00%	95.00%	80.00%	60.00%	70.00%
6	a03	F	0.00%	100.00%	50.00%	10.00%	100.00%	55.00%	26.67%	100.00%	63.33%
7	a21	M	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	73.33%	100.00%	86.67%
8	a24	M	10.00%	100.00%	55.00%	10.00%	100.00%	55.00%	26.67%	66.67%	46.67%
		Mean:	**8.75%**	**86.25%**	**47.50%**	**28.75%**	**95.00%**	**61.88%**	**45.83%**	**77.50%**	**61.67%**
		S.D.:	***11.26%***	***31.14%***	***11.34%***	***31.82%***	***7.56%***	***13.87%***	***20.91%***	***25.93%***	***15.84%***
COP											
1	c01	F	0.00%	100.00%	50.00%	60.00%	60.00%	60.00%	73.33%	73.33%	73.33%
2	c05	F	10.00%	80.00%	45.00%	40.00%	100.00%	70.00%	13.33%	60.00%	36.67%
3	c10	F	10.00%	80.00%	45.00%	20.00%	70.00%	45.00%	13.33%	60.00%	36.67%
4	c20	F	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	6.67%	100.00%	53.33%
5	c25	F	10.00%	80.00%	45.00%	20.00%	100.00%	60.00%	0.00%	100.00%	50.00%
6	b12	F	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	13.33%	100.00%	56.67%
7	b14	F	20.00%	70.00%	45.00%	100.00%	90.00%	95.00%	100.00%	80.00%	90.00%
8	c34	M	50.00%	60.00%	55.00%	100.00%	90.00%	95.00%	93.33%	60.00%	76.67%
9	a26	M	20.00%	90.00%	55.00%	20.00%	100.00%	60.00%	93.33%	100.00%	96.67%
10	a29	M	60.00%	60.00%	60.00%	100.00%	90.00%	95.00%	80.00%	86.67%	83.33%
11	a31	M	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	86.67%	100.00%	93.33%
		Mean:	**16.36%**	**83.64%**	**50.00%**	**45.45%**	**90.91%**	**68.18%**	**52.12%**	**83.64%**	**67.88%**
		S.D.:	***20.63%***	***15.67%***	***5.00%***	***39.08%***	***13.75%***	***18.88%***	***41.72%***	***17.73%***	***22.17%***
CON											
1	a16	F	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	75.00%	100.00%	87.50%
2	a38	F	10.00%	100.00%	55.00%	30.00%	70.00%	50.00%	60.00%	100.00%	80.00%
3	b18	F	40.00%	40.00%	40.00%	40.00%	60.00%	50.00%	55.00%	93.33%	74.17%
4	c07	F	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%
5	a05	F	0.00%	100.00%	50.00%	10.00%	100.00%	55.00%	53.33%	100.00%	76.67%
6	b10	F	20.00%	60.00%	40.00%	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%
7	a28	M	20.00%	80.00%	50.00%	0.00%	100.00%	50.00%	0.00%	66.67%	33.33%
8	b24	M	30.00%	60.00%	45.00%	30.00%	30.00%	30.00%	30.00%	30.00%	30.00%
9	c29	M	10.00%	70.00%	40.00%	10.00%	60.00%	35.00%	0.00%	6.67%	3.33%
10	c33	M	10.00%	90.00%	50.00%	10.00%	80.00%	45.00%	0.00%	13.33%	6.67%
		Mean:	**14.00%**	**80.00%**	**47.00%**	**13.00%**	**80.00%**	**46.50%**	**27.33%**	**71.00%**	**49.17%**
		S.D.:	***13.50%***	***21.60%***	***5.37%***	***14.94%***	***24.49%***	***7.84%***	***30.77%***	***39.25%***	***30.42%***

	Username	Gender	Pre (/e/)	Pre (/æ /)	Pre (avg)	Post (/e/)	Post (/æ /)	Post (avg)	TC (/e/)	TC (/æ /)	TC (avg)
HSP											
1	a01	F	0.00%	100.00%	50.00%	100.00%	100.00%	100.00%	100.00%	86.67%	93.33%
2	a02	F	100.00%	10.00%	55.00%	90.00%	100.00%	95.00%	100.00%	66.67%	83.33%
3	a11	F	20.00%	90.00%	55.00%	60.00%	100.00%	80.00%	93.33%	80.00%	86.67%
4	a37	F	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%
5	b02	F	100.00%	0.00%	50.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
6	b04	F	100.00%	0.00%	50.00%	90.00%	100.00%	95.00%	80.00%	100.00%	90.00%
7	b38	M	60.00%	40.00%	50.00%	100.00%	100.00%	100.00%	100.00%	60.00%	80.00%
8	a25	M	100.00%	10.00%	55.00%	100.00%	90.00%	95.00%	100.00%	46.67%	73.33%
9	c38	M	100.00%	0.00%	50.00%	100.00%	10.00%	55.00%	100.00%	13.33%	56.67%
		Mean:	**75.56%**	**27.78%**	**51.67%**	**93.33%**	**77.78%**	**85.56%**	**97.04%**	**61.48%**	**79.26%**
		S.D.:	***39.72%***	***40.24%***	***2.50%***	***13.23%***	***41.47%***	***19.76%***	***6.76%***	***35.87%***	***16.65%***
HIP											
1	a17	F	10.00%	0.00%	5.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2	c02	F	100.00%	0.00%	50.00%	100.00%	70.00%	85.00%	100.00%	66.67%	83.33%
3	c03	F	90.00%	10.00%	50.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4	c04	F	100.00%	40.00%	70.00%	100.00%	60.00%	80.00%	100.00%	66.67%	83.33%
5	c06	F	20.00%	50.00%	35.00%	50.00%	50.00%	50.00%	46.67%	33.33%	40.00%
6	c35	M	60.00%	60.00%	60.00%	100.00%	100.00%	100.00%	100.00%	66.67%	83.33%
7	c36	M	100.00%	0.00%	50.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
8	c37	M	100.00%	60.00%	80.00%	100.00%	80.00%	90.00%	100.00%	60.00%	80.00%
9	c27	M	30.00%	70.00%	50.00%	80.00%	100.00%	90.00%	66.67%	46.67%	56.67%
		Mean:	**67.78%**	**32.22%**	**50.00%**	**92.22%**	**84.44%**	**88.33%**	**90.37%**	**71.11%**	**80.74%**
		S.D.:	***38.33%***	***29.49%***	***21.36%***	***17.16%***	***20.07%***	***16.20%***	***19.75%***	***24.27%***	***20.53%***
HSN											
1	b01	F	100.00%	20.00%	60.00%	100.00%	40.00%	70.00%	100.00%	33.33%	66.67%
2	b03	F	100.00%	90.00%	95.00%	100.00%	100.00%	100.00%	100.00%	73.33%	86.67%
3	b05	F	40.00%	0.00%	20.00%	90.00%	80.00%	85.00%	73.33%	46.67%	60.00%
4	b06	F	100.00%	0.00%	50.00%	100.00%	60.00%	80.00%	93.33%	60.00%	76.67%
5	b07	F	60.00%	10.00%	35.00%	100.00%	0.00%	50.00%	40.00%	26.67%	33.33%
6	b08	F	100.00%	0.00%	50.00%	100.00%	40.00%	70.00%	100.00%	33.33%	66.67%
7	b25	M	70.00%	0.00%	35.00%	100.00%	60.00%	80.00%	100.00%	30.00%	65.00%
8	b21	M	100.00%	0.00%	50.00%	90.00%	80.00%	85.00%	100.00%	13.33%	56.67%
9	c39	M	100.00%	20.00%	60.00%	90.00%	30.00%	60.00%	100.00%	0.00%	50.00%
		Mean:	**85.56%**	**15.56%**	**50.56%**	**96.67%**	**54.44%**	**75.56%**	**89.63%**	**35.19%**	**62.41%**
		S.D.:	***22.97%***	***29.20%***	***21.13%***	***5.00%***	***30.46%***	***14.88%***	***20.58%***	***22.43%***	***15.30%***
HIN											
1	a04	F	100.00%	0.00%	50.00%	100.00%	40.00%	70.00%	100.00%	6.67%	53.33%
2	a10	F	100.00%	20.00%	60.00%	100.00%	80.00%	90.00%	100.00%	60.00%	80.00%

3	b11	F	10.00%	100.00%	55.00%	100.00%	100.00%	100.00%	100.00%	66.67%	83.33%
4	c08	F	20.00%	40.00%	30.00%	100.00%	100.00%	100.00%	100.00%	66.67%	83.33%
5	c09	F	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	100.00%	20.00%	60.00%
6	c15	F	60.00%	30.00%	45.00%	80.00%	40.00%	60.00%	73.33%	40.00%	56.67%
7	b37	M	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	100.00%	20.00%	60.00%
8	b39	M	20.00%	60.00%	40.00%	100.00%	100.00%	100.00%	100.00%	60.00%	80.00%
		Mean:	**63.75%**	**31.25%**	**47.50%**	**97.50%**	**57.50%**	**77.50%**	**96.67%**	**42.50%**	**69.58%**
		S.D.:	***41.38%***	***35.23%***	***9.26%***	***7.07%***	***43.34%***	***22.52%***	***9.43%***	***24.15%***	***13.15%***
LSP											
1	b16	F	10.00%	60.00%	35.00%	70.00%	100.00%	85.00%	93.33%	86.67%	90.00%
2	b17	F	50.00%	30.00%	40.00%	40.00%	40.00%	40.00%	80.00%	33.33%	56.67%
3	a14	F	100.00%	0.00%	50.00%	90.00%	60.00%	75.00%	86.67%	20.00%	53.33%
4	b20	F	60.00%	60.00%	60.00%	80.00%	80.00%	80.00%	100.00%	73.33%	86.67%
5	b29	M	50.00%	40.00%	45.00%	50.00%	100.00%	75.00%	93.33%	33.33%	63.33%
6	b35	M	100.00%	0.00%	50.00%	80.00%	60.00%	70.00%	93.33%	60.00%	76.67%
7	a23	M	100.00%	0.00%	50.00%	100.00%	40.00%	70.00%	60.00%	0.00%	30.00%
		Mean:	**67.14%**	**27.14%**	**47.14%**	**72.86%**	**68.57%**	**70.71%**	**86.67%**	**43.81%**	**65.24%**
		S.D.:	***34.50%***	***27.52%***	***8.09%***	***21.38%***	***25.45%***	***14.56%***	***13.33%***	***30.76%***	***21.07%***
LIP											
1	c12	F	100.00%	20.00%	60.00%	100.00%	10.00%	55.00%	60.00%	26.67%	43.33%
2	c18	F	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	93.33%	73.33%	83.33%
3	c21	F	60.00%	80.00%	70.00%	70.00%	100.00%	85.00%	93.33%	40.00%	66.67%
4	c23	F	100.00%	0.00%	50.00%	70.00%	20.00%	45.00%	40.00%	6.67%	23.33%
5	c24	F	100.00%	0.00%	50.00%	40.00%	100.00%	70.00%	60.00%	46.67%	53.33%
6	b27	M	20.00%	10.00%	15.00%	90.00%	100.00%	95.00%	80.00%	20.00%	50.00%
7	c32	M	100.00%	0.00%	50.00%	70.00%	50.00%	60.00%	86.67%	6.67%	46.67%
		Mean:	**82.86%**	**15.71%**	**49.29%**	**77.14%**	**54.29%**	**65.71%**	**73.33%**	**31.43%**	**52.38%**
		S.D.:	***31.47%***	***29.36%***	***16.94%***	***21.38%***	***45.41%***	***18.58%***	***20.37%***	***23.95%***	***18.83%***
LSN											
1	a12	F	80.00%	100.00%	90.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
2	a15	F	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%
3	a19	F	100.00%	0.00%	50.00%	70.00%	80.00%	75.00%	73.33%	20.00%	46.67%
4	b19	F	100.00%	0.00%	50.00%	100.00%	30.00%	65.00%	100.00%	13.33%	56.67%
5	c22	F	100.00%	0.00%	50.00%	100.00%	30.00%	65.00%	100.00%	33.33%	66.67%
6	c26	M	60.00%	0.00%	30.00%	60.00%	0.00%	30.00%	33.33%	0.00%	16.67%
7	c31	M	100.00%	10.00%	55.00%	60.00%	100.00%	80.00%	33.33%	40.00%	36.67%
		Mean:	**91.43%**	**15.71%**	**53.57%**	**84.29%**	**48.57%**	**66.43%**	**77.14%**	**29.52%**	**53.33%**
		S.D.:	***15.74%***	***37.35%***	***17.96%***	***19.88%***	***44.13%***	***22.31%***	***31.47%***	***34.61%***	***25.96%***
LIN											
1	a09	F	100.00%	10.00%	55.00%	100.00%	10.00%	55.00%	100.00%	46.67%	73.33%
2	a22	F	100.00%	0.00%	50.00%	90.00%	100.00%	95.00%	100.00%	33.33%	66.67%
3	c13	F	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	100.00%	26.67%	63.33%

4	c17	F	100.00%	0.00%	50.00%	70.00%	70.00%	70.00%	33.33%	60.00%	46.67%
5	c19	F	100.00%	20.00%	60.00%	60.00%	80.00%	70.00%	66.67%	66.67%	66.67%
6	a03	F	0.00%	70.00%	35.00%	10.00%	70.00%	40.00%	40.00%	0.00%	20.00%
7	a21	M	100.00%	0.00%	50.00%	60.00%	60.00%	60.00%	60.00%	26.67%	43.33%
8	a24	M	100.00%	0.00%	50.00%	70.00%	80.00%	75.00%	66.67%	20.00%	43.33%
		Mean:	**87.50%**	**12.50%**	**50.00%**	**70.00%**	**58.75%**	**64.38%**	**70.83%**	**35.00%**	**52.92%**
		S.D.:	***35.36%***	***24.35%***	***7.07%***	***29.28%***	***35.23%***	***17.00%***	***26.89%***	***21.89%***	***17.77%***
COP											
1	c01	F	100.00%	0.00%	50.00%	40.00%	100.00%	70.00%	80.00%	73.33%	76.67%
2	c05	F	90.00%	0.00%	45.00%	100.00%	0.00%	50.00%	100.00%	26.67%	63.33%
3	c10	F	60.00%	10.00%	35.00%	80.00%	10.00%	45.00%	93.33%	13.33%	53.33%
4	c20	F	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%
5	c25	F	100.00%	0.00%	50.00%	60.00%	60.00%	60.00%	33.33%	6.67%	20.00%
6	b12	F	100.00%	0.00%	50.00%	100.00%	20.00%	60.00%	100.00%	0.00%	50.00%
7	b14	F	20.00%	70.00%	45.00%	100.00%	100.00%	100.00%	86.67%	86.67%	86.67%
8	c34	M	100.00%	40.00%	70.00%	100.00%	100.00%	100.00%	93.33%	40.00%	66.67%
9	a26	M	30.00%	10.00%	20.00%	10.00%	90.00%	50.00%	100.00%	33.33%	66.67%
10	a29	M	60.00%	70.00%	65.00%	90.00%	100.00%	95.00%	86.67%	80.00%	83.33%
11	a31	M	100.00%	0.00%	50.00%	100.00%	30.00%	65.00%	100.00%	40.00%	70.00%
		Mean:	**78.18%**	**18.18%**	**48.18%**	**80.00%**	**55.45%**	**67.73%**	**88.48%**	**36.36%**	**62.42%**
		S.D.:	***30.60%***	***28.22%***	***13.28%***	***30.66%***	***43.90%***	***21.02%***	***19.57%***	***31.60%***	***18.74%***
CON											
1	a16	F	100.00%	0.00%	50.00%	100.00%	10.00%	55.00%	100.00%	0.00%	50.00%
2	a38	F	100.00%	10.00%	55.00%	100.00%	10.00%	55.00%	100.00%	13.33%	56.67%
3	b18	F	60.00%	30.00%	45.00%	40.00%	40.00%	40.00%	40.00%	6.67%	23.33%
4	c07	F	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%
5	a05	F	60.00%	20.00%	40.00%	10.00%	50.00%	30.00%	66.67%	13.33%	40.00%
6	b10	F	60.00%	20.00%	40.00%	40.00%	0.00%	20.00%	20.00%	20.00%	20.00%
7	a28	M	100.00%	0.00%	50.00%	100.00%	20.00%	60.00%	100.00%	0.00%	50.00%
8	b24	M	100.00%	10.00%	55.00%	100.00%	0.00%	50.00%	100.00%	26.67%	63.33%
9	c29	M	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%	100.00%	0.00%	50.00%
10	c33	M	100.00%	0.00%	50.00%	100.00%	10.00%	55.00%	100.00%	6.67%	53.33%
		Mean:	**88.00%**	**9.00%**	**48.50%**	**79.00%**	**14.00%**	**46.50%**	**82.67%**	**8.67%**	**45.67%**
		S.D.:	***19.32%***	***11.01%***	***5.30%***	***34.79%***	***17.76%***	***12.70%***	***30.01%***	***9.45%***	***13.97%***

	Username	Gender	Pre (/ʊ/)	Pre (/uː/)	Pre (avg)	Post (/ʊ/)	Post (/uː/)	Post (avg)	TC (/ʊ/)	TC (/uː/)	TC (avg)
HSP											
1	a01	F	0.00%	100.00%	50.00%	60.00%	100.00%	80.00%	90.00%	80.00%	85.00%
2	a02	F	0.00%	100.00%	50.00%	60.00%	100.00%	80.00%	60.00%	80.00%	70.00%
3	a11	F	0.00%	100.00%	50.00%	60.00%	100.00%	80.00%	90.00%	90.00%	90.00%
4	a37	F	20.00%	40.00%	30.00%	60.00%	60.00%	60.00%	10.00%	60.00%	35.00%
5	b02	F	0.00%	40.00%	20.00%	60.00%	100.00%	80.00%	90.00%	80.00%	85.00%
6	b04	F	20.00%	60.00%	40.00%	40.00%	100.00%	70.00%	90.00%	80.00%	85.00%
7	b38	M	20.00%	60.00%	40.00%	100.00%	100.00%	100.00%	70.00%	100.00%	85.00%
8	a25	M	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	100.00%	80.00%	90.00%
9	c38	M	20.00%	80.00%	50.00%	20.00%	100.00%	60.00%	70.00%	90.00%	80.00%
		Mean:	**8.89%**	**75.56%**	**42.22%**	**55.56%**	**95.56%**	**75.56%**	**74.44%**	**82.22%**	**78.33%**
		S.D.:	***10.54%***	***26.03%***	***10.93%***	***21.86%***	***13.33%***	***12.36%***	***27.44%***	***10.93%***	***17.32%***
HIP											
1	a17	F	40.00%	60.00%	50.00%	60.00%	100.00%	80.00%	100.00%	80.00%	90.00%
2	c02	F	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	90.00%	80.00%	85.00%
3	c03	F	20.00%	40.00%	30.00%	60.00%	100.00%	80.00%	60.00%	90.00%	75.00%
4	c04	F	0.00%	80.00%	40.00%	20.00%	100.00%	60.00%	80.00%	100.00%	90.00%
5	c06	F	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	50.00%	80.00%	65.00%
6	c35	M	20.00%	40.00%	30.00%	40.00%	100.00%	70.00%	90.00%	90.00%	90.00%
7	c36	M	0.00%	100.00%	50.00%	80.00%	100.00%	90.00%	90.00%	90.00%	90.00%
8	c37	M	20.00%	60.00%	40.00%	40.00%	100.00%	70.00%	80.00%	90.00%	85.00%
9	c27	M	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	90.00%	100.00%	95.00%
		Mean:	**11.11%**	**75.56%**	**43.33%**	**44.44%**	**######**	**72.22%**	**81.11%**	**88.89%**	**85.00%**
		S.D.:	***14.53%***	***26.03%***	***8.66%***	***19.44%***	***0.00%***	***9.72%***	***16.16%***	***7.82%***	***9.35%***
HSN											
1	b01	F	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	60.00%	80.00%	70.00%
2	b03	F	0.00%	100.00%	50.00%	20.00%	80.00%	50.00%	40.00%	90.00%	65.00%
3	b05	F	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	60.00%	80.00%	70.00%
4	b06	F	0.00%	100.00%	50.00%	40.00%	60.00%	50.00%	40.00%	40.00%	40.00%
5	b07	F	20.00%	60.00%	40.00%	100.00%	60.00%	80.00%	70.00%	60.00%	65.00%
6	b08	F	0.00%	100.00%	50.00%	0.00%	80.00%	40.00%	50.00%	80.00%	65.00%
7	b25	M	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	30.00%	80.00%	55.00%
8	b21	M	60.00%	0.00%	30.00%	60.00%	80.00%	70.00%	80.00%	80.00%	80.00%
9	c39	M	20.00%	60.00%	40.00%	60.00%	100.00%	80.00%	50.00%	90.00%	70.00%
		Mean:	**11.11%**	**80.00%**	**45.56%**	**37.78%**	**84.44%**	**61.11%**	**53.33%**	**75.56%**	**64.44%**
		S.D.:	***20.28%***	***34.64%***	***7.26%***	***30.73%***	***16.67%***	***13.64%***	***15.81%***	***15.90%***	***11.30%***
HIN											
1	a04	F	20.00%	60.00%	40.00%	40.00%	80.00%	60.00%	60.00%	50.00%	55.00%
2	a10	F	20.00%	60.00%	40.00%	60.00%	80.00%	70.00%	80.00%	80.00%	80.00%

3	b11	F	0.00%	80.00%	40.00%	20.00%	80.00%	50.00%	60.00%	80.00%	70.00%
4	c08	F	20.00%	60.00%	40.00%	40.00%	100.00%	70.00%	100.00%	90.00%	95.00%
5	c09	F	0.00%	100.00%	50.00%	20.00%	60.00%	40.00%	60.00%	50.00%	55.00%
6	c15	F	20.00%	60.00%	40.00%	60.00%	100.00%	80.00%	80.00%	90.00%	85.00%
7	b37	M	0.00%	100.00%	50.00%	60.00%	60.00%	60.00%	10.00%	80.00%	45.00%
8	b39	M	20.00%	80.00%	50.00%	20.00%	100.00%	60.00%	70.00%	80.00%	75.00%
		Mean:	**12.50%**	**75.00%**	**43.75%**	**40.00%**	**82.50%**	**61.25%**	**65.00%**	**75.00%**	**70.00%**
		S.D.:	***10.35%***	***17.73%***	***5.18%***	***18.52%***	***16.69%***	***12.46%***	***26.19%***	***16.04%***	***17.11%***
LSP											
1	b16	F	20.00%	80.00%	50.00%	40.00%	100.00%	70.00%	80.00%	70.00%	75.00%
2	b17	F	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	80.00%	90.00%	85.00%
3	a14	F	0.00%	80.00%	40.00%	20.00%	80.00%	50.00%	60.00%	80.00%	70.00%
4	b20	F	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	90.00%	90.00%	90.00%
5	b29	M	0.00%	80.00%	40.00%	40.00%	100.00%	70.00%	60.00%	90.00%	75.00%
6	b35	M	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	80.00%	90.00%	85.00%
7	a23	M	0.00%	80.00%	40.00%	0.00%	80.00%	40.00%	50.00%	80.00%	65.00%
		Mean:	**2.86%**	**88.57%**	**45.71%**	**28.57%**	**94.29%**	**61.43%**	**71.43%**	**84.29%**	**77.86%**
		S.D.:	***7.56%***	***10.69%***	***5.35%***	***15.74%***	***9.76%***	***12.15%***	***14.64%***	***7.87%***	***9.06%***
LIP											
1	c12	F	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	30.00%	80.00%	55.00%
2	c18	F	20.00%	60.00%	40.00%	20.00%	100.00%	60.00%	40.00%	90.00%	65.00%
3	c21	F	0.00%	80.00%	40.00%	40.00%	80.00%	60.00%	70.00%	100.00%	85.00%
4	c23	F	0.00%	100.00%	50.00%	40.00%	60.00%	50.00%	10.00%	80.00%	45.00%
5	c24	F	0.00%	100.00%	50.00%	20.00%	80.00%	50.00%	20.00%	70.00%	45.00%
6	b27	M	20.00%	40.00%	30.00%	20.00%	80.00%	50.00%	40.00%	90.00%	65.00%
7	c32	M	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	60.00%	80.00%	70.00%
		Mean:	**5.71%**	**82.86%**	**44.29%**	**31.43%**	**85.71%**	**58.57%**	**38.57%**	**84.29%**	**61.43%**
		S.D.:	***9.76%***	***24.30%***	***7.87%***	***10.69%***	***15.12%***	***9.00%***	***21.16%***	***9.76%***	***14.35%***
LSN											
1	a12	F	0.00%	100.00%	50.00%	60.00%	100.00%	80.00%	80.00%	100.00%	90.00%
2	a15	F	0.00%	80.00%	40.00%	20.00%	100.00%	60.00%	20.00%	70.00%	45.00%
3	a19	F	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	40.00%	40.00%	40.00%
4	b19	F	0.00%	100.00%	50.00%	20.00%	80.00%	50.00%	20.00%	60.00%	40.00%
5	c22	F	0.00%	100.00%	50.00%	40.00%	80.00%	60.00%	40.00%	80.00%	60.00%
6	c26	M	0.00%	80.00%	40.00%	20.00%	100.00%	60.00%	20.00%	100.00%	60.00%
7	c31	M	0.00%	80.00%	40.00%	0.00%	60.00%	30.00%	60.00%	60.00%	60.00%
		Mean:	**0.00%**	**91.43%**	**45.71%**	**28.57%**	**88.57%**	**58.57%**	**40.00%**	**72.86%**	**56.43%**
		S.D.:	***0.00%***	***10.69%***	***5.35%***	***19.52%***	***15.74%***	***15.74%***	***23.09%***	***22.15%***	***17.49%***
LIN											
1	a09	F	0.00%	100.00%	50.00%	0.00%	80.00%	40.00%	0.00%	60.00%	30.00%
2	a22	F	40.00%	40.00%	40.00%	60.00%	80.00%	70.00%	60.00%	60.00%	60.00%
3	c13	F	0.00%	100.00%	50.00%	40.00%	80.00%	60.00%	20.00%	40.00%	30.00%

4	c17	F	20.00%	60.00%	40.00%	60.00%	100.00%	80.00%	40.00%	60.00%	50.00%
5	c19	F	20.00%	40.00%	30.00%	40.00%	100.00%	70.00%	70.00%	80.00%	75.00%
6	a03	F	0.00%	100.00%	50.00%	20.00%	80.00%	50.00%	20.00%	70.00%	45.00%
7	a21	M	20.00%	80.00%	50.00%	0.00%	100.00%	50.00%	0.00%	40.00%	20.00%
8	a24	M	0.00%	100.00%	50.00%	20.00%	100.00%	60.00%	60.00%	80.00%	70.00%
		Mean:	**12.50%**	**77.50%**	**45.00%**	**30.00%**	**90.00%**	**60.00%**	**33.75%**	**61.25%**	**47.50%**
		S.D.:	***14.88%***	***27.12%***	***7.56%***	***23.90%***	***10.69%***	***13.09%***	***27.74%***	***15.53%***	***20.00%***
COP											
1	c01	F	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	60.00%	80.00%	70.00%
2	c05	F	0.00%	100.00%	50.00%	20.00%	80.00%	50.00%	10.00%	100.00%	55.00%
3	c10	F	0.00%	80.00%	40.00%	40.00%	100.00%	70.00%	20.00%	80.00%	50.00%
4	c20	F	0.00%	80.00%	40.00%	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%
5	c25	F	0.00%	100.00%	50.00%	40.00%	80.00%	60.00%	10.00%	80.00%	45.00%
6	b12	F	0.00%	100.00%	50.00%	40.00%	100.00%	70.00%	10.00%	100.00%	55.00%
7	b14	F	40.00%	60.00%	50.00%	60.00%	80.00%	70.00%	40.00%	100.00%	70.00%
8	c34	M	40.00%	60.00%	50.00%	40.00%	100.00%	70.00%	60.00%	100.00%	80.00%
9	a26	M	20.00%	40.00%	30.00%	60.00%	60.00%	60.00%	40.00%	70.00%	55.00%
10	a29	M	0.00%	60.00%	30.00%	60.00%	60.00%	60.00%	60.00%	60.00%	60.00%
11	a31	M	0.00%	100.00%	50.00%	40.00%	40.00%	40.00%	60.00%	80.00%	70.00%
		Mean:	**9.09%**	**80.00%**	**44.55%**	**36.36%**	**81.82%**	**59.09%**	**33.64%**	**86.36%**	**60.00%**
		S.D.:	***16.40%***	***21.91%***	***8.20%***	***21.57%***	***20.89%***	***10.44%***	***24.20%***	***14.33%***	***10.95%***
CON											
1	a16	F	0.00%	100.00%	50.00%	40.00%	40.00%	40.00%	20.00%	100.00%	60.00%
2	a38	F	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	0.00%	60.00%	30.00%
3	b18	F	0.00%	80.00%	40.00%	20.00%	60.00%	40.00%	20.00%	50.00%	35.00%
4	c07	F	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%
5	a05	F	0.00%	80.00%	40.00%	20.00%	100.00%	60.00%	20.00%	100.00%	60.00%
6	b10	F	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	0.00%	70.00%	35.00%
7	a28	M	0.00%	100.00%	50.00%	0.00%	80.00%	40.00%	0.00%	100.00%	50.00%
8	b24	M	0.00%	80.00%	40.00%	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%
9	c29	M	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%	0.00%	100.00%	50.00%
10	c33	M	0.00%	80.00%	40.00%	0.00%	80.00%	40.00%	0.00%	100.00%	50.00%
		Mean:	**0.00%**	**92.00%**	**46.00%**	**8.00%**	**86.00%**	**47.00%**	**6.00%**	**88.00%**	**47.00%**
		S.D.:	***0.00%***	***10.33%***	***5.16%***	***13.98%***	***21.19%***	***6.75%***	***9.66%***	***19.89%***	***10.33%***

APPENDIX L

RESULTS OF THE PRODUCTION DATA SUBMITTED TO 5-WAY ANOVA

(PRETEST VS. POSTTEST)

<u>Between-subjects factor:</u>

Perception Training (HVPT, LVTP, none)

Production Training (with, without)

Perceptual Training Intensity (standard, intensive, none)

<u>Within-subjects factor (repeated-measures):</u>

Test (pretest, posttest)

Vowel (front, mid, back)

Main Effects / Interactions	**df (between-subject)**	**df (within-subject)**	**Mean Square**	**F**	**Sig.**
Test	1	75	45987.516	239.497	.000
Test * Perception Training	1	75	3716.184	19.353	.000
Test * Production Training	1	75	4160.041	21.665	.000
Test * Perceptual Training Intensity	1	75	.463	.002	.961
Test * Perception Training * Production Training	1	75	207.640	1.081	.302
Test * Perception Training * Perceptual Training Intensity	1	75	56.141	.292	.590
Test * Production Training * Perceptual Training Intensity	1	75	35.555	.185	.668
Test * Perception Training * Production Training * Perceptual Training Intensity	1	75	5.285	.028	.869
Vowel	2	150	2715.376	18.108	.000
Vowel * Perception Training	2	150	203.914	1.360	.260
Vowel * Production Training	2	150	107.571	.717	.490
Vowel * Perceptual Training Intensity	2	150	15.177	.101	.904
Vowel * Perception Training * Production Training	2	150	131.143	.875	.419

Vowel * Perception Training * Perceptual Training Intensity	2	150	44.048	.294	.746
Vowel * Production Training * Perceptual Training Intensity	2	150	26.650	.178	.837
Vowel * Perception Training * Production Training * Perceptual Training Intensity	2	150	3.592	.024	.976
Test * Vowel	2	150	412.873	3.444	.034
Test * Vowel * Perception Training	2	150	65.682	.548	.579
Test * Vowel * Production Training	2	150	41.199	.344	.710
Test * Vowel * Perceptual Training Intensity	2	150	4.192	.035	.966
Test * Vowel * Perception Training * Production Training	2	150	85.514	.713	.492
Test * Vowel * Perception Training * Perceptual Training Intensity	2	150	41.371	.345	.709
Test * Vowel * Production Training * Perceptual Training Intensity	2	150	32.686	.273	.762
Test * Vowel * Perception Training * Production Training * Perceptual Training Intensity	2	150	15.584	.130	.878
Perception Training	1	75	3576.156	12.305	.001
Production Training	1	75	3287.615	11.312	.001
Perceptual Training Intensity	1	75	33.941	.117	.734
Perception Training * Production Training	1	75	297.361	1.023	.315
Perception Training * Perceptual Training Intensity	1	75	117.560	.405	.527
Production Training * Perceptual Training Intensity	1	75	10.557	.036	.849
Perception Training * Production Training * Perceptual Training Intensity	1	75	2.738	.009	.923

Results of the Production Data submitted to 4-way ANOVA (TC)

Between-subjects factor:

Perception Training (HVPT, LVTP, none)

Production Training (with, without)

Perceptual Training Intensity (standard, intensive, none)

Within-subjects factor (repeated-measures):

Vowel (front, mid, back)

Main Effects / Interactions	df (between-subject)	df (within-subject)	Mean Square	F	Sig.
Vowel	2	150	784.416	4.017	.020
Vowel * Perception Training	2	150	52.597	.269	.764
Vowel * Production Training	2	150	135.950	.696	.500
Vowel * Perceptual Training Intensity	2	150	60.317	.309	.735
Vowel * Perception Training * Production Training	2	150	157.932	.809	.447
Vowel * Perception Training * Perceptual Training Intensity	2	150	220.583	1.130	.326
Vowel * Production Training * Perceptual Training Intensity	2	150	43.854	.225	.799
Vowel * Perception Training * Production Training * Perceptual Training Intensity	2	150	.866	.004	.996
Perception Training	1	75	11729.385	23.040	.000
Production Training	1	75	9587.183	18.832	.000
Perceptual Training Intensity	1	75	77.008	.151	.698
Perception Training * Production Training	1	75	13.390	.026	.872
Perception Training * Perceptual Training Intensity	1	75	1488.565	2.924	.091
Production Training * Perceptual Training Intensity	1	75	587.499	1.154	.286
Perception Training * Production Training * Perceptual Training Intensity	1	75	172.523	.339	.562

REFERENCES

Abramson, A. S., & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. In B. Hála, M. Romportl & P. Janota (Eds.), *Proceedings of the 6th International Congress of Phonetic Sciences* (pp. 569-573). Prague: Academia.

Akahane-Yamada, R., Bradlow, A. R., Pisoni, D., & Tohkura, Y. (1997). Effects of audio-visual training on the identification of English /r/ and /l/ by Japanese speakers. *Journal of the Acoustical Society of America*, *102*, 3137.

Alexander, J. A., Wong, P. C. M., & Bradlow, A. R. (2005). Lexical tone perception in musicians and nonmusicians. In *9th European Conference on Speech Communication and Technology,* Lisboa, Portugal.

Aliaga-García, C. (2010). Measuring perceptual cue weighting after training: A comparison of auditory vs. articulatory training methods. In K. Dziubalska-Kołaczyk, M. Wrembel, & M. Kul (Eds.), *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech, New Sounds 2010* (pp.12-18). Poland: Poznań.

Aliaga-García, C., & Mora, J. C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. In A. M. A. Watkins, A. S. Rauber and B. O. Baptista (Eds.), *Recent Research in Second Language Phonetics/Phonology: Perception and Production* (pp. 2-31). Newcastle upon Tyne: Cambridge Scholars Publishing.

Alivuotila, L., Hakokari, J., Savela, J., Happonen, R. P., & Aaltonen, O. (2007).

Perception and imitation of Finnish open vowels among children, naïve adults, and trained phoneticians. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 361-364), Saarbrucken, Germany.

Allen, M. M. (2013). Intervention Efficacy and Intensity for Children with Speech Sound Disorders. *Journal of Speech, Language, and Hearing Research, 56(3),* 865-877.

Altenberg, E.P., & Vago, R.M. (1983). Theoretical implications of an error analysis of second language phonology production. *Language Learning, 33*, 427-447. Reprinted in G. Ioup, & S.H. Weinberger (Eds.), *Interlanguage Phonology*, (pp.148-164), Cambridge: Newbury House Publishers, 1987.

Altmann, H., Berger, I., & Braun, B. (2012). Asymmetries in the perception of non-native consonantal and vocalic length contrasts. *Second Language Research*, *28(4),* 387-413.

Alves, U. K., & Magro, V. (2011). Raising awareness of L2 phonology: explicit instruction and the acquisition of aspirated /p/ by Brazilian Portuguese speakers. *Letras de Hoje*, *46(3),* 71-80.

Anderson, J. (1987). The markedness differential hypothesis and syllable structure difficulty. In G. Ioup, & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 279-291). Cambridge, MA: Newbury House.

Archibald, J. (1998). Second Language Phonology, Phonetics, and Typology. *Studies in*

Second Language Acquisition, 20(1), 189-211.

Arteaga, D. L. (2000). Articulatory Phonetics in the First Year Spanish Classroom. *Modern Language Journal 84(3),* 339-54.

Asher, J. J. & García, R. (1969). The optimal age to learn a second language. *The Modern Language Journal, 53,* 334-341.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effect of early experience. *Child Development, 52,* 1135-1145.

Baese-Berk, M. (2010). *An examination of the relationship between speech perception and production* (Unpublished doctoral dissertation). Northwestern University, Illinois.

Baker, E. (2012). Optimal intervention intensity. *International Journal of Speech-Language Pathology, 14,* 401-409.

Baker, W., & Trofimovich, P. (2000). Korean-English bilinguals' perception and production of American-English vowels: Effects of age, experience, and native language phonology. *Paper presented at the American Association of Applied Linguistics*, Vancouver.

Bambra, L.M., Warren, S. E. (1993). Masses trials revisited: Appropriate applications in functional skill training. In Gable, R. A., Warren, S. E. (Eds), *Strategies for teaching students with mild to severe mental retardation* (pp. 165-190). London: Kingsley.

Barrie, M. (2003). Contrast in Cantonese Vowels. *Toronto Working Papers in*

Linguistics, 20, 1-19.

Beddor, P. S., Harnsberger, J. D., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: *Acoustic structures and their perceptual correlates. Journal of Phonetics, 30,* 591-627.

Beddor, P. S., & Gottfried, T. L. (1995). Methodological issues in cross-language speech perception research with adults. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 207-232). Timonium, MD: York Press.

Benson, B. (1988). Universal preference for the open syllable as an independent process in interlanguage phonology. *Language Learning, 38,* 221-242.

Bernard, J. R., & Mannell, R. H. (1986). A study /h_d/ words in Australian English. *Working Papers of the Speech, Hearing and Language Research Centre,* Macquarie University.

Best, C. T. (1984). Discovering messages in the medium: Speech perception and the prelinguistic infant. In H. E. Fitzgerald, B. Lester and M. Yogman (Eds.), *Theory and research in behavioral pediatrics* (pp. 139-180). New York: Plenum.

Best, C. T. (1993). Emergence of language-specific constraints in perception o non-native speech contrasts: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life,* (pp. 289-304). Dordrecht, the Netherlands: Kluwer Academic Publisher.

Best, C. T. (1994a). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman and H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Best, C. T. (1994b). Learning to perceive the sound pattern of English. In C. Rovee-Collier and L. Lipsitt (Eds.), *Advances in Infancy Research* (pp. 217-304). Norwood, NJ: Ablex.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistics experience: Issues in cross-language research* (pp. 171-206). Timonium, MD: York Press.

Best, C. T. (1999). Development of language-specific influences on speech perception and production in pre-verbal infancy. In J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences (pp. 1261-1264).* San Francisco/Berkeley, CA: Department of Linguistics.

Best, C. T., Hallé, P. A., Bohn, O.-S., & Faber, A. (2003). Cross-language perception of nonnative vowels: Phonological and phonetic effects of listeners' native languages. In M. J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2889-2892). Barcelona: Causal Productions.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109(2),* 775-794.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 345-360.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on crosslanguage perception of approximants. *Journal of Phonetics, 20,* 305-350.

Best, C. T., Traill, A., Carter, A., Harrison, K. D., & Faber, A. (2003). !Xóõ click perception by English, Isizulu, and Sesotho listeners. In M. J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 853-856). Barcelona: Causal Productions.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. Munro and O.-S. Bohn (Ed.), *Second language speech learning* (pp.13-34). Amsterdam: John Benjamins Publishing.

Bhatia, T. K. (1995). Acquisition of Voicing and Aspiration in Second Language Development. In V. Gambhir (Ed.), *Teaching and Acquisition of South Asian Languages* (pp. 182-196). Philadelphia: University of Pennsylvania Press.

Bialystok, E., & Hakuta, K. (1994). *In other words: The science and psychology of second language acquisition*. New York: Basic Books.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language, 68,* 706-755.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language, 44*, 235-249.

Blake, B. J. (1985). Short a' in Melbourne English. *Journal of the International Phonetic Association, 15,* 6-20.

Boersma, P. (1998). *Functional phonology* (Doctoral dissertation). University of Amsterdam, the Netherlands. The Hague: Holland Academic Graphics.

Boersma, P., Escudero P., & Hayes, R. (2003). Learning abstract phonological from auditory phonetic categories: an integrated model for the acquisition of language-specific sound categories. In M. J. Sole, D. Recansens & J. Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences, 1013-1016.* Barcelona: Causal Productions.

Boersma, P., & Weenink, D. (2005). *PRAAT: Doing phonetics by computer.* Retrieved from http://www.fon.hum.uva.nl/praat/.

Bohn, O.-S. (1995). Cross-language speech perception in adults: L1 transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp.275-300). Timonium, MD: York Press.

Bohn, O.-S. (1998). Linguistic relativity in speech perception: An overview of the influence of language experience on the perception of speech sounds from infancy to adulthood. In S. Niemeier and R. Dirven (Eds.), *Evidence for Linguistic Relativity*. Amsterdam and Philadelpha: John Benjamins.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11,* 303-328.

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14,* 131-158.

Bohn, O.-S., & Steinlen, A. K. (2003). Consonantal context affects cross-language perception of vowels. In M. J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2289-2292). Barcelona: Causal Productions.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis*. Mawah, NJ: Lawrence Erlbaum Associates.

Bongaerts, T., Planken, B. & Schils, E. (1995). Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton and Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 30-50). Multilingual Matters, Clevedon, UK.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19,* 447-465.

Bosch, L. & Sebastian-Galles, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech, 46,* 217-243.

Bowen, T., & Marks, J. (1992). *The pronunciation book: Student centred activities for pronunciation work.* London: Longman.

Bradlow, A. R. (2008). Training non-native language sound patterns: Lessons from training Japanese adults on the English /ɹ/ and /l/ contrast. In J. G. Hansen-Edwards & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition.* Philadelphia: John Benjamins Publishing Company.

Bradlow, A. R., & Bent. T. (2003). *Listener adaptation to foreign accented speech.* Paper presented at the International Congress of Phonetic Sciences, Barcelona, Spain.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics, 61,* 977-985.

Bradlow, A. R., Pisoni, D. B., Yamada, R. A., & Tohkura, Y. (1997). Training Japanese listeners to identify English /ɹ/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101,* 2299-2310.

Brown, A. (1991). *Pronunciation Models.* Singapore: University Press.

Brown, A. (1995). Minimal pairs: minimal importance? *ELT Journal*, *49*, 169-175.

Butcher, A. (2006). Formant frequencies of /hVd/ vowels in the speech of South Australian females. In P. Warren & C. Watson (Eds.), *Proceedings of the 11th Australian International Conference on Speech Science & Technology, 6-8 December 2006* (pp. 449-453). Auckland, New Zealand.

Carlisle, R. S. (1991). The influence of environment on vowel epenthesis in

Spanish/English interphonology. *Applied Linguistics, 12,* 76-95.

Carmell, T. (1997). *Spectral Cues Homepage.* Retrieved June 3, 2010, from http://speech.bme.ogi.edu/tutordemos/SpectrogramReading/ipa/ipahome.html

Carney, A.E., Widin, G. P., & Viemeister, N. F. (1977). Noncategorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America, 62,* 961-970.

Carruthers, S. W. (2007). Articulatory Training on Facial Movements Using the Webcam Pronunciation Mirror. *TESOL Working Paper Series, 5(1).* Retrieved on 10th, February, 2012 from http://www.hpu.edu/CHSS/LangLing/TESOL/ProfessionalDevelopment/200720TWPspring07/carruthers%205-1.pdf.

Castino, J. (1996). Impact of a Phonetics Course on FL Learners' Acquisition of Spanish Phonology. *Selecta: Journal of the Pacific Northwest Council on Foreign Languages, 17,* 55-58.

Catford, J. C., & Pisoni, D. B. (1970). Auditory vs. articulatory training in exotic sounds. *Modern Language Journal, 54(7),* 477-481.

Cebrian, J. (2002). Phonetic similarity, syllabification and phonotactic constraints in the acquisition of a second language contrast. *Toronto Working Papers in Linguistics Dissertation Series.* University of Toronto, Toronto, Canada.

Cenoz, J., & Lecumberri, L. G. (1999). The Effect of Training on the Discrimination of English Vowels. *International Review of Applied Linguistics, 37(4),* 261-275.

Chan, A. Y. W. (2010). Advanced Cantonese ESL learners' production of English speech sounds: Problems and strategies. *System, 38*, 316-328.

Chan, A. Y. W. (2012). Advanced Cantonese ESL learners' production of English speech sounds: Problems and strategies. *TESOL Quarterly, 45,* 718-248.

Chan, A. Y. W., & Li, D. C. S. (2000). English and Cantonese Phonology in Contrast: Explaining Cantonese ESL Learners’ English Pronunciation Problems. *Language Culture and Curriculum, 13,* 67-85.

Chan, C. P. H. (2002). *An investigation into the perception (and production) of English word-initial consonants by native speakers of Cantonese* (Unpublished Master’s thesis). The University of Hong Kong, Hong Kong.

Chang, L. F. (1975). *A comparative analysis of English and Cantonese Phonological Systems* (Unpublished Master’s thesis). The University of Hong Kong, Hong Kong.

Cichocki, W., House, A. B., & Lister, A. C. (1997). Cantonese Speakers and the Acquisition of French Nasal Vowels. *Revue Parole, 1,* 1-16.

Corder, S. P. (1971). Idiosyncratic errors and error analysis. *International Review of Applied Linguistics*, 9, 147-159.

DeKeyser, Robert (2003). Implicit and explicit learning. In C. Doughty and M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp.313-348). Malden, MA: Blackwell.

Denton, C. A., Cirino, P. T., Barth, A. E., Romain, M., Vaughn, S., Wexler, J., Francis D. J., Fletcher, J. M. (2011). An experimental study of scheduling and duration of “tier 2” first-grade reading intervention. *Journal of Research on Educational*

Effectiveness, *4*, 208-230.

Derwing, T. M., Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 20,* 1-16.

Derwing, T., Thomson, R., & Munro, M. (2006). English pronunciation and fluency development in Mandarin and Slavic speakers. *System, 34,* 183-193.

Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association 27,* 47-55.

Deterding, D, Kirkpatrick, A., & Wong, P. M. J. (2008). The pronunciation of Hong Kong English. *English World-Wide: A journal of varieties of English, 29,* 148-175.

Diehl, R. L., & Kluender, K. R. (1987). On the categorization of speech sounds. In S, Harnad (Ed.), *Categorical Perception* (pp. 226-253). London: Cambridge University Press.

Diehl, R. L., & Kluender. K. R. (1989). On the objects of speech perception. *Ecological Psychology, 1(2),* 121-144.

Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech Perception. *Annual Review of Psychology, 55,* 149-179.

Durie, M., & Hajek, J. (1994). A revised standard phonemic orthography for Australian English vowels. *Australian Journal of Linguistics, 14,* 93-107.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27,* 315-330.

Eckman, F. R. (1981a). On predicting phonological difficulty in second language acquisition. *Studies in Second Language Acquisition, 4,* 18-30.

Eckman, F. R. (1981b). On the naturalness of Interlanguage phonological rules. *Language Learning, 31,* 195-216.

Eimas, P. D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics, 16,* 513-521.

Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. *Perception and Psychophysics, 18,* 341-347.

Eimas, P. D., & Miller, J. L. (1980a). Contextual effects in infant speech perception. *Science, 209,* 1140-1141.

Eimas, P. D., & Miller, J. L. (1980b). Discrimination of the information for manner of articulation. *Infant Behavior and Development, 3,* 367-375.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. & Vigorito, J. (1971). Speech perception in infants. *Science, 171,* 303-306.

Ellis, R. (1990). *Instructed second language learning.* Oxford: Blackwell.

Escudero, P. (2001). The role of the input in the development of L1 and L2 sound contrasts: language-specific cue weighting for vowels. In A. H.-J. Do, L. Dominguez & A. Johansen (Eds.). *Proceedings of the 25th Annual Boston University Conference on Language Development* (pp. 250-261). Cascadilla.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization* (Doctoral dissertation, Ultrecht University). Retrieved from http://www.fon.hum.uva.nl/theses/PaolaEscuderoPhD2005.pdf.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26,* 551-585.

Escudero, P., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *Journal of the Acoustical Society of America, 128 (5),* EL254-EL260.

Escudero, P.,& Vasiliev, P. (2011). Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *Journal of the Acoustical Society of America. 130(5),* EL277-EL283.

Escudero, P., & Williams, D. (2011). Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners. *Journal of the Acoustical Society of America, 129,* EL1-EL7.

Famoso, C. F., Schwartz, P. N., & DaSilva, A. (1998). Comparison of American English vowel production and identification by speakers of Luso-Portuguese. *Paper presented at the 16th International Congress on Acoustics and the 135th Meeting of Acoustical Society of America,* Seattle.

Fein, B. G., Golman, M. G., Kone, H. J., & McClintock, C. R. (1956). Effective utilization of staff time in public school speech correction. *Journal of Speech and Hearing Disorders, 21,* 283-291.

Ferguson, C. A., & Macken, M. A. (1980). Phonological development in children: play and cognition. *Papers and Reports on Child Language Development, 18,* 138-177.

Fey, M.E., Warren, S.F., Brady, N., Finestack, L.H., Bredin-Oja, S. L., Fairchild, M., Sokol, S., & Yoder, P. J. (2006). Early effects of responsivity education/prelinguistic milieu teaching for children with developmental delays and their parents. *Journal of Speech, Language, and Hearing Research, 49,* 526-547.

Field, J. (2005). Intelligibility and the listener: the role of lexical stress. *TESOL Quarterly, 39(3),* 399-423.

Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics, 15,* 47-65.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84,* 70-79.

Flege, J. E. (1989). Chinese subjects’ perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America, 86,* 1684-1697.

Flege, J. E. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Huebner and C Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 249-290). Philadelphia: John Benjamins Publishing Company.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America, 93,* 1589-1608.

Flege, J. E. (1995a). Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics, 16,* 425-442.

Flege, J. E. (1995b). Second-language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistics experience: Issues in cross-language research* (pp. 233-272). Timonium, MD: York Press.

Flege, J. E. (1999). Age of learning and second-language speech. In D. Birdsong (Ed.), *New perspectives on the critical period hypothesis for second language acquisition* (pp. 101-132). Mahwah, NJ, Lawrence Erlbaum.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, and A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217-244). Trier, Germany: Wissenschaflicher Verlag Trier.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer and N. Schiller (Eds.) *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319-355). Berlin: Mouton de Gruyter.

Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics, 34,* 153-175.

Flege, J. E., & Bohn, O. S. (1989). The perception of English vowels by native Spanish speakers. *Journal of the Acoustical Society of America, 85,* S85(A).

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25,* 437-470.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15,* 67-83.

Flege, J. E. & Fletcher, K. L. (1992), Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America, 91,* 370-389.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25,* 169-186.

Flege, J. E., Frieda, E. M., Walley, A. C., & Randazza, L. A. (1998). Lexical Factors and Segmental Accuracy in Second Language Speech Production. *Studies in Second Language Acquisition, 20(1),* 155-187.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition, 23,* 527-552.

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26,* 1-34.

Flege, J. E., Mackay, I. R., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America, 106,* 2973-2987.

Flege, J. E., MacKay, I., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics, 23,* 567-598.

Flege, J. E., Munro, M., & MacKay, I. (1995). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97,* 3125-3134.

Flege, J. E., Munro, M. J., & Skelton, L. (1992). Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America, 92,* 128-143.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech, 38,* 25-55.

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Journal of the Acoustical Society of America, 99,* 1161-1173.

Fowler, C. A. (1981). Production and perception of coarticulation among stress and unstressed vowels. *Journal of Speech and Hearing Research, 24,* 127-139.

Fowler, C. A. (1984). Segmentation of coarticulated speech in perception. *Perception & Psychophysics, 36,* 359-368.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics, 14,* 3-28.

Fowler, C. A. (1989). Real Objects of Speech Perception: A Commentary on Diehl and Kluender. *Ecological Psychology, 1,* 145-160.

Fowler, C. A. (1990). Sound-producing sources as objects of perception: Rate normalization and nonspeech perception. *Journal of the Acoustical Society of America, 88(3),* 1236-1249.

Fowler, C. A. (1994). Speech perception: direct realist theory. In R. E. Asher (Ed.) *The Encyclopedia of Language and Linguistics* (pp.4199-4203). Oxford: Pergamon.

Fowler, C. A. (1996). Listeners do hear sounds not tongues. *Journal of the Acoustical*

Society of America, 99, 1730-1741.

Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 349-366.

Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception and Psychophysics, 62,* 1668-1680.

Francis, A. L., Nusbaum, H. C., & Fenn, K. (2007). Effects of training on the acoustic phonetic representation of synthetic speech. *Journal of Speech Language and Hearing Research*, *50,* 1445-1465.

Frieda, E. M., Walley, A. C., Flege, J. E., & Sloane, M. E. (2000). Adults' perception and production of the English vowel /i/. *Journal of Speech, Learning and Hearing Research, 43,* 129-143.

Fries, C. C. (1945). *Teaching and learning English as a foreign language.* Ann Arbor: University of Michigan Press.

Gerken, L., & Aslin, R. N. (2005). Thirty years of research on infant speech perception: the legacy of Peter W. Jusczyk. *Language Learning and Development, 1,* 5-21.

Gibson, E. J. (1969). *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts.

Gibson, E. J. (1991). *An odyssey in learning and perception (Learning, development, and conceptual change).* Cambridge, Mass: MIT Press.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Boston, MA: Houghton Mifflin.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Boston MA: Houghton Mifflin.

Gilichinskaya, Y. D., & Strange, W. (2010). Perceptual assimilation of American English vowels by inexperienced Russian listeners. *Journal of the Acoustical Society of America, 128,* EL80-EL85.

Gillam, R. B., Loeb, D. F., & Friel-Patti, S. (2001). Looking back: A summary of five exploratory studies of Fast For Word. *American Journal of Speech-Language Pathology, 10*, 269-273.

Gillam, R. B., Loeb, D. F., Hoffman, L. M., Bohman, T., Champlain, C. A., Thibodeau, L., et al. (2008). Efficacy of Fast ForWord language intervention in school-age children with language impairment: A randomized controlled trial. Journal of Speech, Language, and Hearing Science, 51, 97-119.

Goldinger, S. (1990). Effects of talker variability on self-paced serial recall. *Research on speech perception.* Bloomington: Indiana University Press.

Goldinger, S. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166-1183.

Goldinger, S. (1997). Perception and production in an episodic lexicon. In K. Johnson and J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp.33-66). New York: Academic Press.

Goldinger, S., Pisoni, D., & Logan, J. S. (1991). On the nature of talker variability effects on serial recall of spoken word lists. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, 17, 152-162.

Gómez Lacabex, E. (2009). Perception and Production of Vowel Reduction in Second Language Acquisition (Unpublished doctoral dissertation). Universidad del Pais Vasco: Leios.

González-Bueno, M. (1997). The effects of formal instruction on the acquisition of Spanish stop consonants. In W. Glass & A. T. Pérez-Leroux (Eds.), *Contemporary perspectives on the acquisition of Spanish, vol. 2: Production, processing, and comprehension* (pp. 57-75). Somerville, MA: Cascadilla Press.

Gordon, J., Darcy, I., & Ewer, D. (2012). *Pronunciation Teaching and Learning: Effects of Explicit Phonetic Instruction in the L2 classroom.* Poster presented at Pronunciation in Second Language Learning and Teaching (PSLLT) 2012, Vancouver, Canada.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds “l” and “r”. *Neuropsychologia, 9,* 317-323.

Gottfried, T. L., & Beddor, P. S. (1988). Perception of temporal and spectral information in French vowels. *Language and Speech, 31,* 57-75.

Gray, S. (2003). Word-learning by preschoolers with specific language impairment: What predicts success? *Journal of Speech, Language, and Hearing Research, 46,* 56-57.

Gray, S. (2005). Word-learning by preschoolers with specific language impairment: Effect of phonological or semantic cues. *Journal of Speech, Language, and Hearing Research, 48,* 1452-1467.

Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology, 25,* 577-588.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107,* 2711-2724.

Guion, S. G., & Pederson, E. (2002). The role of orienting attention for learning novel phonetic categories. *The Institute of Cognitive and Decision Sciences at the University of Oregon, Technical Reports,* No. 02-6.

Guion, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.-S. Bohn & M. J. Munro (Eds.), *Second-language speech learning: The role of language experience in speech perception and production: A festschrift in honour of James E. Flege* (pp. 57-77). Amsterdam: John Benjamins.

Hahn, L. D. (2004). Primary Stress and Intelligibility: Research to Motivate the Teaching of Suprasegmentals. *TESOL Quarterly, 38(2),* 201-223.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical period hypothesis for second language acquisition. *Psychological Science, 14,* 31-38.

Handbook of the International Phonetic Association. (1999). Cambridge, UK: Cambridge University Press.

Hardison, D. (2003). Learning to form new L2 phonetic categories in sentence context. In M. J. Sole, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth*

International Congress of Phonetic Sciences (pp. 515-518). Barcelona, Spain.

Harnsberger, J. D. (2000). A cross-language study of the identification of non-native nasal consonants varying in place of articulation. *Acoustical Society of America, 108,* 764- 783.

Harnsberger, J. D. (2001). On the relationship between identification and discrimination on non-native nasal consonants. *Acoustical Society of America, 110,* 489-503.

Hattori, K. (2009). *Perception and Production of English /r/-/l/ by Adult Japanese Speakers* (Unpublished doctoral dissertation). University College London, London.

Harroti, K., & Iverson, P. (2008). English /r/-/l/ pronunciation training for Japanese speakers. *Journal of the Acoustical Society of America, 123,* 3327.

Hattori, K., & Iverson, P. (2009). English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America, 125,* 469-479.

Hashway, K. (n.d.). April Fools. *Reading Composition Worksheet for 5th grade*. Retrieved October 21, 2011, from www.superteacherworksheets.com.

Hazan, V., & Barrett, S. (2000). The development of phonemic categorisation in children aged 6 to 12. *Journal of Phonetics, 28,* 377-396.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication, 47,* 360-378.

Heeren, W., & Schouten, M. E. H. (2008). Perceptual development of phoneme contrasts: How sensitivity changes along acoustic dimensions that contrast phoneme categories. *Journal of the Acoustical Society of America, 124,* 2291-2302.

Hensman, B. (1969). The phonological problems of Chinese students of English as a second language. *Studium, 2,* 197-210.

Hewlett, N. (1990). Processes of development and production. In P. Grunwell (Ed.), *Developmental speech disorders* (pp. 15-38). Edinburgh, U.K.: Churchill Livingstone.

Higgins, J. (2013). *Minimal pairs for English RP.* Retrieved 5 July 2013 from http://myweb.tiscali.co.uk/wordscape/wordlist/minpairs.html.

Hillenbrand, J. M., & Clark, M. J. (2000). Some effects of duration on vowel recognition. *Journal of Acoustical Society of America, 108,* 3013-3022.

Hillenbrand, J., Getty, L., Clark, M., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97,* 3099-3111.

Hillenbrand, J. M., & Nearey, T. M. (1999). Identification of resynthesized /hVd/ utterances: Effects of formant contour. *Journal of the Acoustical Society of America, 109,* 3509-3523.

Hirata, Y. (2004). Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *Journal of the Acoustical Society of America, 116,* 2384-2394.

Hirata, Y. & Kelly, S. (2010). Effects of lips and hands on auditory learning of second-language speech sounds. *Journal of Speech, Language, and Hearing Research, 53(2),* 298-310.

Hirata, Y., Whitehurst, E., & Cullings, E. (2007). Training native English speakers to identify Japanese vowel length contrast with sentences at varied speaking rates. *Journal of the Acoustical Society of America, 121,* 3837-3845.

Højen, A., & Flege, J. (2006). Early learners' discrimination of second-language (L2) vowels. *Journal of the Acoustical Society of America. 119,* 3072-3084.

Holt, L. L., & Lotto, A. J. (2008). Speech perception within an auditory cognitive science framework. *Current Directions in Psychological Science, 17,* 42-46.

Houston, D. M. (2011). Infant Speech Perception. In R. Seewald and A. M. Tharpe (Eds.), *Comprehensive Handbook of Pediatric Audiology* (pp. 47-62). San Diego, CA: Plural Publishing, Inc.

Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics, 30,* 225- 249.

Hung, T. T. N. (2002). Towards a Phonology of Hong Kong. In K. Bolton (Ed.), *Hong Kong English: Autonomy and Creativity* (pp. 119-140). Hong Kong: Hong Kong University Press.

Ingram, J. C. L., & Park, S. G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics, 25,* 343-370.

Iverson, P. & Evans, B. G. (2007a). Auditory training of English vowels for first language speakers of Spanish and German. In: *16th International Congress of Phonetic Sciences*, pp. 1625-1628, Saarbrucken, Germany.

Iverson, P. & Evans, B. G. (2007b) Learning English vowels with different first language vowel systems: Perception of formant targets, formant movement, and duration. *Journal of the Acoustical Society of America, 122,* 2842-2854.

Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first language vowel systems II: Auditory training for native Spanish and German speakers. *Journal of the Acoustical Society of America, 126(2),* 866-877.

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America, 97,* 553-562.

Iverson, P., & Kuhl, P. K. (1996). Influences of phonetic identification and category goodness of American listeners' perception of /r/ and /l/. *Journal of the Acoustical Society of America, 99,* 1130-1140.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. S. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87,* 47-57.

Iverson, P., Pinet, M. & Evans, B. G. (2011). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics, 33(1),* 145 - 160.

Jakcobson, R. (1941). *Child Language, aphasia, and phonological universals.* The Hague: Mouton.

Jamieson, D., & Morosan, D. (1986). Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception and*

Psychophysics, 40, 205-215.

Jenkins, J. (2000). *The phonology of English as an international language: new models, new norms, new goals.* Oxford: University Press.

Jenkins, J. J., Strange, W., & Trent, S. A. (1999). Content-independent dynamic informationfor the perception of coarticulated vowels. *Journal of the Acoustical Society of America, 106,* 438-448.

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the U.S. *Applied Psycholinguistics, 24,* 131-161.

Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *Journal of the Acoustical Society of America, 119,* 1118-1130.

Johnson, J. & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Johnson, K. (1997). Speech perception without speaker normalization: an exemplar model. In K. Johnson and J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145-166). California: Academic Press.

Jongman, A., & Wade, T. (2007). Acoustic variability and perceptual learning: The case of non-native accented speech. In O.-S. Bohn & M. J. Munro (Eds.), *Second language speech learning: The role of language experience in speech perception and production: A festschrift in honour of James E. Flege,* (pp. 133-166).

Amsterdam: John Benjamins.

Jusczyk, P. W. (1989). *Developing phonological categories form the speech signal.* Paper presented at International Conference on Phonological Development, Stanford University.

Jusczyk, P. W. (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.

Kachru, B. (1992). *The Other Tongue: English across cultures.* University of Illinois Press.

Kim, Y. H., & Hazan, V. (2010). Individual variability in the perceptual learning of L2 speech sounds and its cognitive correlates. In Dziubalska-Kolaczyk, K. and Wrembel, M. and Kul, M., (Eds.) *New Sounds 2010: Proceedings of the Sixth International Symposium on the Acquisition of Second Language Speech.* School of English, Adam Mickiewicz University, Poznań, Poland.

King, R. D. (1967). A measure of functional load. *Studia Linguistica, 21(1),* 1-14.

Kingston, J. (2003). Learning foreign vowels. *Language and speech, 46,* 295-349.

Kingston, J., & Diehl, R. L. (1994). Phonetic knowledge. *Language, 70,* 419-454.

Kingston, J, & Diehl, R. L. (1995). Intermediate properties in the perception of distinctive feature values. In B. Connell, A. Arvaniti, (Eds.), *Phonology and Phonetics: Papers in Laboratory Phonology* (pp. 7-27). London: Cambridge University Press.

Kochetov, A. (2004). Processing of secondary articulation contrasts in different syllable positions: Universal and language-particular factors. *Cahiers Linguistiques d'Ottawa, 32*, 24-43.

Kondaurova, M., & Francis, A. L. (2004). Perception of the English tense/lax vowel contrast by native speakers of Russian. *Journal of the Acoustical Society of America, 116,* 2572-2573.

Kondaurova, M. (2008). *Training to ignore vs. training to attend: The distribution of selective attention in the acquisition of a foreign phonetic contrast* (Unpublished doctorial dissertation). Purdue University, Indiana.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50,* 93-107.

Kuhl, P. K. (1992). Psychoacoustics and speech perception: internal standards, perceptual anchors, and prototypes. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 293-332). Washington, DC: American Psychological Association.

Kuhl, P. K. (1993). Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In B. de-Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 259-274). Dordrecht, Netherlands: Kluwer Academic Publishers.

Kuhl, P. K. (2000a). Language, mind, and brain: Experience alters perception. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp.99-115). Cambridge, MA: MIT Press.

Kuhl, P. K. (2000b). *A new view of language acquisition.* National Academy of Sciences

USA, 97, 11850-11857.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience, 5(11),* 831-843.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society, B, 363,* 979-1000.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: implications for the "critical period". *Language Learning and Development, 1,* 237-264.

Kuhl, P. K., & Iverson, P. (1995) Linguistic experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistics experience: Issues in cross-language research* (pp. 121-154). Timonium, MD: York Press.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science, 218,* 1138-1141.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: vocal imitation and developmental change. *Journal of the Acoustical Society of America, 100,* 2425-2438.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100,* 9096-9101.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992).

Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age. *Science, 255,* 606-255.

Ladefoged, P. (2005). *Vowels and consonants: an introduction to sounds of languages.* Oxford: Blackwell.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.* University of Michigan Press: Ann Arbor.

Lado, R. (1964). *Language teaching: A scientific approach.* New York: McGraw-Hill, Incorporation.

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics, 26,* 227-247.

Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four and six and a half month infants from Spanish environments. *Journal of Experimental Child Psychology, 20*, 215-225.

Leather, J. (1983). Second language pronunciation and teaching. *Language Learning, 16,* 198-219.

Leather, J. (1997). Interrelation of perceptual and productive learning in the initial acquisition of second-language tone. In A. James and J. Leather (Eds.), *Second-Language Speech: Structure and Process* (pp. 75-101). Berlin: Mouton de Gruyter.

Leather, J. (1999). Second-Language Speech Research: An Introduction. *Language*

Learning 49(1), 1-56.

Leather, J. & James, A. (1996). Second language speech. In W. C. Ritchie & T. K. Bhatia, (Eds.), *Handbook of Second Language Acquisition* (pp.269-316). San Diego, CA: Academic Press.

Lee, J., Perrachione, T., Dees, T., & Wong, P. C. M. (2007). *Differential effects of stimulus variability and learners' pre-existing pitch perception ability in lexical tone learning by native English speakers.* Paper presented in 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.

Lee, J. B., Kaye, R. C., & Cherney, L. R. (2009). Conversational script performance in adults with non-fluent aphasia: Treatment intensity and aphasia severity. *Aphasiology, 23,* 885-897.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105,* 1455-1468.

Lee, W. (1968). Thoughts on contrastive linguistics in the context of language teaching. In J. Altis (Ed.), *Report of the nineteenth Annual Round Table Meeting on Linguistics and Language Studies* (pp.185-194). Washington D. C.: Georgetown University.

Lengeris, A. (2009). *Individual differences in second-language vowel training* (Doctoral dissertation). Retrieved from UCL Discovery.

Lenneberg, E. H. (1967). *Biological Foundations of Language.* New York, NY: Wiley.

Leow, R. (1997). Attention, awareness, and foreign language behavior. *Language*

Learning, 47, 467-501.

Leow, R. (1998). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics, 19,* 133-159.

Leow, R. P. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania, 84,* 496-509.

Leung, C. B., & Brice, A. E. (2012). An Analysis of Phonological Processes Involved in Spoken English of Hong Kong Primary Pre-service Teachers. *Language Testing in Asia, 2(2),* 42-60.

Liberman, A. M. (1996). Introduction: some assumptions about speech and how they changed. *Speech: A Special Code.* Cambridge, MA: MIT Press.

Liberman, A. M., Cooper F. S., Shankweiler, D. P., & Studert-Kennedy M. (1967). Perception of the speech code. *Psychological Review, 74,* 431-461.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21,* 1-36.

Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science, 245,* 489-494.

Lisker, L., & Abramson, A. S. (1967). The voicing dimension: Some experiments in comparative phonetics. *Haskins Laboratories Status Report on Speech Research, SR11,* 9-15.

Liu, H.-M., Kuhl, P. K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6,* F1-F10.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /ɹ/ and /l/ II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America, 94,* 1242-1255.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tokhura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /ɹ/ and /l/. III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, *96,* 2076-2087.

Lo, S. K. (2007). *The Markedness Differential Hypothesis and the Acquisition of English Final Consonants by Cantonese ESL Learners in Hong Kong* (Unpublished MPhil thesis). City University of Hong Kong, Hong Kong

Lo, W. K. (2000). *Cantonese Phonology and Phonetics: An Engineering Introduction.* Unpublished manuscript. The Chinese University of Hong Kong, Hong Kong.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /ɹ/ and /l/: A first report. *Journal of the Acoustical Society of America*, *89*, 874-886.

Logan, J. S., & Pruitt, J. S. (1995). Methodological issues in training listeners to perceive non-native phonemes. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 351-377). Baltimore, MD: York Press.

Long, M. (1990). Maturational constraints on language development. *Studies in Second*

Language Acquisition, 12, 251-285.

Lord, G. (2005). (How) can we teach foreign language pronunciation? The effects of a phonetics class on second language pronunciation. *Hispania, 88(3),* 557-567.

Lotto, A. J. (2000). Language acquisition as complex category formation. *Phonetica, 57,* 189-196.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19,* 37-66.

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English /ɹ/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics, 2*, 369-390.

Mackay, I. R. A., Flege, J. E. & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholiguistics, 27,* 153-183.

Mackay, I. R. A., Meador, D., & Flege, J. E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, *58,* 103-125.

MacLeod, A. N., Stoel-Gammon, C., & Wassink, A. B. (2009). Production of high vowels in Canadian English and Canadian French: a comparison of early bilingual and monolingual speakers. *Journal of Phonetics, 37(4),* pp 374-387.

Major, R. C. (1987). A model for interlanguage phonology. In G. L. Ioup & S. L. Weinberger (Eds.), *Interlanguage phonology* (pp. 101-124). Rowley, MA: Newbury House.

Major, R. C. (1994). Current trends in interlanguage phonology. In M. Yavas (Ed.), *First* *and second language phonology* (pp.181-204). San Diego: Singular.

Major, R. C. (1998). Interlanguage phonetics and phonology - an introduction. *Studies in Second Language Acquisition, 20,* 131-137.

Major R., & Faudree, M. (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition, 18,* 69-90.

Major, R. C., & Kim, E. (1996). The similarity differential rate hypothesis. *Language Learning*, 46, 465-496.

Makarova, A. O. (2010). *Acquisition of Three Vowel Contrasts by Russian Speakers of American English* (Unpublished doctoral dissertation). Harvard University, Cambridge.

Martin, C., Mullennix, J., Pisoni, D., & Summers, W. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning Memory Cognition, 15,* 676-684.

Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: rounded vowels in English learners of German. *Bilingualism, Language and Cognition, 13*, 279-297.

Massaro, D. W., Bigler, S., Chen, T., Perlman, M., / Ouni, S. (2008). Pronunciation training: the role of eye and ear. *INTERSPEECH-2008*, 2623-2626.

Matthews, S. and Yip, V. (1994). *Cantonese: A Comprehensive Grammar.* London: Routeledge.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82,* B101-B111.

McAllister, R., Flege, J. E., & Piske, T. (2002) The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30,* 229-258.

McCandliss, B. D., Fiez, J. A., Protopapas, A, Conway, M., & McClellang, J. L. 2002. Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of Plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience, 2,* 89-108.

McClaskey, C. L., Pisoni, D. B., & Carrell, T. D. (1983). Transfer of training of a new linguistic contrast in voicing. *Perception and Psychophysics, 34,* 323-330.

McCrocklin, S. (2012). Effect of audio vs. video listening exercises on aural discrimination of vowels. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference, Sept. 2011. (pp. 178-186).* Ames, IA: Iowa State University.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746-748.

McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition, 95(2),* B15-B26.

McNerny, M. & Mendelsohn, D. (1992). Suprasegmentals in the pronunciation class: Setting priorities. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 185-196). Oxford: Oxford University Press.

Meng, H., Zee, E., & Lee W. S. (2007). *A Contrastive Phonetic Study between*

Cantonese and English to Predict Salient Mispronunciations by Cantonese Learners of English. Unpublished article. The Chinese University of Hong Kong.

Miller, J. D., Wier, C. C., Pastore, R. E., Kelly, W. J., & Dooling, R. J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: an example of categorical perception. *Journal of the Acoustical Society of America*, *60*, 410-417.

Minnik-Fox, M. & Maeda, K. (1999.) Categorization of American English vowels by Japanese speakers. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. Baily (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1437-1440). San Fransciso.

Morley, J. (1994). *Pronunciation Pedagogy and Theory.* Virginia: Teachers of English to speakers of other languages.

Morrison, G. S. (2006). *L1 & L2 Production and Perception of English and Spanish Vowels A Statistical Modeling Approach* (Doctoral dissertation, University of Alberta). Retrieved from http://geoff-morrison.net/documents/Morrison%20(2006)%20L1%20%26%20L2%20Production%20and%20Perception%20of%20English%20and%20Spanish%20Vowels%20-%20A%20Statistical%20Modelling%20Approach.pdf.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation and instruction. *Studies in Second Language Acquisition, 21,* 81-108.

Moyer, A. (2004). *Age, accent and experience in second language acquisition: An integrated approach to critical period inquiry.* Clevedon, UK: Multilingual

Matters.

Moyer, A. (2009). Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. In M. Young-Scholten & T. Piske (Eds.), *Input matters in SLA* (pp. 159-174). Bristol, UK: Multilingual.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, *85,* 365-378.

Munro, M. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness rating. *Language and Speech*, 36, 39-66.

Munro, M. J. & Derwing, T.M. (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning, 45,* 73-97.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49, (Supplement 1),* 285-310.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34,* 520-531.

Munro, M., Flege, J., & MacKay, I. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics, 17,* 313-334.

Munro, M., & Mann, V. (2005). Age of immersion as a predictor of foreign accent.

Applied Psycholinguistics, 26, 311-341.

Munro, M. J., Wang, X., & Li, C. H. (2000). *Asymmetries in English high vowel production by Cantonese speakers: Findings and methodological issues.* Manuscript submitted for publication.

Nagao, K., Lim, B.-J., & de Jong, K. (2003). Perceptual acquisitions of non-native syllable structures by native listeners of Japanese. In *16th International Congress of Phonetic Sciences*, Barcelona, Spain.

Neufeld, G. G. (1977). Language Learning Ability in Adults: A Study on the Acquisition of Prosody and Articulatory Features. *Working Papers on Bilingualism, 12,* 46-60.

Neufeld, G. G. (1978). On the Acquisition of Prosodie and Articulatory Features in Adult Language Learning. *The Canadian Modern Language Review, 34,* 163-174.

Neufeld, G. G. (1979). Towards a Theory of Language-Learning Ability. *Language Learning, 29,* 227-241.

Ng, M., & Chu, R. (2009). An acoustical and perceptual study of vowels produced by alaryngeal speakers of Cantonese. *Folia Phoniatrica et Logopaedica, 61,* 97-104.

Nielsen, K. Y. (2007). The Interaction between Spontaneous Imitation and Linguistic Knowledge. *Working Papers in Phonetics, 105,* 125-137.

Nishi, K., & Kewley-Port, D. (2005). Training Japanese listeners to identify American English vowels. *Journal of the Acoustical Society of America, 117,* 2401.

Nishi, K., & Kewley-Port, D. (2007a). Second language vowel perception training: Effects of set size, training order, and native language. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1621-1624). Saarbrucken, Germany.

Nishi, K., & Kewley-Port, D. (2007b). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research, 50,* 1496-1509.

Nishi, K., & Kewley-Port, D. (2008). Nonnative speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech, Language, and Hearing Research, 51(6),* 1480-1493.

Nobre-Oliveira, D. (2007). Effects of Perceptual Training on the Learning of English Vowels in Non-native Settings. In A. S. Rauber, M. A. Watkins, & B. O. Baptista (Eds.), *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech* (pp.382-389). Florianópolis, Brazil: Federal University of Santa Catarina.

Nosofsky, R. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psycholinguistics: General, 115,* 39-57.

Nosofsky, R. (1987). Attention and learning processes in the identification and categorization of integral stimuli, *Journal of Experimental Psycholinguistics: Learning Memory Cognition, 3,* 7-18.

Ohala, J. J. (1996). Speech perception is hearing sounds, not tongues. *Journal of Acoustic Society of America, 99,* 1718-1725.

Oyama, S. (1976). A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research, 5(3),* 261-283.

Patkowski, M. S. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning, 30,* 449-472.

Patkowski, M. S. (1990). Age and accent in second language: A reply to James Emil Flege. *Applied Linguistics, 11,* 73-89.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12,* 445-449.

Pederson, E. & Guion-Anderson, S. G. (2010). Orienting attention during phonetic training facilitates learning. *Journal of the Acoustic Society of America, 127,* 1-6.

Pennington, M. C. (1987). Acquiring proficiency in English phonology: Problems and solutions for the Japanese learner. *Gaikokugo Kyoiku Kiyo, 16,* 1-20. Nagoya, Japan: Nagoya Gakuin University.

Pennington, M. C. (1996). *Phonology in English language teaching: An international approach.* London: Longman.

Pennington, M. C. (1997). Phonology in language teaching: Essentials of theory and practice. In K. Bardovi-Harlig & B. S. Hartford (Eds.), *Beyond methods: Companion components in language teacher education* (pp. 67-87). New York: McGraw-Hill.

Pennington, M. C. (1998). The teachability of phonology in adulthood: A re-examination. *International Review of Applied Linguistics, 46,* 323-341.

Perrachione T. K., Lee, J., Ha L.Y., Wong P. C. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *Journal of Acoustical Society of America, 130,* 461-472.

Perry, F. L. (2005). *Research in applied linguistics: Becoming a discerning consumer.* New Jersey: Lawrence Erlbaum Associates.

Piper, T. & Cansin, D. (1988). Factors influencing the foreign accent. *The Canadian Modern Language Review, 44,* 334-342.

Piske, T., Flege, J.E., Mackay, I.R.A. & Meador, D. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica, 59,* 49-71.

Piske, T., Mackay, I.R.A. & Flege, J.E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29,* 191-215.

Pisoni, D. (1977). Identification and discrimination of the relative onset time of two component tones: Implications for voicing perception in stops. *Journal of Acoustic Society of America, 61,* 1352-1361.

Pisoni, D. (1990). Effects of talker variability on speech perception: implications for current research and theory. *Proceedings of the 1990 International Conference on Spoken Language Processing, Kobe, Japan* (pp. 1399-407). Tokyo: The Acoustical Society of Japan.

Pisoni, D. (1992). Some comments on invariance, variability and perceptual normalization in speech perception. *Proceedings of the 1992 International Conference on Spoken Language Processing, Banff, Canada* (pp. 587-590). Edmonton: University of Alberta.

Pisoni, D. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson and J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9-32). New York: Academic Press.

Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology, 8,* 297-314.

Pisoni, D. B., & Lively, S. E. (1995). Variability and invariance in speech perception: A new look at some old problems in perceptual learning. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 433-459). Baltimore: York Press.

Pisoni, D., Nusbaum, H., Luce, P., & Slowiaczek, L. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication, 4*, 75-95.

Podlipsky, V. J. (2005). Greek perception and production of an English vowel contrast: A preliminary study. *Journal of the Acoustical Society of America, 117,* 2401.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America, 89*, 2961-2977.

Polka, L. (1992). Characterizing the influence of native language experience on adult speech perception. *Perception & Psychophysics, 52,* 37-52.

Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America, 97,* 1286-1296.

Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America, 100*, 577-592.

Polka, L., & Werker, J. F. (1994). Developmental changes in the perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 421-435.

Posner, M., & Keele, S. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353-363.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25-42.

Pruitt, J.S. (1993). Comments on "Training Japanese listeners to identify English /r/ and /l/: A first report" [J. S. Logan, S. E. Lively, and D. B. Pisoni, J. Acoust. Soc. Am. 89, 874-886 (1991)]. *Journal of the Acoustical Society of America, 94,* 1146-1147.

Pruitt, J. S. (1995). Perceptual training on Hindi dental and retroflex consonants by native English and Japanese speakers. *Journal of the Acoustical Society of America, 97,* 3417.

Pruitt, J. S., Jenkins, J. J., & Strange, S. (2006). Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese. *Journal of the Acoustical Society of America, 119,* 1684-1696.

Richards, J. C., Platt, J., & Platt, H. (1992). *Longman Dictionary of Language Teaching and Applied Linguistics.* Essex, England: Longman.

Riches, N. G., Tomasello, M., Conti-Ramsden, G. (2005). Verb learning in children with SLI: Frequency and spacing effects. *Journal of Speech, Language, and Hearing Research, 42,* 1397-1411.

Rochet, B. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in* *cross-language speech research* (pp. 379-410). Timonium, MD: York Press.

Rogers, C. L. (1997). *Intelligibility of Chinese-Accented English* (Unpublished doctoral dissertation). Indiana University, Indiana.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

Schmidt, R. W. (1994a). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.

Schmidt, R. W. (1994b). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 165-209). San Diego, CA: Academic Press.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209-231.

Sewell, A. (2009). World Englishes, English as a Lingua Franca and the case of Hong Kong English. *English Today, 25*, 37-43.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics, 3,* 243-261.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Applied Psycholinguistics, 84,* 127-190.

Shriberg, L. D., Kwiatkowski, J., & Hoffmann, K. (1984). A Procedure for Phonetic Transcription by Consensus. *Journal of Speech and Hearing Research, 27,* 456-465.

Simard, D., & Wong, W. (2001). Alertness, orientation, and detection: The conceptualization of attentional functions in SLA. *Studies in Second Language Acquisition, 23(1),* 103-124.

Simões, A. R. M. (1996). Phonetics in Second Language Acquisition: An Acoustic Study of Fluency in Adult Learners of Spanish. *Hispania, 79(1),* 87-95.

Sleve, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science, 17,* 675- 681.

So, L. K. H., & Thelwall, R. (1992). Hong Kong Spoken Cantonese Database. In J. Pittam (Ed.), *Proceedings of the Fourth Australian International Conference on Speech Science and Technology,* (pp. 718-722). Brisbane: University of Queensland Printery.

So, L. K. H., & Wang, J. (1996). Acoustic distinction of Cantonese long and short vowels. In P. McCormack & A. Russell (Eds.), *Proceedings of the Sixth International Conference on Speech Science and Technology,* (pp. 379-384). Adelaide, South Australia.

Sperbeck, M., Strange, W., & Ito, K. (2005). Training Japanese L2 learners to perceive difficult American vowel contrasts. *Journal of the Acoustical Society of America, 117,* 2400.

Stevens, K. N. (1972). The Quantal Nature of Speech: Evidence from

Articulatory-Acoustic Data. In E. E. David, Jr. & P. B. Denes (Eds.), *Human Communication: A Unified View* (pp.51-66), New York: McGraw-Hill.

Stevens, N. K., & Blumstein, S. E. (1981). The search for invariant acoustic correlates of phonetic features. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp.1-38). Hillsdale, NJ: Erlbaum.

Stevens, K. N., & Klatt, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America, 55,* 653-659.

Stevens, K. N., Liberman, A. M., Studdert-Kennedy, M., & Ohman, S. E. G. (1969). Crosslanguage study of vowel perception. *Language and Speech, 12,* 1-23.

Strange, W. (1986). Speech input and the development of speech perception. In J. Kavanagh (Ed.), *Otitis Media and Child Development* (pp. 12-26). Parkton, MD: York Press.

Strange, W. (1992). Learning non-native phoneme contrasts: Interactions among subject, stimulus, and task variables. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka, (Eds.), *Speech Perception, Production and Linguistic Structure* (pp. 197-219). Ohmasha. Japan: IOS Press.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 3-45). Baltimore: York Press.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., & Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by

Japanese listeners. *Journal of the Acoustical Society of America, 109,* 1691-1704.

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /ɹ-l/ by Japanese adults learning English. *Perception and Psychophysics, 36,* 131-145.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: the re-education of selective perception. In, J. G. Hansen-Edwards &, M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 153-191). Amsterdam and Philadelphia: John Benjamins Publishing Company.

Strange, W., Yamada, R. A., Kubo, R., Trent, S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics, 26,* 311-344.

Streeter, L.A. (1976). Language perception of 2-month-old infants show both innate mechanisms and experience. *Nature, 259,* 39-41.

Sundara, M., Polka, L., & Genesee, F. (2006). Language experience facilitates discrimination of /d-th/ in monolingual and bilingual acquisition of English. *Cognition, 100(2),* 369-388.

Sussman, H. M., Fruchter, D., Hilbert, J., & Sirosh, J. (1998). Linear correlates in the speech signal: the orderly output constraint. *Behavioural Brain Science, 21,* 241-299.

Tahta, S., Wood, M. & Loewenthal, K. (1981). Foreign accents: Factors relating to transfer of accent from the first language to a second language. *Language and Speech, 24,* 265-272.

Tajima, K., Kato, H., Rothwell, A., Akahane-Yamada, R., & Munhall, K. G. (2008). Training English listeners to perceive phonemic length contrasts in Japanese. *Journal of the Acoustical Society of America, 123,* 397-413.

Taylor, I. & Taylor M. M. (1990). *Psycholinguistics: learning and using language.* Prentice-Hall: Englewood Cliffs.

Thomson, R. I. (2007). *Modelling L1/L2 interactions in the perception and production of English vowels by Mandrin L1 speakers: a training study* (Unpublished doctoral dissertation). University of Alberta, Alberta.

To, C. K. S., Cheung, P. S. P. & McLeod, S. (2012). A population study of children's acquisition of Hong Kong Cantonese consonants, vowels, and tones. *Journal of Speech, Language, and Hearing Research, 56,* 103-122.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16,* 183-203.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34,* 33-58.

Trehub, S. E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psvchology, 9,* 91-96.

Trehub, S.E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development, 47,* 466-472.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals:

Effects of L2 exprierence on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28,* 1-30.

Trofimovich, P., Baker, W., & Mack, M. (2001). Context and experience-based effects on the learning of vowels in a second language. *Studies in the Linguistic Sciences, 31,* 167-186.

Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: a longitudinal study. *Child Development, 75,* 1067-1084.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Crosslanguage tests on adults and infants. *Journal of the Acoustical Society of America, 120(4),* 2285-2294.

Tsushima, T, & Hamada, M. (2005). Relation between perception and production ability during a speech training course. *Journal of the Acoustical Society of America, 117(4),* 2427.

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: Language-specific developmental changes in speech perception abilities. *Proceedings of 1994 International Conference on Spoken Language Processing,* Yokohama, Japan (pp. 1695-1698). Acoustical Society of Japan.

Ukrainetz, T. A., Ross, C. L., & Harm, H. M. (2009). An investigation of treatment scheduling for phonemic awareness of kindergartners who are at risk for reading disabilities. *Language, Speech, and Hearing Services in the Schools, 40,* 86-100.

Van Borsel, J. (1989). The reliability of phonetic transcription: a practical note. *Child Language Teaching and Therapy, 5*, 327-333.

Van Hattum, R. J. (1959). Evaluating elementary school speech therapy. *Exceptional Children, 25*, 411-415.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research.* Norwood, NJ: Ablex.

VanPatten, B. (2002). Processing instruction: An update. *Language Learning, 52,* 755-803.

Wang, X. (1997). The acquisition of English vowels by Mandarin ESL learners: A study of production and perception (Master's thesis). Retrieved from Summit, the SFU Institutional Repository.

Wang, X. (1998). Effects of first language on native Mandarin speakers' perception of English vowels. In K. Lee & M. Oliverie (Eds.), *Proceedings of The 14th Northwestern Linguistics Conference, March 7-8, 1998*. Simon Fraser University, BC, Canada.

Wang, X. (2002). *Training Mandarin and Cantonese Speakers to identify English vowel contrasts: long-term retention and effects on production* (Doctoral dissertation, Simon Fraser University). Retrieved from http://zimmer.csufresno.edu/~xinw/contents.pdf.

Wang, X., & Munro, M. (1998). *English vowel production by Beijing Mandarin speakers*. Paper presented at the 16th International Congress on Acoustics and the 135th Meeting of the Acoustical Society of America, Seattle.

Wang, X., & Munro, M. J. (1999). *The perception of English tense-lax vowel by native Mandarin speakers: The effect of training on attention to temporal and spectral cues.* Paper presented at the 14th International Congress of Phonetic Sciences, San Francisco.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32,* 539-552.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America, 106,* 3649-3658.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America, 113,* 1033-1043.

Warren, S. F., Fey, M. E., Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews, 13(1),* 70-77.

Weinreich, U. (1953). *Languages in contact.* New York: Linguistic Circle of New York.

Werker, J. F. (1994). Cross-language speech perception: Development changes does not involve loss. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 93-120). Cambridge, MA: MIT Press.

Werker, J. F. (1995). Age-related changes in cross-linguistic speech perception: Standing at the crossroads. In W. Strange (Ed.), *Speech perception in linguistic* *experience: Theoretical and methodological issues* (pp. 155-169). Timonium, MD: York Press.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development, 1,* 197-234.

Werker, J. F., Gillbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development,* 52, 1866-1878.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24,* 672-683.

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics, 37,* 35-44.

Werker, J. F., & Polka, L. (1993). Developmental changes in speech perception: new challenges and new directions. *Journal of Phonetics, 21,* 83-101.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of nonnative speech sounds. *Canadian Journal of Psychology, 37,* 278-286.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development, 7*, 49-63.

Weston, A. J., & Harber, S. K. (1975). The effects of scheduling on progress in paired-stimuli articulation therapy. *Language, Speech, and Hearing Services in Schools, 6,* 96-101.

Whitehill, T. L., Ciocca, V., Chan J. C-T., & Samman, N. (2006). Acoustic analysis of vowels following glossectomy. *Clinical Linguistics and Phonetics, 20,* 135-140.

Whitman, R., & Jackson, K. (1972). The unpredictability of contrastive analysis. *Language Learning, 22,* 29-41.

Wong, J. W. S. (2012). Training the Perception and Production of English /e/ and /æ/ of Cantonese ESL Learners: A Comparison of Low vs. High Variability Phonetic Training. In F. Cox, K. Demuth, S. Lin, K. Miles, S. Palethorpe, J. Shaw & I. Yuen (Eds.), *Proceedings of the 14th Australasian International Conference on Speech Science and Technology, 3-6 December 2012* (pp. 37-40). Sydney, Australia: Macquarie University.

Wong, J. W. S. (2013a). Does proficiency matter? - Effects of High Variability Phonetic Training on the Perception and Production of English Vowels by Cantonese ESL learners with high and low proficiency levels. In Mompeán, J. & Fouz, J. (Eds.), *Proceedings of the 3rd International Conference on English Pronunciation: Issues & Practices (EPIP 3)* (pp.121-124). Murcia, Spain: University of Murcia.

Wong, J. W. S. (2013b, to appear). The Effects of Training Diversity in Training the Perception and Production of English Vowels /ɪ/ and /i:/by Cantonese ESL learners. *Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech 2013).* Lyon, France.

Wong, W. S. (2010). *Training the Perception and Production of English Vowels /e/ and /æ/ by Cantonese-speaking Secondary School Students* (Unpublished MPhil Thesis). The Chinese University of Hong Kong, Hong Kong.

Yamada, R. (1993). Effects of extended training on /r/ and /l/ identification by native speakers of Japanese. *Journal of the Acoustical Society of America*, 93, 2391.

Yamada, R. (1995). Age and acquisition of speech language sounds: Perception of American English /ɹ/ and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 305-320). Baltimore, MD: York Press.

Yamada, R. (1999). Toward further understanding of second language speech learning: An approach utilizing speech technology. *Journal of the Acoustical Society of America, 105,* 1032.

Yamada, T., Yamada, R. A., & Strange, W. (1996). Perceptual learning of Japanese mora syllables by native speakers of American English. *Proceedings of the International Conference on Spoken Language Processing,* Yokohama, Japan.

Yoder, P. J. & Warren, S. E. (2001). Relative treatment effects of two prelinguistic communication interventions on language development in toddlers with developmental delays vary by maternal characteristics. *Journal of Speech, Language, and Hearing Research, 44*, 224-237.

Yoder, P. J. & Warren, S. E. (2002). Effects of prelinguistic milieu teaching and parent responsivity education on dyads involving children with intellectual disabilities. *Journal of Speech, Language, and Hearing Research, 45,* 1297-1310.

Zhang, C. C., Peng, G.., & Wang, W. S.-Y. (2011). Effect of language experience on the categorical perception of Cantonese vowel duration. In P. Cosi, R. De Mori, G. Di Fabbrizio, & R. Pieraccini (Eds.), *Proceedings of Interspeech 2011: 12th Annual Conference of the International Speech Communication Association* (pp. 169-172).

Florence, Italy.