

**THE ACQUISITION OF ENGLISH VOWELS BY MANDARIN  
ESL LEARNERS: A STUDY OF PRODUCTION AND  
PERCEPTION**

by

Xinchun Wang

B.A., Shanghai Teachers' University, 1982

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF ARTS

in the Department  
of  
Linguistics

© Xinchun Wang 1997

SIMON FRASER UNIVERSITY

September 1997

All rights reserved. This work may not be  
reproduced in whole or in part, by photocopy  
or other means, without permission of the author



National Library  
of Canada

Acquisitions and  
Bibliographic Services

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque nationale  
du Canada

Acquisitions et  
services bibliographiques

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file* *Votre référence*

*Our file* *Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

88

0-612-24263-3

Canada

## APPROVAL

Name: Xinchun Wang.

Degree: Master of Arts

Title of Thesis: The Acquisition of English Vowels by Mandarin ESL Learners: A Study of Production and Perception

Examining Committee:

Chair: Dr. Nancy Hedberg

---

Dr. Murray Munro  
Senior Supervisor  
Assistant Professor of Linguistics

---

Dr. Zita McRobbie  
Assistant Professor of Linguistics

---

Dr. Ocke-Schwen Bohn, External Examiner  
Professor, Chair of English Linguistics  
English Department  
Aarhus University, Denmark

Date Approved: September 24, 1997

## Abstract

It is generally agreed that adult second language (L2) speech learning is strongly influenced by the learners' first language (L1). This study examines the effect of the L1 vowel system on native Mandarin speakers' production and perception of English vowels.

Fifteen native Mandarin speakers from Beijing who had been living in Canada between 0.5 and 6 years participated as speakers and listeners, and 15 native speakers of Canadian English participated as a comparison group for the production test. The isolated English vowels [i ɪ eɪ ε æ u ʊ ou ɒ ʌ], produced in a carrier sentence by both groups, were identified by four native English listeners. The results showed that, in general, the Mandarin speakers' productions of the vowels that have Mandarin counterparts were as intelligible as the native English speakers' productions and were significantly more intelligible than the vowels lacking obvious Mandarin counterparts.

In a perception task, the Mandarin speakers identified the five front English vowels, [i ɪ eɪ ε æ], produced by 10 native English speakers in /bVt/ utterances. The results of the perception test were not completely consistent with those of the production test, indicating that L2 production may not always be related to perception in a straightforward way and that the relationship between perception and production is a complex issue.

A third experiment compared the temporal and spectral properties of the Mandarin vowels and the Mandarin-accented English vowels with those of the English vowels. The results revealed phonetic differences between Mandarin [eɪ u ʊ a] and English [eɪ u ʊ ɒ ʌ] in

general. Most Mandarin-accented English vowels deviated from the native English norms in terms of spectral values. The deviations in acoustic properties from the native norms seemed to be the major cause of the low intelligibility scores for the Mandarin-accented vowels. In addition, the spectral overlap in the vowel pairs [i]-[ɪ] and [ɛ]-[æ] apparently caused intelligibility problems. The results suggested that the dense distribution of L2 vowels in the area of the vowel space where the L1 vowel system is less crowded may have caused some of the difficulties the L2 learners experienced in producing and perceiving these vowels.

This study also examined the usefulness of two hypotheses in predicting problems faced by Mandarin speakers learning the English vowel system: the Contrastive Analysis Hypothesis and the Speech Learning Model. The data do not fully support either of the hypotheses.

To my parents  
with deep appreciation and love

## Acknowledgments

First of all, I wish to express my deepest gratitude to my supervisor Dr. Murray Munro for his persistent encouragement and invaluable help throughout the years of my graduate studies in the Linguistics Department at Simon Fraser University. This thesis would not have been possible were it not for his guidance in helping me to focus on a topic, his advice on the design of all the experiments, and his careful appraisal of all the drafts.

Special thanks also goes to Dr. Zita McRobbie for first introducing me to experimental phonetics and for her critical supervision of this thesis. The insightful and critical comments by Dr. Ocke-Schwen Bohn were greatly appreciated.

I thank Darleen Crawshaw and Lorna Fadden for their special friendship over the years and for their important contribution to the editing and proof-reading of an earlier draft.

I also thank my fellow graduate students Susan Russell, Kyoung-Ja Lee, Siok Lee, Caroline Morgan, Baoning Fu, Janine Tool and Tara McFarlane for their long-term support and friendship. My thanks is extended to the 34 participants whose contribution made this study possible.

Finally, I thank my husband Yong, my son Ziyue, and my parents back home in Shanghai for their support throughout my years of graduate studies.

## TABLE OF CONTENTS

Title page.....	i
Approval.....	ii
Abstract.....	iii
Dedication.....	v
Acknowledgments.....	vi
Table of Contents.....	vii
List of Tables.....	ix
List of Figures.....	x
Chapter I. Introduction.....	1
1.1. The Effect of Age: The Critical Period Hypothesis.....	1
1.2. The Effect of Language Transfer.....	4
1.2.1. The Contrastive Analysis Hypothesis.....	4
1.2.2. The Perceptual Assimilation Hypothesis.....	5
1.2.3. The Speech Learning Model.....	7
1.3. Non-native Language Transfer Factors.....	11
1.3.1. Linguistic Desensitization Hypothesis.....	11
1.3.2. Second Language Experience.....	13
1.4. The Relationship between Perception and Production.....	14
1.5. Mandarin Chinese Vowels.....	16
1.6. Research Topics.....	21
1.6.1. Outline of the Present Study.....	23
Chapter II. Experiment 1: Intelligibility Test on Mandarin Speakers' English Vowel Productions.....	24
2.1. Method.....	25
2.1.1. Participants.....	25
2.1.2. Recordings.....	26
2.1.3. Stimulus Preparation.....	28
2.1.4. Listeners.....	29
2.1.5. Listening Test.....	29
2.2. Results.....	30
2.2.1. Individual Speaker Differences.....	34
2.2.2. Native English Vowel Differences.....	37
2.3. Discussion.....	38
2.4. Conclusion.....	41
2.5. Limitations.....	43
Tables for Chapter II.....	45
Figures for Chapter II.....	50

Chapter III. Experiment 2: The Perception of English Front Vowels by Mandarin Speakers.....	52
3.1. Method.....	53
3.1.1. Participants.....	53
3.1.2. Materials and Procedures.....	54
3.2. Results.....	55
3.3. Discussion.....	57
3.4. Comparison of Experiment 1 and Experiment 2.....	59
3.5. Conclusion.....	63
Tables for Chapter III.....	67
Figures for Chapter III.....	69
 Chapter IV: Experiment 3; Acoustic Measurements.....	 72
4.1. Method.....	76
4.2. Results.....	77
4.2.1. Mandarin Vowels [i eɪ u ou ɑ] and Their English Counterparts.....	77
4.2.1.1. Duration.....	77
4.2.1.2. Spectral Data.....	77
4.2.1.3. Discussion.....	80
4.2.2. Comparison of Native English and Mandarin- accented vowels [i ɪ eɪ ɛ æ u ʊ ou ɒ ʌ].....	82
4.2.2.1. Duration.....	82
4.2.2.2. Spectral Data.....	83
4.2.2.3. Discussion.....	87
4.2.2.4. Individual Data.....	91
4.3. Conclusion.....	92
Tables for Chapter IV.....	95
Figures for Chapter IV.....	99
 Chapter V. General Discussion.....	 104
Chapter VI. Conclusion.....	115
Appendices	
Appendix 1. English List.....	121
Appendix 2. Mandarin List.....	122
Appendix 3. Language Background Questionnaire.....	124
References.....	126

## List of Tables

### Chapter II

Table 2-1	Background information for the 15 Beijing Mandarin speakers.....	45
Table 2-2	Confusion Matrix (%) for Mandarin speakers' productions of English [i ɪ eɪ ε æ u ʊ ou ɒ ʌ].....	46
Table 2-3	Correct % of identifications (mean & standard deviations) of English and Mandarin speakers' 10 English vowel productions by 4 listeners.....	47
Table 2-4	Correct % of identifications of 15 Mandarin speakers' productions of 10 English vowels and standard deviations.....	48
Table 2-5	Correct % of identifications of 15 English speakers' productions of 10 English vowels and standard deviations.....	49

### Chapter III

Table 3-1	Correct % of identifications of English [i ɪ eɪ ε æ] in /bVt/ by 15 Mandarin listeners.....	67
Table 3-2	Confusion Matrix (%) for Mandarin listeners' perceptions of English [i ɪ eɪ ε æ].....	68

### Chapter IV

Table 4-1	Mean durations (msec) and standard deviations of isolated English vowels [i eɪ u ʊ ɒ ʌ] and Mandarin vowels [i eɪ u ʊ ɑ].....	95
Table 4-2	Mean normalized F1 and F2 frequencies (Hz) and standard deviations of English [ɪ u ɒ ʌ eɪ ou] and Mandarin [ɪ u ɑ eɪ ou].....	96
Table 4-3	Mean durations and standard deviations for English [i ɪ eɪ ε æ u ʊ ou ɒ ʌ] produced by English and Mandarin speakers.....	97
Table 4-4	Mean normalized F1 and F2 frequencies (Hz) and standard deviations for English [i ɪ eɪ ε æ u ʊ ou ɒ ʌ] produced by both groups.....	98

## List of Figures

Chapter II	
Figure 2-1	The mean % of correct identifications of each English vowel by four listeners.....50
Figure 2-2	Individual consistencies in Mandarin speakers' productions of different English vowels.....51
Chapter III	
Figure 3-1	Individual Mandarin listeners' perception scores (%) for English [i ɪ eɪ ε æ] produced by 10 native English speakers.....69
Figure 3-2	Mean correct identifications (%) of English [i ɪ eɪ ε æ] in /bVt/ by Mandarin listeners.....70
Figure 3-3	Mandarin speakers' perception and production scores (%) on English front vowels.....71
Chapter IV	
Figure 4-1	Mean F1 and F2 of English [i eɪ u ou ʌ] produced by 15 English speakers and Mandarin [i eɪ u ou ʌ] produced by 15 Mandarin speakers.....99
Figure 4-2	Mean F1 & F2 of English [i ɪ eɪ ε æ u ʊ ou ɒ ʌ] produced by English and Mandarin speakers, and Mandarin [i eɪ u ou ʌ] produced by Mandarin speakers.....100
Figure 4-3	F1& F2 plot of English [i ɪ ε æ] produced by 15 English and 15 Mandarin speakers.....101
Figure 4-4	F1& F2 plot of English [u ʊ] produced by 15 English and 15 Mandarin speakers.....102
Figure 4-5	F1& F2 plot of English [ou] at measurement (a) and [ɒ] produced by English and Mandarin speakers.....103

# CHAPTER I

## INTRODUCTION

Adults who learn a second language generally do not produce the phones of the second language (L2) in a perfectly native-like manner. This problem remains one of the most enduring and challenging research topics in the field of speech perception and production in foreign language learning. Researchers have examined this issue from various perspectives.

This chapter will discuss some of the important factors that influence foreign/second language speech learning, including the effect of age and first language interference. In addition, some common theoretical approaches to second language speech learning will be discussed along with previous empirical studies. The concepts and hypotheses that will be reviewed are the Critical Period Hypothesis, the Contrastive Analysis Hypothesis, Flege's Speech Learning Model, Best's Perceptual Assimilation Model, and Bohn's Linguistic 'Desensitization Hypothesis.

### 1.1. THE EFFECT OF AGE: THE CRITICAL PERIOD HYPOTHESIS

It is widely believed that second/foreign language learning, especially in the area of phonology or speech learning, is strongly influenced by a learner's age. This effect of age is best expressed by the Critical Period Hypothesis (CPH). The concept of critical age in language learning was introduced by Lenneberg who stated that "foreign accent cannot be overcome easily after puberty" (Lenneberg, 1967, p. 176). According to the CPH, the sounds of an L2 cannot be learned perfectly

once a neurologically-based critical period has been passed (Lenneberg, 1967, Patkowski, 1990, Long, 1990).

The exact age of learning a second language at which a speaker's accent will be detected has been found to be inconsistent in previous research. For example, when Patkowski (1990) re-examined his 1980 study with 67 subjects, he found that the age of 15 was the turning point at which learners' accents became clearly detectable. In a study by Thompson (1991), Russian immigrants who had arrived in the U.S. before 10 years of age were perceived to have an accent. Flege (1988) found that native Mandarin speakers whose age of arrival in the U.S. was 7.6 were judged to be significantly less authentic (accented) in pronunciation than a group of native English speakers. Tähta, Wood, & Loewenthal found in their study (1981) that subjects who began learning English at the age of eight or earlier rarely produced a foreign accent. However, those who began learning English between the age of 9 and 11 were not accent free. In Long's (1990) summary of studies addressing the effect of age on second language phonological learning, he concluded that a native-like pronunciation is impossible for many individuals if their first exposure is not before the age of six and for the remainder by about the age of 12.

However, some researchers have been less quick to conclude that these findings are best captured in terms of a Critical Period. Flege argued that "although this may be true, such a conclusion fails to provide insight into how L2 learning differs from L1 acquisition, or what actually causes foreign accent". (Flege, 1995, p. 234). While all agree that age of learning is an important factor in foreign accent, some researchers believe that there is no clear cutting line or 'critical period'. In the study by Flege, Munro and Mackay (1995), the experimenters

assessed the degree of perceived foreign accent in a set of English sentences produced by native Italian (NI) subjects who differed in their age of learning English. These NI speakers, who had lived in Canada for over 30 years on average, began to learn English between the ages of 3 and 21. Native English listeners used a quasi-continuous scale to rate English sentences for degree of accent. The ratings revealed that there was a linear relation between the NI subjects' age of learning (AOL) and their degree of perceived foreign accent. In general, the later in life the NI subjects began to learn English, the stronger their English sentences were judged to be accented. However, there was no abrupt decline in perceived accent around a particular age group. In another study conducted by Munro, Flege and Mackay (1996), 11 English vowels were produced by the same Italian subjects. Their English vowel productions were rated for degree of foreign accent by 10 native English-speaking listeners. A parallel finding of a relationship between the age of arrival and the degree of accentedness was observed. The later the speakers began to learn English, the stronger their degree of accentedness was judged to be on almost every vowel. Another important discovery was that the AOL at which an accent first became detectable varied considerably across both listeners and vowels. (Munro et al., 1996). Both studies confirmed that there is no sharp "cutting line" due to age. "If a critical period exists, it apparently does not result in a sharp discontinuity in L2 pronunciation ability" (Flege, 1995, p. 234).

To summarize, researchers generally agree that age influence L2 speech learning. Yet, no consensus has been reached on whether the critical period is, in fact, "critical."

## **1.2. THE EFFECT OF LANGUAGE TRANSFER**

Another widely-recognized factor in the learning of a second language is the influence of the speakers' native language. Such an influence is especially noticeable in the area of phonetics and phonology. Nonetheless, there appears to be some disagreement on the degree and nature of this factor. Odlin (1989) pointed out that transfer can take place at both the phonemic and phonetic levels in L2 speech learning. Typology, or cross-linguistic frequencies of certain segments, common phonological rules, and syllable structure differences between the native and target languages are other factors that can affect the level of difficulty in learning the sounds of the L2. In this section, some different theoretical approaches towards language transfer will be discussed.

### **1.2.1. The Contrastive Analysis Hypothesis**

Perhaps the strongest claim about the influence of first language on the learning of a second language is the Contrastive Analysis Hypothesis (CAH). Lado, (1957) was the first to propose the CAH. He believed that learners tend to transfer both the forms and the meanings of their native language to the target language. Therefore, he considered it important to compare the structures of the native language and the target language in order to predict the learners' difficulties. The cross-linguistic differences in the CAH approach cover all aspects of language as well as culture. This discussion will focus only on segments. According to the CAH, the areas of difficulty in L2 phonemes are mainly caused by the differences between the phonological systems of the L1 and L2. For example, a segment is predicted to be difficult to learn if it

exists in the L2 but not in the L1. A learner will have trouble perceiving and producing the new segment. As a result, he will substitute some other phoneme from his native sound system. (Lado, 1957). This model accounts for some of the common problems with certain L2 learners. For example, the problem with Japanese speakers in learning the English segments /ɪ/ and /i/ is commonly believed to be caused by the lack of the /ɪ/ and /i/ contrast in the Japanese sound system. In fact, several studies have shown that Japanese speakers have difficulty perceiving this distinction (Sheldon & Strange, 1982, Flege, Naoyuki & Mann, 1995, among others). In his study, Cebrian (1996) found that the CAH was useful in isolating the potential possibility of transferring L1 phonological rules by Catalan speakers when learning English. However, the CAH has been criticized for not being able to account for the degree of difficulty corresponding to typological markedness in the prediction of problems that L2 learners face (Eckman, 1977, Odlin, 1989). It also fails to explain the individual differences among learners who have less difficulty with particular sound segments that are predicted to be a problem because of their absence in L1. For example, Munro, Flege, Mackay (1996) found that there were important individual differences in native Italian speakers' English vowel productions. The present study involving Mandarin speakers will further test the CAH by examining their success in learning to produce sounds that are missing from their L1 phonological inventory.

### **1.2.2. Perceptual Assimilation Hypothesis**

Direct transfer of L1 speech sounds into the L2 does not explain all the problems encountered in L2 speech learning (Odlin, 1989, Bohn, 1995). For instance, the effect of native language experience on the

perception of non-native language phones has been treated in terms of developmental changes (Werker & Polka, 1993, Best, 1994). In summarizing previous studies on the ability of infants and adults to discriminate non-native sounds, Werker and Polka stated that infants have perceptual abilities to discriminate phonetic contrasts on a broad, language-universal basis. However, language specific experience has proven to have an effect on the perception of non-native speech by the end of the first year in life. Adults generally have shown decreased ability to discriminate non-native phonetic contrasts in order to facilitate or enhance the perception of native phonetic contrasts (Werker & Polka, 1993).

One explanation of such an effect is that L2 learners or bilingual speakers tend to perceive the sounds of the L2 in terms of L1 phonology. (Best, 1994, Flege, 1995). For the adult listeners, the influence of the native phonological system causes perceptual assimilation of non-native phonemes to native phoneme categories apparently on the basis of articulatory similarities and discrepancies (Best, 1994). During the acquisition of the first language, speech perception becomes functional or 'attuned to' the contrastive phonemes of the L1 sounds, either because phonetically distinct sounds are "assimilated" by a single functional category, or because the L1 phonology filters out features (or properties) that are different phonetically but not phonologically. According to Best, the nature of the effect of language experience on the perception of non-native segments is 'an adjustment of selective attention'. It is important to point out that the effect of the native language on the perception of non-native segmental contrasts is neither 'absolute nor permanent' (Best, 1994). It is worth noting that Werker and Polka (1993) have also reported that

adults can discriminate some non-native phonetic distinctions without training.

Although the Perceptual Assimilation Hypothesis will not be directly tested in this study, it does provide some insight into the issues that will be reviewed here.

### **1.2.3. The Speech Learning Model**

Flege examined the influence of the L1 on L2 speech learning from a different perspective. The differences between native and non-native speech sounds are classified quantitatively by comparing the phonetic differences between related L1 and L2 segments. Over the years, Flege (1987, 1991a, 1995) has developed the Speech Learning Model (SLM) to account for the L1 factors he has noticed in his studies.

The SLM predicts that success in L2 production depends on the establishment of new phonetic categories for the segments that exist in the L2 but not in the L1. According to the model's first hypothesis (H1), "sounds in L1 and L2 are related perceptually to one another at a position-sensitive allophonic level rather than at a more abstract phonemic level" (Flege, 1995, p. 239). This suggests that L2 learners will relate an L2 sound to the nearest allophones of an L1 phoneme, not simply to a phoneme. The cross-language phonetic differences are not necessarily filtered out or assimilated to the L1 phonological system, but may be perceived by bilinguals.

If learners can perceive the phonetic differences between the L1 and L2 phones, it is possible that they will establish a new category for the L2 phone. The SLM's second hypothesis (H2) predicts that "a new phonetic category can be established for an L2 sound that differs phonetically from the closest L1 sound if bilinguals discern at least

some of the phonetic difference between the L1 and L2.” (Flege, 1995, p. 239). Eventually, learners will also be able to produce the sound when its phonetic category is established. As predicted by Flege in his Hypothesis 7 (H7) of SLM, “the production of a sound eventually corresponds to the properties represented in its perceptual phonetic category representation.” (Flege, 1995, p. 239).

The SLM also states that success or failure in establishing a new phonetic category depends largely on the phonetic distance between the L1 and L2 sounds. The chances of cross-language phonetic differences being distinguished increase with the degree of cross-language phonetic differences. In his earlier studies, Flege used less quantitative terms such as “new” and “similar” phones to describe the phonetic differences between L1 and L2 sounds. According to his earlier predictions, phones from the L2 that are “new” have no counterparts in the L1. “Similar” L2 phones, on the other hand, differ systematically from an easily identifiable counterpart in the L1. (Flege, 1987). For example, the French /y/ is a “new” phone for English learners of French while the French /t/ and /u/ are both considered to be “similar” phones despite the fact that they differ considerably from the English counterparts in terms of the VOT value for /t/ and the F2 value for /u/. According to this hypothesis, L2 learners will have difficulty in producing authentically L2 phones that differ acoustically from phones in L1 (similar phones) unless they establish a phonetic category for the L2 phones. The phenomenon referred to as ‘equivalence classification’ will block adult L2 learners from establishing a phonetic category for the “similar” but not for the “new” L2 phones.

In one of his studies, Flege (1987) tested groups of monolingual French and English speakers as well as native French and English

speakers who differed in experience in each others' language. The study examined the productions of the "similar" phones /t/ and /u/, and the "new" phone /y/ for the English subjects. The results showed that native English subjects who were experienced in French did not differ much from French monolinguals in producing the French /y/. However, the French and English subjects produced /u/ in their L2 with the F2 values differing significantly from those of the native speakers. They all produced /t/ in their L2 with mean VOT values that either closely resembled the L1's phonetic norms or fell between the phonetic norms of L1 and L2. Flege concluded that equivalence classification prevented experienced L2 learners from producing "similar" L2 phones, but not "new" L2 phones, authentically. (Flege, 1987). In other words, there was an upper limit on the extent to which L2 learners approximate the phonetic norms of the L2 for similar but not new phones.

However, a number of problems have arisen for this early version of the Speech Learning Model. One of the problems lies in the classification of "new" and "similar" phones when comparing native and non-native phones. Generally, there is a lack of a concrete standard for cross-language grouping. Blankenship (1991) questioned the standard for the term "similar" in her study of vowels. She noted the difficulty in determining how close together the vowels must be in order to be categorized as "similar". Also it was not clear whether the similarity between vowels should be characterized exclusively in spectral terms. Later, this "new" versus "similar" distinction of phones was modified to a more general term of cross-language phonetic differences. "The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that the phonetic difference between the sounds will be discerned." (Flege, 1995, p. 239). This

difference can be interpreted as more quantitative in nature. If related phones in the L1 and L2 are placed on a continuum, then, the greater or smaller distance in phonetic similarity will fall at different points along the scale rather than at the two extreme ends identified as "new" or "similar". This revised version of the SLM still needs further testing.

Another problem with the SLM is that it cannot account for the results generated in other empirical studies. If cross-language phonetic differences are really the basis for either the success or failure in speech perception and production in second language learning, then the "strange" new sounds in the L2 that have the greatest differences from L1 sounds should not be problematic for the learners. However, a more complicated phenomenon is found in the results from other empirical studies. For example, Munro, Flege and Mackay (1996) studied the English vowel productions of native Italian speakers grouped according to their age of learning (AOL). The subjects' productions of the vowels were presented to native speakers of English for both identification and accent ratings. The results indicated that the absence of a particular vowel from the L1 vowel system did not seem to be a critical factor for accent. For example, neither /æ/ nor /ɜ:/ occurs in Italian, yet these vowels were reported to be the second best and worst produced respectively. Vowel quality differences, in terms of simple acoustic (first two formant values) distance, between the English and Italian vowels did not seem to account for the results observed. Therefore, the hypothesis that English vowels which are acoustically "close" to Italian vowels should be less well produced than vowels that are more distant was not confirmed. Some aspects of the Speech Learning Model will be tested in the present study.

### 1.3. NON-NATIVE LANGUAGE TRANSFER FACTORS

This section will discuss only the phonologically relevant problems in L2 sound perception and production that are *not* directly caused by L1 transfer.

#### 1.3.1. Linguistic Desensitization Hypothesis

Flege and Bohn (1989) reported that native Spanish speakers used duration cues to identify the synthetic [i]-[ɪ] vowel continua that differed in linearly equal steps in formant frequencies and in duration. Spanish has no duration contrast and has only one vowel /i/ in the vowel space where English has two. Therefore, their use of the duration cues to distinguish the foreign vowel contrast cannot be due to a perceptual strategy transferred from the L1. The result seems to support the hypothesis proposed by Bohn and Flege (1990) which states that whenever L2 speakers cannot discriminate L2 vowel contrasts on the basis of spectral cues, they will rely on duration cues. This hypothesis was referred to as 'linguistic desensitization' or 'language-independent perceptual principle.' Bohn (1995) provided further evidence for this hypothesis in a later discussion that involved more L1 speaker groups.

In Bohn's 1995 study, a number of experiments were conducted to test native German, Spanish and Mandarin speakers' English vowel perception using synthesized tokens. The German speakers identified a *bet* to *bat* continuum while the Spanish and Mandarin speakers identified a *beat* to *bit* continuum. A group of native English speakers also participated as comparison group. The results showed that English listeners identified both *beat* to *bit* and *bet* to *bat* continua almost

exclusively on the basis of spectral cues. Differences in duration had little effect on the native English speakers' responses. The German, Spanish, and Mandarin listeners relied heavily or exclusively on the duration cue in identifying vowels from the English vowel pairs that do not exist in their first languages. (Both Spanish and Mandarin have only one category /i/ in the acoustic area where English has two, /i/ and /ɪ/. German has /ɛ/ in the area where English has /ɛ/ and /æ/.) Bohn stated that German listeners' use of duration cues in non-native vowel pair identification can be partly explained by the transferring of a native perceptual strategy. The German vowel system uses both spectral and duration to contrast vowel pairs. The Spanish and Mandarin listeners' reliance on duration cues cannot be attributed to the use of an L1 strategy for identifying non-native vowel pairs. Neither Spanish nor Mandarin uses duration cues to differentiate vowel contrasts in their vowel systems. Bohn concluded that "the use of the duration cue does not indicate reliance on a native perceptual strategy but reflects a general speech perception strategy that takes over whenever information conveyed by spectral differences is insufficient". (Bohn, 1995, p. 300).

Based on the data from these studies and some earlier ones, Bohn proposed the 'Desensitization Hypothesis' which states that "whenever spectral differences are insufficient to differentiate vowel contrasts because previous linguistic experience did not sensitize listeners to these spectral differences, duration differences will be used to differentiate the non-native vowel contrast." (Bohn, 1995, p. 294). In the above mentioned studies, both the native Spanish and native Mandarin speakers may be said to be 'linguistically desensitized' to spectral differences between /i/ and /ɪ/ due to the fact that both

languages have only one vowel category in the acoustic space where English has two. The native German speakers' desensitization to the spectral cue for /æ/ may be explained by the same reason. The results confirmed the fact that some of the learners' problems in perception are not necessarily caused by transfer but by the lack of sensitization towards the L2 vowels. The present study will further test this hypothesis in both the production and perception.

### 1.3.2. Second Language Experience

The amount of L2 experience is another relevant factor in L2 speech learning. Early stage learners may differ from experienced learners in both their perception and production. Bohn and Flege (1990, 1992) found that the more experienced group of native German English learners, not the less experienced, were more successful in the identifying and producing the English /æ/. However, L2 experience showed no effect on the perception and production of "similar" vowels /i/-/ɪ/. In other words, the results suggest that L2 experience affected some vowels but not all.

In another study by Flege (1991b), groups of native Spanish speakers, experienced and inexperienced English L2 learners, and monolingual Spanish speakers identified English front vowels using Spanish vowel categories. L2 experience was found to have an effect on the awareness of /ɪ/ (Spanish has only /i/ while English has /i/ and /ɪ/ in this area of vowel space). However, it had no effect on the perception of the English vowel /æ/ for any of the groups. Experienced and inexperienced English L2 learners, as well as the monolingual Spanish speakers identified English /æ/ with Spanish /a/ most of the time. (Spanish does not have the /æ/ category).

In a more recent study on adult Japanese learners' productions of English /ɪ/ and /I/ conducted by Flege, Naoyuki & Mann (1995), it was found that experienced Japanese learners who had lived in the United States for an average of 21 years produced /ɪ/ and /I/ tokens that were as intelligible as those produced by native English speakers. The majority of the experienced adult Japanese learners' productions of /ɪ/ and /I/ also received ratings that fell within the range of the native English speakers' productions. In contrast, the productions of English /ɪ/ and /I/ by Japanese speakers who lived in the United States for an average of 2 years were significantly less intelligible. In addition, their production was often judged to be strongly accented. The findings suggest that L2 experience and length of residence in the L1 environment had an effect on the success of learning the new L2 consonant contrast that does not exist in the L1. The significance of this study is the extreme difference in length of residence of the two Japanese groups (19 years apart). It seems that many years of exposure to the L1 environment really helped the Japanese speakers to produce the new consonant distinction that is not existent in their L1.

#### **1.4. THE RELATIONSHIP BETWEEN PERCEPTION AND PRODUCTION**

The effective learning of an L2 segment involves both perception and production. It is often assumed that production has a perceptual basis and that good perception generally precedes good production. In first language acquisition, young children all go through the stage of perceiving the sounds of their L1 before they learn to produce them. In a foreign language classroom, it is not uncommon for teachers to

practise the discrimination of certain L2 sounds before they actually let the students try to produce them. Furthermore, even such common terms as 'imitating' and 'mimicking' all imply a perceptual basis for production.

Previous studies concerning L2 segmental perception have very seldom involved parallel production studies. In theory, Flege's Speech Learning Model strongly implies that production has a perceptual basis. According to the SLM, an L2 speaker's success in perceiving at least some differences between the L1 and L2 sounds is considered to be the basis for the establishment of new phonetic categories for the L2 sounds. This success in perception will presumably lead to successful production of the L2 sounds but not vice versa. If a bilingual speaker is unable to discriminate an L2 vowel from either the neighboring L2 vowels, or from the L1 vowels that are different from the L2 vowel phonetically, the L2 vowel will be produced inaccurately (Flege, 1995). Therefore, L2 vowel production errors arise from the failure to discriminate certain L2 vowels. Flege (1995) reported that native Korean subjects failed to discriminate the English front vowel pairs /i/-/ɪ/ and /ɛ/-/æ/ in a perception test. The Korean subjects also had problems in producing these vowel pairs. The results showed that their productions of the intended /i/ and /ɪ/ tokens (in bVt context) were often misidentified as each other, as were their /ɛ/ and /æ/ tokens.

Some counter-evidence to the proposal that perception precedes production was found in a study by Sheldon & Strange (1982), who examined the production and perception of the English /ɪ/ and /i/ by native Japanese speakers. They found that Japanese learners were somewhat more successful in producing a contrast between /ɪ/ and /i/ than in perceiving the distinction between the two sounds. They

concluded that perceptual mastery of an L2 contrast does not always precede the learners' ability to produce 'acceptable tokens' of contrasting phonemes. Blankenship (1991) also reported in her study that although Spanish bilinguals were not always able to perceive differences in L2 vowels, they were able to produce all the necessary L2 vowel contrasts in both a reading task and in an interview. These findings were in contrast to the common belief that perception precedes production. The present study will test whether production has a perceptual basis using natural vowel tokens.

### 1.5. MANDARIN CHINESE VOWELS

Previous studies have indicated that for both theoretical and empirical reasons, the relationship between the sounds of different languages cannot simply be evaluated by comparing the phoneme inventories (Bohn & Flege, 1992). Since the sounds in two languages often show different acoustic and articulatory characteristics, a phonetic description is necessary (Odlin, 1989). Mandarin Chinese, a term familiar to the western world, is known as Standard Chinese, or *Putonghua* in Mainland China. The dialectal norm for Mandarin is the speech of the capital city, Beijing. Mandarin Chinese has five vowel phonemes<sup>1</sup> (Cheng, 1966, Svantesson, 1984). They are:

i   y                      u  
                                    e  
                                    a

---

<sup>1</sup> The mid vowel [ɤ] with the retroflex coloring, or "er-hua" final, is used as a morphological ending which is heard only in the northern areas. Mandarin speakers in other areas usually are unable to produce it and, therefore, tend to avoid it. (Norman, 1988). It is not under consideration in the present study.

In spite of the fact that the vowel form [o] occurs in the Pinyin system as in [ou], it is treated as an allophone of /ɛ/ and occurs only in a specific environment (Cheng, 1966). In fact, the mid vowel /e/ has many “surface forms” that cover a range from the front to the back. In order to identify the surface vowels in Mandarin, the allophones of the five phonemes are given below:

i	y	ɨ	ɯ	u
e		ə	ɤ	o
ɛ				ɔ
æ		a		ɑ

The high front, unrounded vowel /i/ has three surface forms:

- [ɯ] (high back, unrounded) occurs only after dental sibilants [ts] [ts'] and [s] such as in [sɯ] “silk”.
- [ɨ] (high, central unrounded) occurs only after retroflex fricatives and affricates [tʂ], [tʂ'], [ʂ], and [ʐ], e.g., [ʂɨ] “poem”.
- [i] occurs elsewhere, e.g., [di] “drop”

The first two allophones of /i/ are sometimes referred to as apical vowels (Dow, 1972, Karlgren, cited from Chao, 1968 & from Howie, 1976, Shih, 1995).

The mid vowel /e/ has the most allophones. They cover a wide range and have different descriptions. The following is a summary of the most common descriptions of the surface forms of /e/ and their environments (Howie, 1976, Chao, 1968, Wu, 1994).

- [e] occurs before [i] as in [pei] “cup”.
- [ɛ] occurs after glide or medial [i] and [y], e.g., [iɛ] “leaf” and [yɛ] “moon”.
- [o] occurs before [u] as in /gou/[gou] “hook”.

---[ɤ] sometimes described as [ɤʌ] occurs in an open syllable as a monophthong as in [gɤ] “song”.

---[ɔ] occurs after the labial consonants in open syllables, e.g., [pɔ] “wave” and after [u] as in [ʃuɔ] “say”.

---[ə] occurs in closed syllables before nasals as in [gən] “root”.

The low vowel phoneme /a/ also has several allophones. The exact number of surface forms for this vowel, however, is controversial and different descriptions are not uncommon. A summary of the common descriptions of the allophones of /a/ and their environments are listed below:

---[ɑ] occurs before [u] and [ŋ] as in [gɑu] “tall”.

---[æ] occurs after [i] and before [n] as in [tiæn] “sky”.

---[a] occurs before all other occurrences of [n], as in [pan] “move” and before [i] as in [gai] “should”.

---[A] occurs elsewhere in an open syllable as in [mA] “mother” and [iA] “duck”.

According to previous reports (Dow, 1972, Howie, 1976, Chao, 1968, Wu, 1994), the Mandarin /a/, has a medium quality between the front [a] and the back [ɑ]. It can also be represented by Chao's low, central [A] (Chao, 1968), which is not part of the IPA. It is [A] in syllable final position as in [iA] and [uA]. However, it is more posterior before the back vowel [u] and is represented as /ɑu/ [ɑu]. Some linguists disregard [A] as an allophone and consider it to be [a] as in [tɕa] (Wu, 1994), while others treat it as [ɑ] as in [xɑ] (Howie, 1976). According to some authors, /a/ has another allophone [æ] which occurs before [n] and after [i] as in [iæn] (Howie, 1976, Wu, 1994, Dow, 1972). There are different explanations and interpretations for this allophone in this

particular environment. Some consider it as [a] as in /iai/ and others as [ɛ] as in /ian/, realized as [iɛn], (Chao, 1968, Li & Thompson, 1981).

The descriptions of the surface vowel forms show that most of the allophones of the five vowel phonemes occur only in diphthongs and closed syllables. Compared with the rich, stressed monophthongs in the English vowel system, Mandarin has only /i/ and /u/ with counterparts in English. Mandarin /a/ does not seem to have an obvious English counterpart when it occurs as a monophthong. Mandarin /e/ is realized as [ɛ] when it occurs in an open syllable as a monophthong. The Mandarin diphthongs [ei] and [ou] are comparable to the English counterparts, the so-called phonetic diphthongs [ei] and [ou]. Other two diphthongs, /ai/ and /au/ are also comparable to the English counterparts [ai] and [au] (sometimes transcribed as [aj] and [aw]). The latter two vowels are excluded from the present study because of the complication of Canadian Raising which changes the quality of these two English diphthongs.

Since cross-language acoustic comparisons of English and Mandarin vowels have been rare, it is difficult to have a comprehensive review of the acoustic comparisons of the vowel qualities. The Canadian English low back vowel [ɒ]<sup>2</sup> is more posterior than the central Mandarin [A] as the phonetic symbols suggest. Compared to the English [u], the Mandarin [u] is both higher and more posterior (Nørmann, 1988). The average F2 value for Mandarin [u] was 620 Hz as measured in a study of one male speaker by Howie (1976), and 761 Hz by Svantesson (1984). The F2 value of Canadian English [u] was measured as 1174 Hz, an average of 10 speakers (five male and five female) in the study by

---

<sup>2</sup> The symbol [ɒ] is used in this study to represent a phonemic merger of /ɑ-ɔ/ in Canadian English (Avis, 1975, cited in Assmann et al, 1982).

Nearey and Assmann (1986). It is important to point out that although [i ε æ u] all occur as allophones of /i e a u/ in diphthongs, they never occur as stressed, syllabic monophthongs. Dow (1972) pointed out that in the falling diphthongs, the second elements [i] and [u] only indicate the directions of the movement, not the maximum points reached. The lax vowel symbols [ɪ] and [ʊ] are used to indicate the movement. Compared with the English lax vowels, which all occur as stressed monophthongs, Mandarin lacks such counterparts in its vowel system. Strictly speaking, the Mandarin surface vowel forms [ɪ] and [ʊ] are the glides [j] and [w] and only occur as on-glide or off-glide. The controversial [æ], as stated previously, occurs only before the nasal [n] and after the glide like [i] in Mandarin. The English lax vowels [ɪ ε æ ʊ] may, therefore, be treated as "new" phones for Mandarin speakers while /i eɪ a u ou/ are "similar" phones that have English counterparts.

The above description of Mandarin vowel phonemes and allophones will provide the basis for the comparisons of English and Mandarin vowels that is the part of this study. Previous studies have reached the agreement that the cross-linguistic comparisons of phones should be carried out at both phonemic and phonetic levels (Bohn & Flege, 1992, Odlin, 1989).

In the present study, the acoustic properties of isolated Mandarin vowels [i u ou eɪ a] produced in the falling tone (see Chapter 2 and Chapter 4 for description) by native Mandarin speakers will be measured and compared to the English counterparts in terms of duration and the first two. The acoustic measurements of these vowels will be carried out with the intention of testing the phonetic differences between these vowel counterparts across languages. Mandarin vowels that have no corresponding English categories will not be considered.

The English lax vowels [ɪ ɛ æ ʊ ʌ] which have no Mandarin counterparts will be studied acoustically and serve as the basis for the intelligibility test.

The purpose of cross-language acoustic analysis of the vowels is to evaluate the phonetic basis for the Mandarin speakers' L2 vowel perception and production study rather than to depict the whole vowel system of both languages or to classify "new" and "similar" vowel categories. Instead of a broad classification of the vowel inventories of the Mandarin, the study will focus more on the acoustic or physical characteristics of the vowels.

## **1.6. RESEARCH TOPICS**

The purpose of this study is to explore the problems that native Mandarin speakers have with the perception and production of English vowels. The main focus is to discover which English vowels are perceived and produced well by native Mandarin speakers and which ones are not. It will be considered whether success or failure in the production of English vowels is explicable in terms of sound correspondences between the native language (Mandarin) and the L2 (English). With regard to foreign language speech learning theory, this study will test the Contrastive Analysis Hypothesis which predicts specific problems in L2 speech learning by comparing the inventories of the L1 and L2. Flege's Speech Learning Model will also be examined. Furthermore, an attempt will be made to explore the individual speaker differences in L2 vowel perception and production, a problem that is important but has received relatively little attention in previous studies.

Previous studies including those summarized above have all contributed to research in the area of foreign language speech learning. However, very limited studies have been conducted to investigate Mandarin speakers' problems while learning to produce English vowels. Among those studies that did include native Mandarin speakers, such as Bohn (1995), only a small number of vowels were actually tested in the perceptual experiment.

A careful examination of previous studies also reveals complexities in the methodology of studying L2 vowels. One problem is that vowels have often been identified or rated for accent in the context of the surrounding consonants, which could have an effect on the listeners' judgment of the vowels. However, the use of isolated vowels in an intelligibility test is not ideal either because in real communication, vowels are rarely produced or heard in isolation. A second problem is the lack of a parallel production test in the perception studies. This limitation has made it difficult to test the relationship between perception and production.

Another common problem with previous cross-linguistic studies is the lack of dialectal control over the subjects. Dialectal differences may have an effect on both the subjects' perception and production. The resulting individual differences in performance may make data interpretation difficult. For example, in the study of Arabic speakers' accented English vowels, Munro had to take into consideration certain Arabic dialectal differences in long and short vowel contrast while analyzing the duration effect on English vowel productions (Munro, 1993). The present study is designed to avoid the above mentioned problems. First, the L1 speakers' dialectal differences are controlled to a very strict region. All the subjects came from Beijing and spoke the

same dialect. Second, the production data collected for the acoustic measurement and the identification test are isolated natural vowels. This limits the influence of the consonant context which could have an effect when studying foreign accents. In addition, an attempt is made to explore the relationship between perception and production by comparing perception and production data for the front vowels.

### **1.6.1. Outline of the Present Study**

Three experiments will be conducted in this study. Experiment 1 is designed to evaluate the intelligibility of a group of Mandarin speakers' English vowel productions. Ten English vowels /i r eɪ ε æ u ʊ ou ɒ ʌ/ produced by both the Mandarin speakers and a comparison group of native English speakers will be submitted to an intelligibility test performed by another group of native English listeners. The purpose is to discover which vowels of the 10 produced by the Mandarin speakers are better identified than the others and which ones are poorly identified. Experiment 2 will test native Mandarin speakers' perceptions of native English speakers' productions of five English front vowels. An attempt will be made to examine how perception relates to production by comparing the results of the two experiments. Experiment 3 will examine and compare the acoustic properties (in terms of the first two formants and duration) of the relevant Mandarin vowels, English speakers' productions of English vowels, and Mandarin speakers' productions of English vowels. The results of the three experiments will be compared and discussed.

## CHAPTER II

### EXPERIMENT 1: INTELLIGIBILITY TEST ON MANDARIN SPEAKERS' ENGLISH VOWEL PRODUCTION

In this experiment, 15 native Mandarin and 15 native English speakers' productions of 10 English vowels [i ɪ eɪ ε æ u ʊ ɔʊ ɒ ʌ] will be identified by four native English listeners. The purpose of this experiment is to evaluate the success of the native Mandarin speakers in producing these vowels. The experiment is designed to address the following issues. First, it will indicate how intelligible the Mandarin accented vowels are to the native English listeners in comparison with the native English speakers' productions of the same 10 vowels. Second, it will establish whether some vowels produced by the Mandarin speakers are more intelligible than others, and if so, which ones. A third issue to be explored is whether any differences in intelligibility are explicable in terms of differences between the vowel systems of the L1 and L2. Furthermore, this experiment will investigate the phenomenon of individual speaker differences in L2 speech learning.

Of the 10 English vowels under consideration, five of them, [ɪ ε æ ʊ ʌ], do not have obvious Mandarin analogs. The other five, [i eɪ u ɔʊ ɒ] have counterparts in Mandarin<sup>1</sup>, at least according to commonly-used transcriptions. It is predicted that Mandarin speakers' productions of the vowels [i eɪ u ɔʊ ɒ] will be more intelligible than their productions of [ɪ ε æ ʊ ʌ]. This is because, according to the Contrastive Analysis

---

<sup>1</sup> The Mandarin /a/ does not have an obvious English counterpart when it is a monophthong. (See Chapter 1 for more description.) The phonetic differences between Canadian English /ɒ/ and Mandarin /a/ will be examined in Chapter 4.

Hypothesis, segments that do not have L1 counterparts will be difficult for L2 speakers to perceive and produce. Learners are expected to substitute their L1 vowels for the closest L2 vowels in both perception and production (Lado, 1957). The results of this experiment will further test the predictions of the Contrastive Analysis Hypothesis.

## **2.1. METHOD**

### **2.1.1. Participants**

Fifteen Mandarin speakers were recruited as paid subjects (seven males and eight females). They all came from Beijing and spoke Standard Mandarin Chinese as verified in an interview with the experimenter. Five of them were born in other cities in China but moved to live in Beijing in early childhood. The rest were born and raised in Beijing. They had a mean age of 32.3 years at the time of the experiment with a range of 26 to 45 years. The Mandarin speakers had been living in Canada for an average of 4 years with a range of 0.5 to 6 years. Their mean age of arrival in Canada was 28.5 years with a range of 22 to 40 years. Most of them were graduate students at Simon Fraser University and the University of British Columbia, and all had high proficiency in English. According to self-report, all the participants had learned English in China before they came to Canada. Age of learning English as a foreign language (age at which English instruction began in China) varied from 6 to 16 years with a mean of 11 years. On average their daily use of English, according to self-report, was 5.3 (range: 3-8) on a scale of 1 to 9, with 1 representing "never" and 9 indicating "only" English. Background information for the individual Beijing Mandarin subjects is given in Table 2-1.

Fifteen native speakers of English also participated as a comparison group. There were equal numbers of male and female speakers (seven males and eight females) in the two groups. The native English participants were all speakers of Canadian English who were born in and grew up in Western Canada (from British Columbia to Ontario). They had a mean age of 30.4 years at the time of the experiment. Most of the native English speakers were graduate or undergraduate students of the Linguistics Department at Simon Fraser University. Other subjects either worked in the same department or were students at the same university. Participants in both groups had normal hearing according to self-report.

### **2.1.2. Recordings**

Individual recordings were made with a Genexxa 33-984D microphone connected to a JVC TD-W709 cassette recorder in a sound-treated room. The subjects first completed a Language Background Form (see Appendix 3) giving general information about their language background. The experimenter then explained the reading tasks to clarify any possible problems concerning the reading lists to the subject. In each case, the subject was told to prepare the reading list by reading silently once before recording to reduce the likelihood of reading errors or hesitations. During the recording phase, if a target word was mispronounced accidentally, the subject was asked to repeat the word with the entire sentence.

The native English participants read only the English list which consisted of 13 English vowels in /bVt/ or /bVk/ words, /bV/ combinations and isolated /V/ production. (See Appendix 1, English

list). They were words or syllables containing the 13 English vowels: /i/, /ɪ/, /eɪ/, /e/, /æ/, /u/, /ʊ/, /oʊ/, /ɒ/, /ɜ/, /ʌ/, /aɪ/, and /ɑu/. The words were written in standard orthography: beat, bit, bait, bet, bat, boot, buwt (same vowel as in 'book'), boat, bot, Bert, but, bite, and bout. In instances where the desired tokens were not real words, for example, 'bot', a "sounds-like" example 'pot' was given just beside the target "word". The real English word 'bought' was avoided for the intended vowel /ɒ/ because English learners are sometimes confused by the many possible pronunciations of "ough" in English (e.g., though, through, cough, rough). The only word that ends with a /k/ instead of /t/, 'book', was given as a special precaution to elicit the correct vowel /ʊ/ just in case some speakers could not produce the 'word' buwt. In order to keep the target vowel in their productions, the speakers were instructed to read the /bVt/ word that contained the target vowel first followed by the /bV/ and then by the isolated /V/. Both the /bVt/ words and /bV/ combinations, as well as the isolated /V/ were read in the carrier sentence "Now I say \_\_\_". The speakers were told to read the list at normal speed.

The Mandarin subjects were recorded reading the English list as well as a Chinese list that was used in Experiment 3. The Chinese list consisted of 37 characters representing the 37 Mandarin Chinese rhymes with an initial consonant, and thirty-two characters of the same rhymes without initials. (See Appendix 2, Chinese list). The characters were read in the sentence frame "Wuo shuo \_\_\_" /wuoʃuo/ (I say \_\_\_). All the characters chosen for the target vowels in the list were in tone four (falling tone) to avoid any pitch variation that might make it difficult to compare across vowels. The falling tone was chosen also

because it resembled the English falling intonation in regular speech. It has been reported in previous studies that both tone and position of a word in a sentence frame can have an effect on the duration of syllable nuclei (Ho, 1976). Ho found that of the four tones in Mandarin Chinese, tone four of an isolated word had the shortest duration. Tone four remained the shortest in duration when occurring sentence finally also. These observations will be taken into account in the analysis reported in Chapter 4.

### **2.1.3. Stimulus Preparation**

The isolated English and Mandarin vowels were carefully separated from the carrier sentence using a waveform editing program. They were digitized at 22.05 kHz with 16-bit resolution and saved in System 7 sound format on an Apple Macintosh computer. Only 10 of the 13 English vowels produced by both the native English and Mandarin speakers were used in this experiment. They were the five front vowels [i i eɪ ε æ] and five back vowels [u u oʊ ɒ ʌ]. Isolated vowels were chosen for the test because they were to be measured and compared to Mandarin vowel counterparts in Experiment 3. Because Mandarin phonotactics permits as codas only /n/ and /ŋ/ and a morpheme suffix -er, and because no one consonant can be found as an onset accompanying all Chinese vowels, isolated vowels of both languages would make the results of the acoustic measurement more comparable cross-linguistically. This issue will be addressed further in the following chapter.

#### **2.1.4. Listeners**

The listeners were four native speakers of Canadian English (one male and three females), one of whom was a trained phonetician (the male listener). The other three were all phonetically trained and were familiar with the IPA symbols. They were graduate or undergraduate students in the Linguistics Department of Simon Fraser University and had completed an advanced level phonetics course. The listeners ranged in age from 24 to 39 years at the time of the experiment with a mean of 28 years, and all had normal hearing according to self-report. Two of them were from British Columbia, one was from Ontario, and one was from Alberta.

#### **2.1.5. Listening Test**

A total of 1500 vowel tokens (10 vowels x 30 speakers x 5 randomizations) were used for the intelligibility test. They were blocked in separate sets of front and back vowels of 150 each for each listening session. Individual listening sessions were held in a sound-treated room. Each listener completed 10 self-paced sessions in approximately eight minutes per session. In most cases, the front and back vowel stimuli (five sessions each) were identified in alternate sessions. The listeners performed the task on a computer using speech playback software (Munro, 1996). The stimuli were all presented at a comfortable listening level through a speaker connected to the computer in a sound-treated room. The listeners were permitted to hear each token only once.

During the front vowel sessions, seven phonetic symbols [i, ɪ, e, ε, æ, ɒ, ʌ] were attached to the relevant keys on the keyboard. The two back

vowels [ɒ ʌ] were also included among the labels for the front vowel task because they provided possible choices for some potentially ambiguous front vowel tokens. During the back vowel sessions, eight phonetic symbols [u ʊ ou ɒ ʌ æ ε ɪ] were used. The three front vowel symbols [æ ε ɪ] were included for the back vowel sessions for similar reasons. Each listener was instructed to press the key labeled with the symbol corresponding to the vowel heard. Different stimulus randomizations were used for each session and for each listener. Before the listening task, the experimenter explained the identification task to the listener carefully to make sure that he or she understood the task thoroughly. The identification data were collected automatically by the computer and saved for subsequent analysis.

## 2.2. RESULTS

Figure 2-1 shows the mean percentage of correct identifications (pooled over four listeners) of the 10 vowel productions by 15 native English and 15 native Mandarin speakers. Overall, it can be readily seen that the native English speakers' vowels were much more intelligible than those of the Mandarin speakers. On average, the Mandarin speakers' English vowels had a correct identification rate of 69% as compared to 93% for those of native English speakers.

A two-way mixed design ANOVA was carried out on the listeners' responses with first language (native English vs. native Mandarin) as a between group factor and vowels (10 levels) as a within group factor. The effects of language [ $F(1, 28) = 60.2, p < 0.0001$ ] and of vowel [ $F(1, 28) = 21.9, p < 0.0001$ ] were significant, as was the interaction of these

factors [ $F(9, 252) = 14.3, p < 0.0001$ ]. Tests of simple main effects revealed ( $p < 0.005$ ) that the Mandarin speakers' productions of [ɪ ɛ æ ʊ ɒ] were significantly less well identified than the corresponding vowel productions by the native English speakers. However, the Mandarin speakers' productions of [i eɪ u ou ʌ] were about as intelligible as those of the native speakers. Figure 2-1 depicts an interesting pattern in the vowels that were produced with high intelligibility to native English listeners. In general, the Mandarin speakers' productions of English vowels that have Mandarin counterparts were significantly more intelligible than those that had no counterparts. However, the vowel [ʌ] was found to be inconsistent with this pattern. The correct identification rate for this vowel was 76% for the native Mandarin speakers and 78% for the native English speakers. No obvious explanation could be found for this exception at this stage. It is important to note that [ʌ] was the second least well identified vowel for the native English speakers. It was the most well identified vowel category without an obvious Mandarin counterpart for the native Mandarin speakers.

Test of simple main effects also revealed that the effect of vowel was significant in both speaker groups. To explore the vowel identification data more carefully, two additional ANOVAs were carried out. First, a one-way repeated measures ANOVA on the vowel identification data for the Mandarin speakers only was carried out to further explore the intelligibility across the 10 vowels within the Mandarin speaker group. A significant effect of vowel was observed [ $F(9, 126) = 20.87, p < 0.0001$ ]. A post hoc Tukey (HSD) test revealed that [ɪ ɛ æ ʊ ɒ] were all significantly less well identified than [i eɪ u ou ʌ], ( $p < 0.05$ ) Although there were some differences in the average correct

identification scores for [ɪ ɛ æ ɒ], the differences proved to be non-significant as the pairwise comparisons indicated. Similarly, the differences in the correct rate of identifications for the [i eɪ u ou ʌ] were not significant either. The pattern again was very clear for vowel categories with and without Mandarin counterparts within the Mandarin speaker group. In general, the native Mandarin speakers' productions of English vowels without Mandarin counterparts were less intelligible to native English listeners than the vowels with Mandarin counterparts. Two important points about the generally less well identified vowel categories need to be examined in detail. First, the Mandarin speakers' [ʊ] was significantly less well identified than all nine other vowels. This distinction was worth noting because unlike the majority of the poorly identified vowel categories [ɪ ɛ æ ɒ] which were not significantly less or better identified than one another as the statistic test showed, [ʊ] was significantly less intelligible than [ɪ ɛ æ ɒ]. The other important discovery was that the identification score for [ʌ], which does not have an obvious Mandarin counterpart, actually put it in the group of better identified vowels which have Mandarin counterparts. It was the only lax vowel that was almost as intelligible as the native English speakers' productions.

To explore further the misidentified Mandarin speakers' vowels, a confusion matrix was created (averaged over the four listeners) and is presented in Table 2-2. The data indicated that the majority of misidentified [ɪ]s were heard as [i]s (37%), (the % correct ID for [ɪ] was only 46%). A considerable number of [ɪ]s were also heard as [eɪ]s (11%). For the poorly identified [ʊ], which had a correct rate of identification of only 15%, the most common misidentification was [u] (47%). A

considerable number of [ʊ]s were also heard as [ou]s (34%). The confusion patterns for these two vowels suggested that the nearest tense vowels of the L1 were substituted by the native Mandarin speakers for the L2 categories that they had not yet learned to produce successfully. The same pattern also applied to [ɒ], whose misidentified tokens were heard as [ou]s most of the time (28%). However, this substitution of the nearest vowel with L1 counterparts could not explain the Mandarin speakers' performance on the other poorly identified vowels. For example, for the misidentified [ɛ]s, more [æ]s were heard (19%) than [eɪ]s (13%). Furthermore, the incorrectly identified [æ]s were almost all heard as [ɛ]s (38%). Only 1% of [æ]s were heard as [eɪ]s. The confusion data for [ɛ] and [æ] indicated that to the native ear, the Mandarin speakers' [ɛ] and [æ] were mostly confused with each other, rather with than [eɪ], the single category that exists in that area of vowel space in their L1 system. The majority of misidentified [ʌ]s were heard as [ɒ]s. The production confusion data will be explored further and compared with perception confusion data in Experiment 2.

In order to explore differences among listeners in the listening task, a two way mixed design ANOVA was carried out with listeners as a between group factor and speaker groups (two levels, native English and native Mandarin) as a within group factor. The effect of listener [ $F(3, 36) = 0.335, p > 0.05$ ] and the interaction between listener and speaker [ $F(3, 36) = 0.059, p > 0.05$ ] were non-significant indicating that the four listeners were fairly consistent in identifying the vowel productions of both groups. The summary of percentage of correct identifications of each vowel by each listener and standard deviations are given in Table 2-3.

### 2.2.1. Individual Speaker Differences

The Mandarin speakers' production data that have been discussed so far were all based on average identifications of 15 speakers by four listeners. In order to explore individual speaker differences, the correct percentage of identifications of 10 vowels with standard deviations for every Mandarin speaker are summarized in Table 2-4. The data suggested that there was very little individual variation in Mandarin speakers' productions of [i eɪ u ou], vowels that have Mandarin counterparts. Of the 15 Mandarin participants, 14 speakers' [i]s, 11 speakers [eɪ]s, and nine speakers [u]s and [ou]s were 100% correctly identified by the four listeners. The Mandarin speakers' productions of [ʊ], which was least well identified among the 10 vowels, were also found to be fairly consistent. For 14 of the 15 speakers, [ʊ] had a correct identification rate below 30%, including six speakers whose productions were totally unintelligible (0%). It is interesting to note, however, that one speaker produced this vowel with a high degree of intelligibility (90%). Considerable inter-speaker differences were found for [ɪ] as it covered a range from 0% to 90%. Three out of 15 speakers produced completely unintelligible tokens (0%); none of them produced a perfectly intelligible [ɪ]. To have a clear view of each speaker's production of these vowels, the correct percentages of identifications of [i eɪ u ʊ ou] were plotted in Figure 2-2.

There were obvious individual differences in the Mandarin speakers' productions of [æ]. Three of the 15 speakers' [æ]s were 100% correctly identified by all the listeners while six speaker's productions of the same vowel had correct identification scores below 15%. Similar but less sharp individual differences were also found for [ɛ]. Six out of

15 speakers' [ɛ]s had a correct identification rate of above 90% (three were 100%), while five of them had correct identification scores that were below 30%. The individual data for these two vowels indicated that some speakers had perfectly intelligible productions of these two vowels (100%) and others had completely unintelligible productions (0%).

Considerable speaker variation in vowel intelligibility was also reflected in each individual speaker's total score on the 10 vowels. First, as indicated in Table 2-4, the best speaker had an average score of 88% as opposed to 51% by the worst speaker. A careful examination of the score on each vowel revealed that the best speaker produced correctly identified vowels 85% of the time or better with the exception of [u], which was only 25% intelligible. The intelligibility of her production of [u] was considerably lower than that of one speaker (90%). This low score on [u] indicated that even the "best" speaker did not master all 10 vowels. The one speaker whose [u] was identified 90% correctly was found to be the second best with an overall intelligibility score of 83%. Five of her 10 vowels were produced perfectly in terms of intelligibility (100%). Her success on [u] suggested that she was the only person who mastered this vowel. However, her production scores for both [ɒ] (35%) and [ɪ] (40%) were lower than the average of the whole group (49% and 46% respectively) indicating that she still had problems with these two vowels. On the other hand, the least successful speaker's productions of four vowels [eɪ u oʊ ʌ] were correctly identified 80% or more of the time, indicating that she did not have serious problems in producing these L2 vowels in terms of intelligibility. Her productions of the other four vowels [ɪ ɛ ʊ ɒ] were correctly identified 25% of the time or less. The

speaker variation observed in the Mandarin speakers' productions of 10 English vowels showed that not a single speaker was completely successful or unsuccessful in producing all the 10 vowels being studied.

To explore whether there was an effect of L2 experience in these speakers' L2 vowel productions, the speakers' language background data collected in the questionnaires were examined. Factors such as the number of years in Canada, the age at which learning English as a foreign language began, the number of years of learning English at school, and daily use of English may affect their success in the production of L2 phones.

As the exploration of the effect of L2 experience on speakers' success in production was not the focus of this study, the selection of the participants was not particularly controlled on L2 experience.<sup>2</sup> Nevertheless, it is still worth the exploration of such an effect in terms of above mentioned factors with the limited data available.

The speaker whose vowel production score was the highest (88%) among the 15 was found to have the shortest length of residence in Canada (only six months), considerably below the mean of three years. In contrast, the speaker whose overall intelligibility score was the lowest (51%) had seven years of residence in Canada, the longest among the 15 speakers. Furthermore, according to self-report, the "best" speaker's daily use of English was less (4) than that of the least successful speaker (8). (The scale was from 1 to 9, with 1 representing 'never' and 9 representing 'only'.) However, the age of learning

---

<sup>2</sup> The controlled data concerning L2 experience would involve at least two groups of L2 speakers (usually with the same L1 background) with considerable difference in the length of residence in the L2 environment. (See p. 13 for more descriptions).

(learning English as a foreign language in China, see also "English instruction began" at p. 25) for the best speaker was nine years, for the least successful speaker, 16 years. The length of time studying English was 19 years for the best speaker but three for the least successful speaker. Although age of learning and years spent studying English were both in favor of the best speaker, the length of residence in the English speaking environment and the daily use of English were apparently not important factors in this case.

As discussed earlier, the data concerning L2 experience and length of residence were not systematically controlled for the purpose of examining the effect of these factors, no firm conclusion can be drawn from the findings here. More controlled studies are needed for the exploration of L2 experience as an effect on L2 vowel production.

### **2.2.2. Native English Vowel Differences**

A one-way repeated measures ANOVA on the vowel identification data for the native English speakers only was also carried out to explore the intelligibility across the 10 vowels. A significant effect was observed [ $F(9, 126) = 8.69, p < 0.0001$ ]. A post hoc Tukey (HSD) pairwise comparison test revealed that although there were some significant differences among vowels, the differences were much less dramatic than those seen in the productions of the Mandarin speakers. There was not a single vowel that was significantly less well identified than all the other nine vowels. In fact, only two vowels, [ɒ] and [ʌ], had significantly lower identification scores than some of the other vowels. No significant difference in intelligibility was found between these two vowels. The individual speaker differences for the native English speakers were

much smaller than those observed in the Mandarin speakers. A summary of the correct percentage of identifications for every native English speakers' 10 vowels, along with standard deviations, is given in Table 2-5. The identification scores for the two less intelligible vowels produced by native English speakers were found to be similar to those reported in a study by Assmann, Nearey & Hogan (1982) which will be discussed in the next section.

### 2.3. DISCUSSION

In general, the Mandarin speakers' productions of English [i eɪ u ou ʌ] were as intelligible as those of the native English speakers. The success may be in part due to the fact that most of these vowels have Mandarin analogs (with the exception of [ʌ]) and that the speakers may have, to some extent, made use of L1 vowel categories in their productions of L2 vowels. Conversely, the native Mandarin speakers' [ɪ ɛ æ ʊ ɒ] were found to be significantly less well identified than their productions of [i eɪ u ou] and than the native English speakers' [ɪ ɛ æ ʊ ɒ] indicating that at least some of them had not learned to produce the vowels without L1 counterparts (except [ɒ]) successfully. However, two important discoveries of exceptions need to be further explored.

First, among vowel categories that do not have obvious Mandarin counterparts, [ʌ] was not significantly less well identified than the vowels that have apparent L1 analogs. On the other hand, [ʌ] was not found to be significantly better identified than [ɒ æ ɛ], suggesting that there were still some differences in degrees of intelligibility between [ʌ] and all the other well identified vowels. It is also important to note that

the native English speakers' [ʌ], which was significantly less well identified than six of the 10 vowels, had a correct rate of identification of 78% as compared to that of 76% for the Mandarin speakers. It appears that English listeners have some problems in identifying this vowel even when it is produced by native English speakers. As in the above mentioned study by Assmann et al. (1982), [ʌ] proved to be the least well identified among the same 10 English vowels that were identical to the ones being tested in the present study. In their study, the authors tested the intelligibility differences of 10 English vowels produced by Western Canadian English speakers (Edmonton area) under different conditions such as consonant environment, isolated vowels, orthographic effect in labeling, speaker information, gated vowel tokens and so on. The results indicated that [ʌ] had the highest error rate in identification under almost all conditions with the exception of gated vowels. For example, in the isolated vowel intelligibility test, a test condition that was similar to the one used in this study, the correct identification for [ʌ] was 32%, which was dramatically lower than the average of 83% for the 10 vowels. The results of the poor identification for [ʌ] observed in both studies may really indicate that native English listeners have difficulty in identifying this particular vowel. No reason for the poor performance on [ʌ] was given in their study. It is not clear as to what is the cause of the poor intelligibility for both [ʌ] and [ɒ] produced by native English speakers observed in this study. However, it is interesting to note that these two vowels are most likely to be confused with each other. The vowel [ɒ] also had a relatively high error rate in the above mentioned study by Assmann et al. In the isolated

vowel context, the % correct ID for [ɒ] was only 73%, 10% lower than the average for the ten vowels being tested.

A second important finding was that the Mandarin speakers' productions of [ʊ] were significantly less well identified than not only the vowels that have L1 counterparts but also the ones that do not have L1 counterparts. Why did the Mandarin speakers have particular difficulty with this vowel? Some possible answers may be revealed upon detailed examination of the confusion data. The confusion pattern for [ʊ] showed that an overwhelming number of [ʊ]s were heard as [u]s and [ou]s, the two adjacent vowels that have Mandarin analogs. A parallel confusion pattern was found for the second least well identified vowel [ɪ]. The majority of misidentified [ɪ]s were heard as [i]s and [eɪ]s, the two nearest vowels that have Mandarin analogs as well. The confusion patterns for these two vowels were similar, although their absolute correct identification scores were different. Although the Mandarin speakers' problems in producing these two vowels were similar as the confusion data suggests, more speakers had trouble with their production of [ʊ] than [ɪ]. The results of Experiment 1 showed that there were apparently more speakers who produced totally unintelligible [ʊ]s than [ɪ]s. One possible reason for this difference may be due to the varying orthography representing the lax back vowel as compared to the relatively simple orthography which represents the front lax vowel. However, as special consideration was given to [ʊ] in the English reading list used in this study, (see Section 2.1.2.) the orthography problem should not be a factor for the poor intelligibility of [ʊ] observed. Another factor might be Mandarin speakers' trouble in keeping the lax vowels in isolation in the reading task, assuming that [ʊ]

could be more intelligible in the word 'book' than in isolation produced in a carrier sentence. The vowel [u] is relatively more restricted in distribution than [ɪ] in English. For example, 'it' /ɪt/, and 'is' /ɪs/ are real words but no words may begin with [u] in English. It may be relatively easier to produce an isolated [ɪ] than an isolated [u]. Further studies are needed before any firm conclusions can be made about this discrepancy in identification scores.

For the other poorly identified Mandarin speakers' back vowel [ɒ], the confusion pattern showed that the majority of the misidentified [ɒ]s were heard as [oʊ]s (28%), an adjacent vowel in the Mandarin system. A considerable number of [ɒ]s were also heard as [ʌ]s, an adjacent vowel in English. The confusion pattern suggested that some of the native Mandarin speakers had avoided simply substituting [oʊ] for [ɒ]. There is also an [ɑ] or [A] (see Chapter 1) in the phonetic space in the Mandarin vowel system. It is possible that some Mandarin speakers produced a more [A]-like English [ɒ] that was misidentified as the English [ʌ]. The phonetic distance between Mandarin [ɑ] and English [ʌ] and [ɒ] will be examined in Experiment 3.

## 2.4. CONCLUSION

The results of this experiment showed that, in general, the 15 Mandarin speakers' productions of English vowels that have Mandarin analogs were significantly more intelligible to the native English listeners than the vowels that do not have Mandarin analogs. The Mandarin speakers' success in their productions of vowels with L1 counterparts was remarkable because their percentage of correct

identifications was comparable to that of the native English speakers' productions of these vowels. In contrast, the Mandarin speakers' productions of the English vowels lacking Mandarin analogs were significantly less successful. The results suggested that there were strong effects of L1 on the production of L2 vowels for the Mandarin speakers. It is possible that some Mandarin speakers simply substituted their L1 vowels in their productions of English vowels. It is also possible that some of them made a little adjustment in an effort to approximate the L2 vowels in their productions. These predictions will be tested in Experiment 3 when acoustic measurements are conducted. Effects of the L1 vowel system on the production of L2 have been reported in previous cross-linguistic studies. Munro (1993) concluded from his study of native Arabic speakers' productions of 10 English vowels that there were 'pervasive' effects of L1 experience on the production L2 vowels.

The results of the intelligibility test seemed to partially support the Contrastive Analysis Hypothesis (CAH) which predicts that sounds that do not exist in the L1 will be more difficult to learn, as the Mandarin speakers did have problems with [ɪ ε æ ʊ]. However, two important points involving the CAH need to be discussed. First, CAH cannot explain the differences in level of difficulty the Mandarin speakers had in their production of vowels that do not have L1 counterparts. For example, the Contrastive Analysis Hypothesis does not predict that [ʊ] would be more difficult than [ɪ ε æ ʌ] for the Mandarin speakers to learn as the results of this study have shown. Secondly, the CAH cannot account for the individual speaker variation in the production of L2 sounds. For example, for the poorly identified

Mandarin accented vowels [u] and [ɪ], each vowel had one speaker whose production was 90% correctly identified. Similarly, three speakers produced a perfect [æ] while another speaker produced a completely unidentifiable [æ] (0%). Therefore, it is important to point out that while some vowels without L1 counterparts are more difficult for some or even a majority of L2 speakers to produce, they are not necessarily difficult for every L2 speaker.

Individual speaker differences in L2 speech learning are an important factor that cannot be ignored in future studies. As discussed in Chapter 1, in a study by Munro, Flege, and Mackay (1996) which involved the study of native Italian speakers' 11 English vowel productions, the results showed that individual differences in speakers' success were found for most of the 11 vowels. They concluded that because of these individual variations, the between-vowel differences in native-like achievement were 'somewhat overshadowed' (Munro et al., 1996). The present vowel production test with native Mandarin speakers demonstrated again that individual speaker variation is an important factor to take into consideration in future studies.

Whether the findings of Mandarin speakers' production of English vowels in this experiment have a perceptual basis will be studied in Experiment 2 in the following chapter.

## **2.5. LIMITATIONS**

This experiment on L2 vowel production entailed careful control over many factors including the speakers' L1 dialect, the number of male and female speakers, and the number of vowels being examined.

The use of isolated vowels without adjacent consonants was another advantage of the present study because consonants can sometimes be a source of accent.

However, limitations of this study must also be noted. First, only one token produced by each speaker representing each vowel was used for identification. Speakers can sometimes show considerable variation in their productions of the same vowel even under same speaking conditions. More tokens of the same vowel would offset the effect of such variation. It may therefore be desirable to include multiple tokens for each vowel segment produced by the same speaker in similar studies in the future.

Another limitation of this study was the speaking context. Although the use of vowels without surrounding consonants in this study can be considered as one of the advantages, isolated vowels for identification are not ideal for an intelligibility test. A complete understanding of vowel production also requires test conditions such as vowels in context in addition to isolated vowels. This is of course because, in natural speech, vowels are rarely produced in isolation as they are in this experiment. These limitations are to be avoided in future studies.

TABLE 2-1

Background information for the 15 Beijing Mandarin speakers

Speaker	Age	Age of Learning	Age of Arrival	Years in Canada	Use of English*
1	28	10	25	3	5
2	26	11	22	4	8
3	38	16	33	6	7
4	26	6	23	3	5
5	36	12	31	5	5
6	31	9	26	5	5
7	33	16	31	3	6
8	44	13	37	7	6
9	34	8	30	4	7
10	31	12	28	3	6
11	28	9	28	0.5	4
12	45	13	40	5	5
13	30	12	26	4	3
14	27	12	25	2	4
15	28	7	23	5	4
<b>Mean</b>	<b>32.5</b>	<b>11</b>	<b>28.5</b>	<b>4</b>	<b>5.3</b>

\* The scale of 1 to 9 represents the Mandarin speakers' estimation of their daily use of English with 1 representing 'never' and 9 'only'.

TABLE 2-2

Confusion Matrix (%) for Mandarin speakers' productions of English

[i ɪ eɪ ε æ u ʊ ou ɒ ʌ] (The percentage of correct identification for each vowel is in bold face.)

Heard	Targets									
	i	ɪ	eɪ	ε	æ	u	ʊ	ou	ɒ	ʌ
i	<b>96</b>	37	1.5							
ɪ		<b>46</b>	0.25	4						
eɪ	4	11	<b>96</b>	13	1					
ε		6	2	<b>63</b>	38					
æ			0.25	19	<b>53</b>	0.25	2			
u						<b>97</b>	47	1.5		
ʊ						2	<b>15</b>	0.5	2.75	5
ou						1	34	<b>95</b>	28	
ɒ					5		4	3	<b>49</b>	17
ʌ				1	3				20	<b>76</b>

TABLE 2-3

Correct % of identifications (mean & standard deviations) of English and Mandarin speakers' 10 English vowel productions by 4 listeners.

Vowel	Native English Speakers				Native Mandarin Speakers							
	Lis.1	Lis.2	Lis.3	Lis.4	Mean	SD	Lis.1	Lis.2	Lis.3	Lis.4	Mean	SD
i	100	100	100	99	100	0	99	93	93	100	96	4
ɪ	99	92	89	93	93	4	72	41	37	35	46	16
eɪ	100	87	99	99	96	6	95	92	100	96	96	3
ɛ	99	88	83	87	89	7	59	72	65	57	63	7
æ	97	100	100	100	99	1	45	48	57	63	53	8
u	100	97	99	100	99	1	89	100	99	99	97	5
ʊ	100	93	91	87	93	5	28	17	5	8	15	9
oʊ	100	97	100	100	99	2	93	89	99	100	95	5
ɒ	81	59	64	83	72	12	67	45	36	49	49	11
ʌ	84	73	77	77	78	5	91	77	80	57	76	15
<b>Mean</b>	96	89	90	93	92	3	74	67	67	66	69	3

TABLE 2-4

Correct % of identifications of 15 Mandarin speakers' productions of 10 English vowels and standard deviations

	i	ɪ	eɪ	ɛ	æ	u	ʊ	oʊ	ɒ	ʌ	F.M	B.M	T.M
M01	100	0	100	90	80	95	0	100	25	65	74	57	66
M02	100	0	100	15	90	100	0	100	95	65	61	72	67
M03	45	25	80	25	55	95	0	95	5	85	46	56	51
M04	100	45	100	100	15	75	30	95	40	90	72	66	69
M05	100	60	100	15	15	100	10	65	60	40	58	55	57
M06	100	40	100	75	100	100	90	100	35	90	83	83	83
M07	100	5	100	30	30	100	5	100	80	100	53	77	65
M08	100	60	100	70	95	100	15	100	10	90	85	63	74
M09	100	75	95	75	100	100	0	100	75	40	89	63	76
M10	100	90	100	45	90	100	0	100	90	75	85	73	79
M11	100	85	100	95	100	100	25	95	85	90	96	79	88
M12	100	75	100	100	10	100	10	95	40	95	77	68	73
M13	100	0	85	20	15	100	0	100	45	75	44	64	54
M14	100	55	100	95	0	95	10	100	40	95	70	68	69
M15	100	80	75	100	5	90	25	85	15	50	72	53	63
Mean	96	46	96	63	53	97	15	95	49	76	71	66	69
SD	14	33	8.4	34	41	6.7	23	9.3	30	20	16	9.1	10

F.M indicates the mean correct % of identifications for front vowels.  
 B.M indicates the mean correct % of identifications for back vowels.  
 T.M indicates the mean correct % of identifications for all 10 vowels.

TABLE 2-5

Correct % of identifications of English speakers'  
productions of 10 English vowels and standard deviations.

	i	ɪ	eɪ	ɛ	æ	u	ʊ	ou	ɒ	ʌ	F.M	B.M	T.M
E01	100	85	95	75	100	100	85	100	85	90	91	92	92
E02	100	100	100	100	100	100	100	100	70	95	100	93	97
E03	100	90	100	50	100	100	95	100	100	45	88	88	88
E04	100	100	90	100	100	100	100	100	60	95	98	91	95
E05	100	95	75	100	95	100	95	100	90	85	93	94	94
E06	100	95	100	90	95	100	100	95	60	80	96	87	92
E07	100	80	100	80	100	100	100	100	0	10	92	62	77
E08	100	100	100	100	100	90	90	100	85	70	100	87	94
E09	100	95	100	95	100	100	95	100	80	70	98	89	94
E10	100	100	100	100	100	100	95	100	45	90	100	86	93
E11	100	100	100	70	100	100	80	100	85	85	94	90	92
E12	95	95	90	95	100	100	100	100	90	80	95	94	95
E13	100	70	100	95	100	100	55	100	45	95	93	79	86
E14	100	100	95	90	100	100	100	95	80	90	97	93	95
E15	100	95	95	95	100	95	100	100	100	90	97	97	97
<b>Mean</b>	100	93	96	89	99	99	93	99	72	78	95	88	92
<b>SD</b>	1.3	8.8	6.9	14	1.8	2.8	12	1.8	26	23	3.6	8.4	5

**F.M** indicates the mean correct % of identifications for front vowels.

**B.M** indicates the mean correct % of identifications for back vowels.

**T.M** indicates the mean correct % of identifications for all 10 vowels.

FIGURE 2-1

The mean % of correct identifications of each English vowel by four listeners

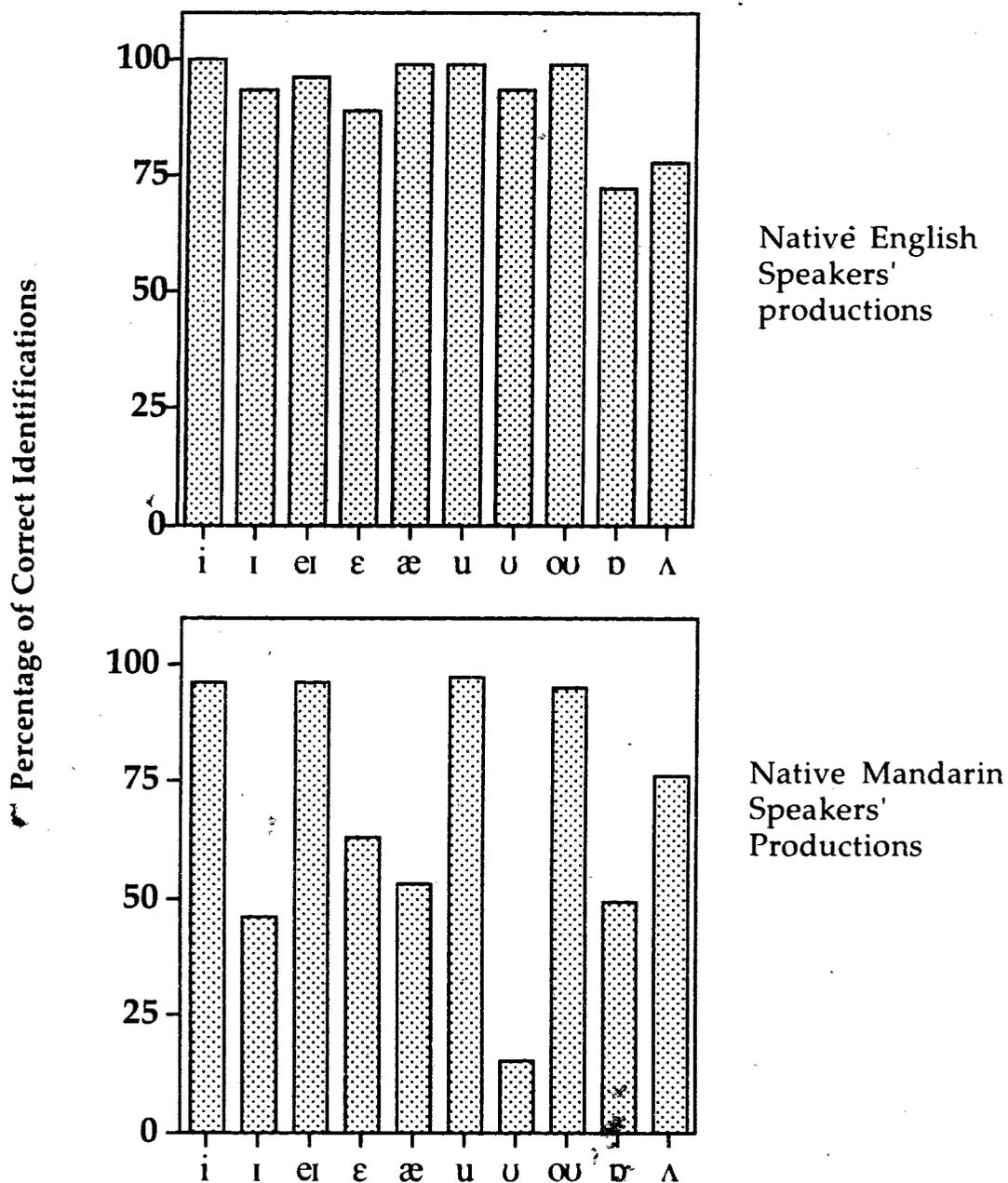
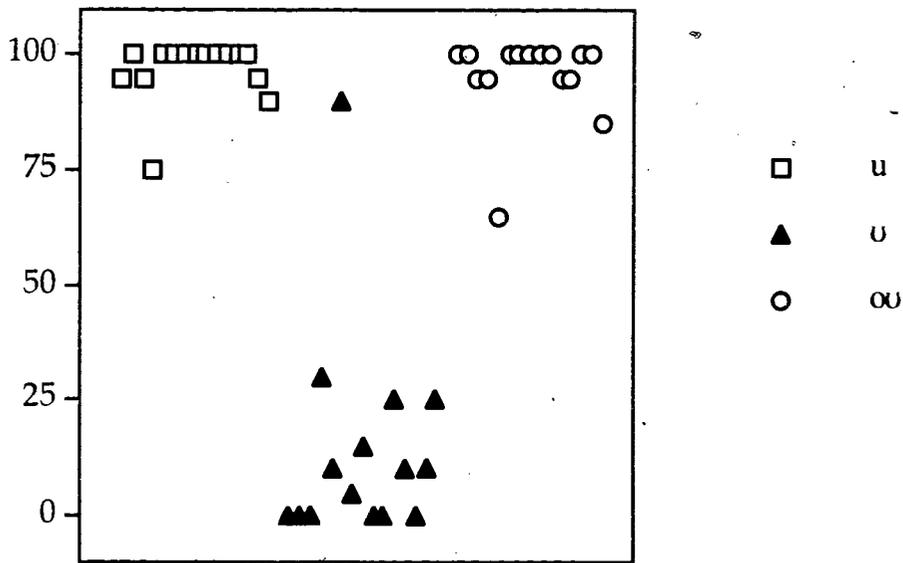
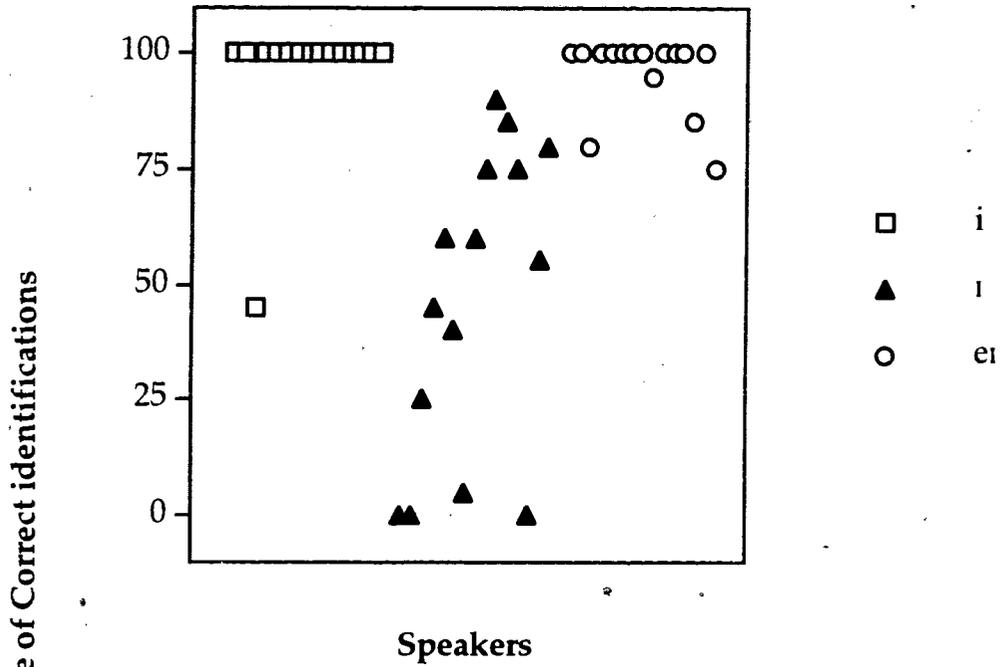


FIGURE 2-2

Individual consistencies in Mandarin speakers' productions of different English vowels



**CHAPTER III**  
**EXPERIMENT 2: THE PERCEPTION OF ENGLISH FRONT VOWELS**  
**BY MANDARIN SPEAKERS**

The results of Experiment 1 indicated that, in general, the 15 Mandarin speakers' productions of English vowels that have Mandarin analogs were significantly more intelligible than those lacking obvious Mandarin counterparts. It is not clear how this pattern in production is related to the speakers' perception. This study will examine the relationship between production and perception of L2 vowels using natural vowel tokens.

The same 15 Mandarin speakers who participated in Experiment 1 will all participate as listeners in Experiment 2. They will identify five English front vowels [i ɪ eɪ ɛ æ] in CV combinations produced by 10 native English speakers. If their success or failure in L2 vowel production is tied to their ability to perceive L1 vowels, it is expected that their success or failure in identifying the five English front vowels should correspond to their production in Experiment 1. Given that their productions of [i eɪ] were significantly more intelligible to the native English listeners than the other vowel productions, it is expected that the Mandarin speakers will be more successful in identifying the English [i eɪ] than [ɪ ɛ æ].

It is often assumed that good perception precedes good production and that failure to produce an L2 sound accurately tends to reflect the lack of a good perceptual representation for that sound. As discussed in Chapter I, Flege's Speech Learning Model posits that an accurate perceptual representation of a sound will eventually be realized in

production (Flege, 1995). Similarly, if a bilingual speaker is unable to discriminate an L2 vowel from the neighboring L2 vowels, and from the L1 vowels that are different phonetically from the L2 vowels, then the L2 vowel will be produced inaccurately. Therefore, L2 vowel production errors may be seen as arising from discriminative failures with certain L2 vowels.

However, some possible counter-evidence to this proposal was reported in a study by Sheldon & Strange (1982), who examined the production and perception of English /ɪ/ and /I/ by native speakers of Japanese. They found that Japanese learners were somewhat more successful in producing the contrast between /ɪ/ and /I/ than in perceiving the distinction between these two sounds. This outcome led them to the conclusion that perceptual mastery of an L2 contrast does not always precede the production of the contrasting phonemes. A parallel result for vowels was found in a study by Blankenship (1991, see Chapter I). The perception test in this study will further explore the relationship between L2 vowel perception and production by adult learners.

### **3.1. METHOD**

#### **3.1.1. Participants**

The same 15 native Beijing Mandarin speakers who participated in Experiment 1 as speakers all participated as listeners in this experiment. The listening task was carried out in a sound-treated room immediately after the Mandarin speakers finished the production task reported in Chapter II.

### 3.1.2. Materials and Procedures

Recordings collected in an earlier study (Munro & Derwing, 1996) were used as listening material. Ten Western Canadian English speakers' productions of the words *beat*, *bit*, *bait*, *bet* and *bat* representing the five English front vowels [i ɪ eɪ ε æ] were all produced in a carrier sentence ("Now I say \_\_\_\_") in sentence final position. They were recorded in a sound treated room on a high quality cassette recorder with a Sony ECM T140 microphone. The target words were carefully separated from the carrier sentences and were digitized at 22.05 kHz with 16-bit resolution on a Macintosh computer and saved as audio files. To prevent distracting effects due to differences in the way the final /t/ was released, waveform editing was used to remove the release of the final [t] at the point of complete constriction.

The tokens for listening identification were recorded on a cassette tape in random order. There were a total of 200 items (10 talkers x 5 words x 4 randomizations). The words for identification were printed on a multiple-choice response sheet with the final (t) in brackets in the order of *bea(t)*, *bi(t)*, *bai(t)*, *be(t)* and *ba(t)* in a single line. There were two sections of 100 numbered lines each, with a space after every five lines. Listeners were instructed to circle the word among the five given words they thought they heard for each item. A voice announced the number of the next item the listener was going to hear on the tape to minimize errors in circling words. The rate at which the listeners heard these words was fixed, and they were not permitted to stop the tape once they started listening. The task, which lasted 20 minutes, was performed in a sound-treated room through headphones with one

listener participating at a time. Five sample items were given to familiarize the subjects with the procedure before the actual task began.

A preliminary screening of tokens was carried out by two trained phoneticians (one male and one female) who were native speakers of Canadian English. The error rates for the two listeners were 2% and 5.5% respectively. The items that were misidentified more than once (N=4) were omitted from any further consideration.

### 3.2. RESULTS

The overall percentage of correct identifications of the vowels by each of the 15 Mandarin listeners is given in Figure 3-1. Across listeners, on average, a total of 73% of the Native English speakers' productions of *beat*, *bit*, *bait*, *bet*, and *bat* were correctly identified. However, there was considerable variation among listeners in the correct rate of identifications. The best listener had a correct identification rate of 87%, while the worst had a much lower rate of 60%. The correct percentage of identification for all five vowels by each individual listener is given in Table 3-1. Obvious differences in the listeners' performance on the vowels can be seen here.

To explore the between-vowel differences, a one-way repeated measures ANOVA with five levels of vowel as a within-group factor was carried out. A significant effect of vowel [ $F(4, 14) = 4.272, p < 0.005$ ] was observed, indicating that some vowels were better identified than others. In order to present a clear view of between-vowel differences indicated by the ANOVA test, the overall correct percentage of identifications of the five front vowels is given in Figure 3-2. A post hoc

Tukey (HSD) test revealed that [eɪ] was significantly less well identified than [i] and [æ] by the Mandarin listeners. No significant differences were found among [ɪ ɛ æ], the three front vowels that lack Mandarin analogs. Moreover, [i] was not significantly better identified than [ɪ ɛ æ] although it was the most intelligible vowel to the Mandarin listeners among the five. Contrary to prediction, [eɪ] was the least well identified vowel among the five.

To investigate whether there were any consistent patterns in the perceptual errors, a confusion matrix giving correct percentages of identification for each vowel was created and is given in Table 3-2. The majority of misidentified [i]s were heard as [ɪ]s (14%), [ɪ]s as [ɛ]s (22%), [eɪ]s as [i]s (23%), [ɛ]s as [æ]s (21%), and [æ]s as [ɛ]s (18%). In general, the confusions indicated a consistent pattern of misidentification of the tokens as "adjacent" vowels. The confusion patterns observed here will be compared with those of the same participants' productions in the next section.

Between-vowel differences were also reflected in each individual listener's best or worst performances on the five vowels. As expected, on average, the best identified vowel was [i]. In fact, eight out of the 15 listeners performed the best in identifying [i] among the five vowels. Only one listener performed the worst on [i]. In contrast, six out of the 15 listeners performed the worst on [eɪ], the most poorly identified vowel of the five. However, one listener's identification of [eɪ] was 100% correct. The average correct identification for [æ] was 79%, which ranked it as the second best. Four out of the 15 listeners performed the best on [æ], including two listeners whose identification scores were 100% correct. However, two listeners' worst performance was on [æ]. Even

though the Mandarin speakers did not identify [ɪ] significantly better or worse than any of the other four vowels, their production of [ɪ] was the least intelligible among the five front vowels in Experiment 1. Two listeners had their lowest identification scores on [ɪ], while one listener performed the best on [ɪ]. Four listeners performed the best and two the worst on [ɛ]. The individual listener's best and worst performances on each vowel are highlighted in Table 3-1.

Considerable listener variation was also observed in the overall perception scores. For example, no single listener appeared to have performed the best on all five vowels. The listener whose average correct identification score was the highest among the 15 subjects heard [i] and [æ] correctly 100% of the time. Yet, at least two other listeners outperformed her on the remaining three vowels [ɪ eɪ ɛ]. Similarly, the listener whose total identification score was the lowest did not perform the worst on every one of the five vowels. The individual differences in perception will be compared with those in production in the next section.

### 3.3. DISCUSSION

Obviously, the perception data failed to support the prediction that both [i] and [eɪ] would be better identified than [ɪ, ɛ, æ]. Even though [i] was the best identified vowel on average, the picture was complicated by the fact that [eɪ] was significantly less well identified than [i] and [æ], and was the least well identified vowel among the five. Given that the Mandarin speakers' productions of [i] and [eɪ] were both highly intelligible to the native English listeners in Experiment 1, and

both have obvious counterparts in Mandarin, it was surprising to see this dramatic discrepancy in the Mandarin listeners' perception scores.

One possible explanation for such a discrepancy might have to do with the differences in the L1 and L2 vowel inventories. While English distinguishes five categories [i ɪ eɪ ɛ æ] in the whole front area, Mandarin contrasts only two [i eɪ]. The squeezing of three additional categories [ɪ ɛ æ] into this area of the vowel space may have a distracting effect on the native Mandarin listeners' identification of [eɪ]. It would probably be much easier for the Mandarin listeners to discriminate only [i] and [eɪ] in the perception test. Such distraction seemed to have relatively less influence on the Mandarin listeners' identifications of [i]. This is probably because [i] is a peripheral vowel. It is interesting to note that [æ], the second best vowel identified by the Mandarin listeners, is the other peripheral vowel in this front area.

The perception data also failed to support the prediction that the Mandarin listeners would have more problems in identifying [ɪ ɛ æ] because these vowels were less successfully produced in Experiment 1 and because they do not have obvious Mandarin counterparts. As the results of the statistical tests indicated, although [i] was the best identified vowel of the five by the Mandarin listeners, it was not significantly different from [ɪ ɛ æ] in terms of correct rate of identification. Moreover, in contrast to the results of the production test in Experiment 1, the Mandarin listeners' perception of [æ] was significantly better than that of [eɪ].

To conclude, data from this identification test failed to support the prediction that native Mandarin listeners would perform better in their identification of the English [i eɪ] than [ɪ ɛ æ]. A comparison with the

results of Experiment 1 revealed discrepancies between the Mandarin speakers' perception and production of the English front vowels. These discrepancies will be further analyzed along with the individual speaker differences in the following section.

#### **3.4. COMPARISON OF EXPERIMENT 1 & EXPERIMENT 2**

As discussed in the above section, there were discrepancies between the Mandarin speakers' production and perception of the five English front vowels, with or without Mandarin counterparts. In order to further explore the relationship of the Mandarin speakers' performance in production to their perception, a correlational analysis was carried out. The means of the perception and production scores on the five front vowels by each speaker (N=15) were compared by computing a Pearson Correlation Coefficient. The correlation  $r(14) = 0.49$ ,  $p < 0.05$  was just significant. While the result showed no strong correlation between perception and production, it did suggest that there was some relationship.

Certain trends were found when the mean scores for both perception and the production of the five front vowels were examined. First, the Mandarin speakers seemed to perceive the "new" vowels [ɪ ε æ] much better than they produced them in terms of the absolute mean scores. This tendency toward better perception than production was especially obvious with [ɪ æ] and is clearly reflected in Figure 3-3. For the vowel [ε], which also lacks an L1 counterpart, such a tendency was less obvious. On the other hand, [i] and especially [eɪ] showed the opposite tendency - better production than perception.

The better performance in perception than in production observed for the "new" vowels [ɪ ɛ æ] has a fairly straightforward interpretation. First, in learning to perceive the English vowels that do not contrast in the L1, Mandarin speakers must learn to discriminate the English categories [ɪ ɛ æ] from their L1 categories [i] and [eɪ]. In addition, they must discriminate the English vowels [ɛ] and [æ] (both are "new") from each other for phonological contrast. In performing the task of identifying the five vowels, [i ɪ eɪ ɛ æ], the Mandarin subjects had to make a forced choice. To some extent, the Mandarin listeners did seem to perceive the differences between the three "new" vowels and the neighboring L1 vowels [i] and [eɪ]. However, they were apparently not always able to make full use of their perceptual knowledge when they produced the same vowels. More evidence was found when the individual listeners'/speakers' performance was analyzed. For example, some cases of total failure in the production of [ɪ] and [æ] were found for a few Mandarin speakers (0% correctly identified by native English listeners). However, not a single Mandarin speaker was found to have scored 0% on the perception of [ɪ ɛ æ].

On the other hand, the observation of a reverse tendency in the relationship between production and perception for [i] and [eɪ] is more difficult to account for. As discussed earlier, the crowding of the three "new" vowels [ɪ ɛ æ] into the front vowel space may have caused some distraction for the Mandarin speakers in the perception task. This may have influenced their performance on [i] and especially on [eɪ]. Apart from this, no further explanation can be offered at this time.

The data presented here suggest that there were differences in the perceptual representations of the vowel categories by the native

English and native Mandarin speakers. To further explore this possibility, the confusion matrices for both the perception and production tests were examined. Fairly complex patterns of confusion were found for some of the vowels. First, when the Mandarin listeners misidentified the native English speakers' [ɪ]s, for the most part, they heard them as [ɛ]s (22%) rather than as [i]s (2%) or [eɪ]s (2%). In contrast, when the Mandarin speakers' productions of [ɪ] were misidentified by the native English listeners, they were mostly heard as [i]s (37%) and [eɪ]s (11%) rather than as [ɛ]s (6%). This confusion indicated that at least some Mandarin speakers had noticed some of the phonetic differences between [i] and [ɪ]. Yet, their phonetic representation of [ɪ] may not have been specified in the same way that native English [ɪ] is specified. As a result, the Mandarin speakers' productions of the English [ɪ] did not meet with the native English speakers' phonetic standard and were heard as [i]. This may be due to their reliance on duration only when producing the [i] and [ɪ] distinction. Such a reliance on duration may have been beneficial in the perception task if durational differences between [i] and [ɪ] were available, but not in the production task, because native English listeners would probably not distinguish [i] from [ɪ] on the basis of duration. Results of previous studies indicated that native speakers of English rely mainly on spectral cues in distinguishing the English [i]-[ɪ] pair (Bohn, 1995, Munro, 1993).

The confusion patterns for [ɛ] and [æ] were similar in both perception and production as they were mistaken for each other most of the time. In spite of the different levels of success in the perception and production tests, the confusion patterns suggested that the Mandarin speakers did not simply substitute [eɪ] for these two "new" categories [ɛ]

and [æ]. Many of them were aware of the phonetic differences between the L1 [eɪ] and L2 [ɛ] and [æ] but still had problems in distinguishing the two close L2 vowels [ɛ] and [æ].

Because much between-listener and between-speaker variation was observed in both the production and perception tests, a comparison of the individual speaker/listener's performances in both experiments was necessary. The production and perception scores to be compared were the average scores of the five front vowels. First, the vowel that had the widest discrepancy between perception and production ([eɪ]) was examined. Of the six listeners who performed the worst in their perception of [eɪ] in Experiment 2, four of them had a perfect production (100%) score for this same vowel in Experiment 1. Of the four, one listener's perception score for [eɪ] was only 22%. The remaining two listeners' productions of [eɪ] earned intelligibility scores of 95% and 75% respectively. In fact, most of the speakers' production and perception scores for [eɪ] were 30% or more apart. The discrepancy between production and perception was indeed enormous for these speakers.

Comparing other vowels, three speakers produced totally unintelligible [ɪ]s in Experiment 1 but their perception scores for the same vowel were 63%, 85%, and 78%. Two of the three speakers received above average (73%) perception scores for [ɪ]. In contrast, the listener whose perception score for [ɪ] was 93%, the highest of all the listeners, had a production score of only 45% that was far below the average. As the examples showed, the discrepancy between the performance in perception and in production was not unusual. However, there were cases in which the speaker performed well or badly on both tasks. For example, four of the speakers' [ɪ]s had very close perception

and production scores that were no more than 5% apart. Two speakers had perfect scores for [æ] in both perception and production. Similarly, the speaker who performed the worst on [ɛ] in production (25%) was also the worst in her perception of [ɛ] (15%). However, the speaker whose perception was the best on [ɛ] (91%) had a production score for [ɛ] that was far below average (45%). For the vowel [i], while most speakers did well in both production and perception, the only speaker whose production of [i] was not 100% correctly identified (45%) was also found to be the one who had the lowest perception score for [i] (55%). To conclude, both discrepancies and consistencies were found between individual speakers' performances in their perception and production of the five front vowels. The implications of this finding will be discussed later.

### 3.5. CONCLUSION

The results of the two experiments showed that there was considerable inconsistency in success between Mandarin speakers' perception and production of the five English front vowels. The expectation that accurate production of L2 sounds has a clear perceptual basis was not fully supported by this study. The statistical analysis indicated a marginal correlation between perception and production of the five front vowels. For the English front vowels that lack Mandarin counterparts, [ɪ ɛ æ], there was a tendency toward better perception than production. For the two front vowels that have Mandarin counterparts, [i eɪ], the opposite tendency, better production than perception, was observed.

The comparison of the confusion patterns in the two experiments suggested that some Mandarin speakers may have established phonetic categories for the "new" vowels [ɪ ε æ] in their perception, even though they still could not produce them as successfully. This was clearly shown by the patterns of the misidentified [ɛ] and [æ] that were mostly confused with each other rather than with [eɪ], the only L1 vowel that is in this area of the vowel space. Similarly, the confusion pattern for [ɪ] indicated that Mandarin listeners did not confuse this "new" vowel with the nearest [i] category in the perception task. Both [ɪ] and [æ] were much better perceived than produced by most Mandarin speakers, although occasional counterexamples were observed. Some speakers' productions of [ɪ] and [æ] were totally unintelligible to the native English listeners. No listeners, however, had a perception score of 0% on any one of the five vowels tested.

In contrast, the vowel [eɪ], which posed no problem for the Mandarin speakers in their production proved to be extremely difficult in perception. This was consistently confirmed by the individual data. Not a single speaker had a perception score that was higher than his or her production score for [eɪ]. At least, when an identification paradigm such as the one used here serves as the criterion, it cannot be concluded that accurate perception must precede accurate production.

However, the conclusion that accurate production of an L2 vowel can precede a good perceptual representation of that vowel cannot be drawn simply based on the Mandarin speakers' performance with [eɪ]. No direct comparison of the present findings can be made with Sheldon and Strange's (1982) finding of Japanese speakers' better production than perception of the English /ɪ/ and /ɪ/. Several arguments can be

provided here. First, English [eɪ] has never been reported as a problem in perception and production for Mandarin learners in any previous studies. Nor has it been described as a problem in any textbooks for pedagogical purposes. In contrast, the English /ɪ/ and /i/ distinction has usually been reported as a problem for Japanese learners of English (Sheldon & Strange, 1982, Flege et al, 1995). From a practical standpoint, Japanese speakers have to discriminate this pair of liquids for intelligibility in both perception and production. English [eɪ], however, is similar to Mandarin [eɪ], and substitution or approximation is not expected to cause serious intelligibility problems. Furthermore, the nature of the identifications of the present test was different from that of Sheldon and Strange's. Five alternatives were included in the identification task, as opposed to only two in their study. As discussed earlier, this may have diverted the Mandarin listeners attention in their perceptual performance. Presumably, the Mandarin speakers' performance on [eɪ] might be significantly better if they were told to identify two vowels instead of five. More studies are needed for a better understanding of the discrepancies found with [eɪ] in the perception and production tests.

In conclusion, it is important to point out the limitations of the two experiments reported so far. One problem is that there was a difference in the tokens used in the two tests. In Experiment 1, the tokens for identification were in /bVt/ syllables with the final /t/ removed, but in Experiment 2, isolated vowels were used. This difference makes the comparison of the results of the two experiments somewhat more difficult to interpret.

The relationship between perception and production is a complex issue. The tendencies noted in Experiment 1 and Experiment 2 need further testing in Experiment 3 when the acoustic properties of the vowels are analyzed. The results of the acoustic analyses will shed light on some of the problems found in the production and perception tests.

TABLE 3-1

Correct % identifications of English [i eɪ ε æ] in /bVt/ by 15 Mandarin listeners.

For each listener, the worst identified vowel is in bold face and the best is underlined.

	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	Total
i	<u>90</u>	73	55	<u>63</u>	<u>83</u>	<u>100</u>	78	<u>95</u>	98	80	<u>95</u>	<u>90</u>	83	<u>75</u>	<u>95</u>	83
ɪ	63	85	55	<u>93</u>	68	78	58	70	53	80	90	78	78	60	85	73
eɪ	64	<u>100</u>	69	92	58	78	50	<b>64</b>	<b>33</b>	<b>22</b>	<b>50</b>	69	<b>61</b>	69	<b>61</b>	63
ε	72	69	53	78	<b>25</b>	81	<b>41</b>	81	84	<u>91</u>	81	59	<u>84</u>	<b>59</b>	81	69
æ	78	<b>64</b>	<u>81</u>	83	67	<u>100</u>	<u>94</u>	89	<u>100</u>	67	94	<b>56</b>	72	61	81	79
<b>Total</b>	73	78	63	82	60	87	64	80	74	68	82	70	76	65	81	73

TABLE 3-2

Confusion Matrix (%) for Mandarin listeners' perceptions of English  
 [i ɪ eɪ ɛ æ] (The % of correct identifications for each vowel is in bold face)

Heard	Target				
	i	ɪ	eɪ	ɛ	æ
i	<b>83</b>	2	23	0	1
ɪ	14	<b>73</b>	12	6	1
eɪ	1	2	<b>63</b>	3	1
ɛ	2	22	1	<b>69</b>	18
æ	0	2	1	21	<b>79</b>

FIGURE 3-1

Individual Mandarin listeners' perception scores (%) for English [i ɪ eɪ ε æ] produced by 10 native English speakers.

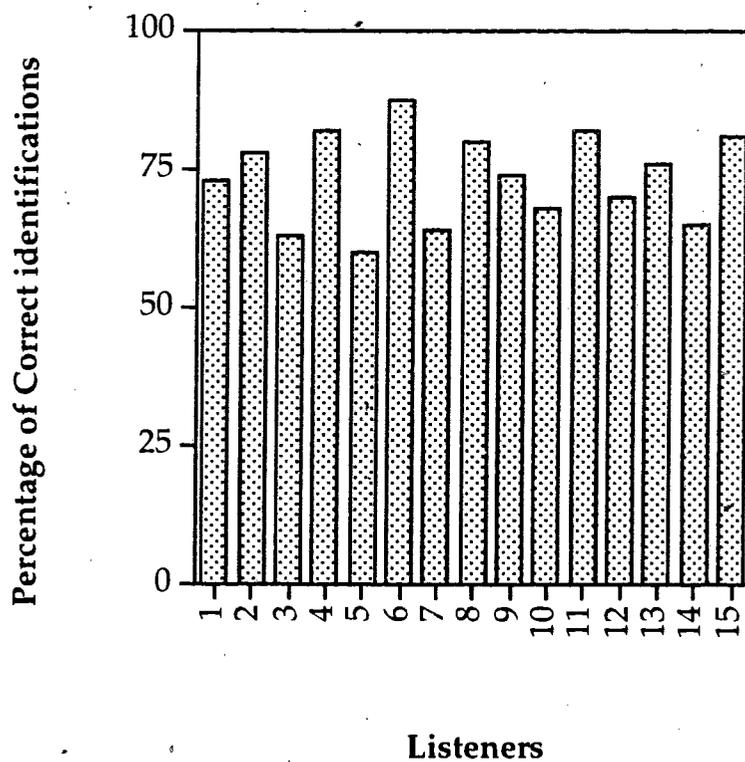
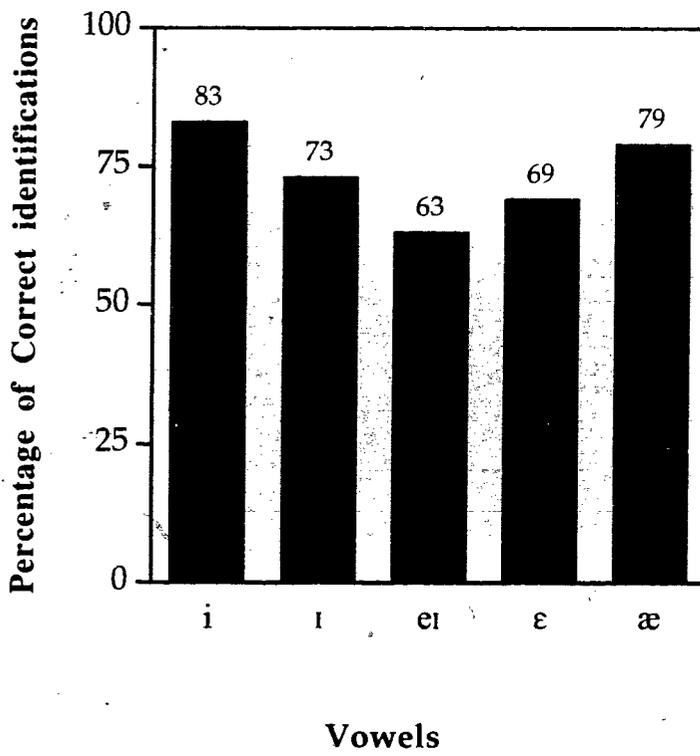


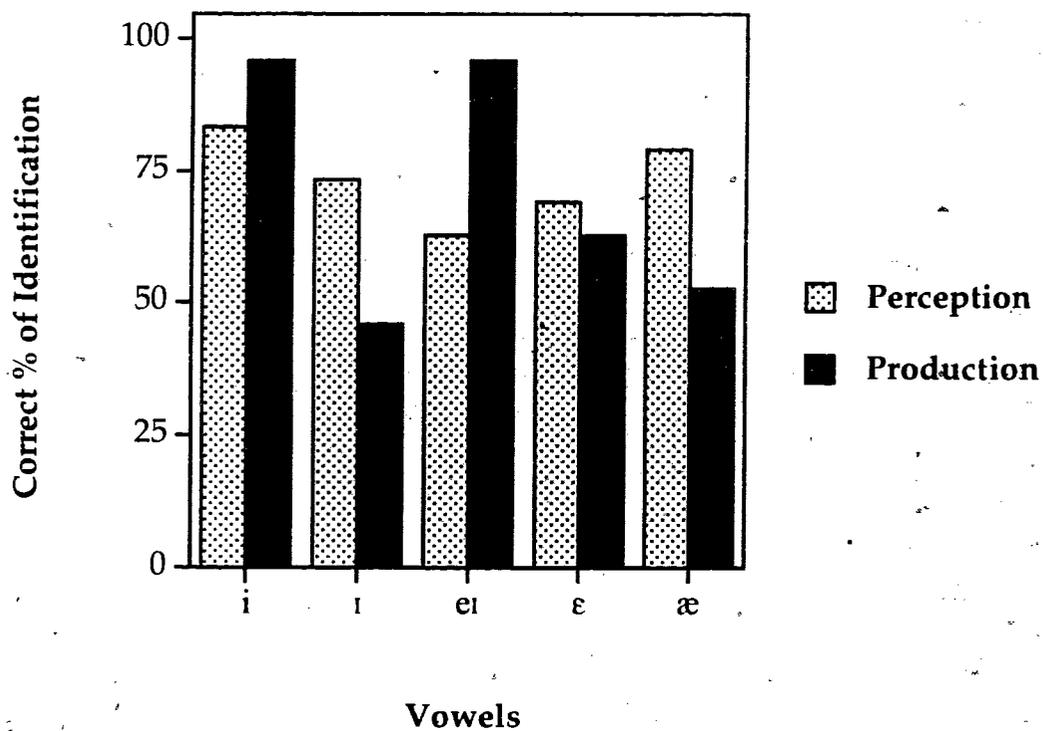
FIGURE 3-2

Mean correct identifications (%) of English [i ɪ eɪ ε æ] in /bVt/ by Mandarin listeners



**FIGURE 3-3**

Mandarin speakers' perception and production scores (%) on five English front vowels



## CHAPTER IV

### EXPERIMENT 3: ACOUSTIC MEASUREMENTS

The purpose of this experiment is to investigate the similarities and differences, in terms of acoustic properties, among English, Mandarin and Mandarin-accented English vowels. There are several reasons for following up the production and perception tests described in Chapter 2 and Chapter 3 with such an analysis. For example, an acoustic analysis might help to determine the extent to which the Mandarin speakers have established separate vowel categories for English vowels, as opposed to simply substituting Mandarin vowels. The acoustic analysis also aims to explore whether Mandarin-accented vowels tend to be 'pulled' (in spectral term) in the direction of the native Mandarin vowels as a result of the effect of L1 on L2 and to what extent they tend to approximate the target English vowels. With regard to L2 speech learning theory, Flege's Speech Learning Model (SLM) will be examined. An early version of the SLM states that "Equivalence Classification" will prevent adult L2 learners from establishing phonetic categories for "similar" but not "new" L2 phones. More recently, however, Flege has modified this proposal to include the notion of 'perceived phonetic dissimilarity' between an L2 sound and the closest L1 sound (Flege, 1995). According to the SLM, the greater the phonetic dissimilarity between the L1 and L2 sounds, the more likely it is that the phonetic differences between the sounds will be discerned. If this is the case, it is expected that the Mandarin-accented vowels that lack obvious English counterparts (new phones with more phonetic dissimilarity) will demonstrate more degree of approximation toward

the native English targets. However, a fine-grained analysis will not be possible here because the speakers' perceptions of phonetic dissimilarity have not been assessed. The acoustic measurement data will be analyzed and compared with those of the two previous experiments in order to explore the relationship between acoustic properties of vowels in terms of formant frequency values and duration, and their degree of intelligibility.

A number of previous studies have compared the phones of two languages in terms of a variety of acoustic properties. For instance, Flege (1987) compared the VOT values of both French and English /t/ and F2 values of French and English /u/ by monolingual speakers of both languages. Measurements of the accented /t/ and /u/ produced by non-native speakers of the two languages were then compared with the native speakers' productions to determine how the nonnative speakers approximated the L2 phones in terms of VOT values for /t/ and in terms of F2 values for /u/. The results indicated that the non-native French and English /u/ differed significantly in F2 values from the native English or French /u/. Similarly, the non-native /t/ either differed significantly from the native /t/ or was intermediate between the values of the monolingual French or English /t/ in VOT values. The study also included a new French vowel for the native English speakers: French /y/. The native English speakers who were experienced in French did not differ significantly from French monolinguals in their production of /y/. Flege concluded that equivalence classification limited their approximation of the native French and English similar phones /u/ and /t/, but not of the new phone /y/.

In another study, Munro (1993) compared the acoustic properties of Arabic and English vowels as well as Arabic-accented vowels in terms of duration, formant values and formant movement. The Arabic-accented English vowels were rated for degree of accentedness by native speakers of English. The results showed that some systematic properties of Arabic-accented vowels, both temporal and spectral, could be explained by the properties of Arabic vowels. For example, the native Arabic speakers probably exaggerated the duration differences between the English tense and lax vowel pairs because their L1 contrasts long and short vowels. In another study, Bohn and Flege (1992) compared the acoustic properties of four English front vowels [i, ɪ, ε, æ] with five German front vowels [i, ɪ, ε, ε:, a] in terms of duration and formant frequencies. Two groups of native German speakers that differed in their experience in English participated. Their productions of the four English vowels were also tested for intelligibility by native English listeners. The results of the intelligibility test and acoustic measurements were then compared to see which acoustic properties the German speakers "carried over" in their productions of English vowels. The results indicated that the amount of L2 experience did not affect the learning of similar vowels, but to some extent, affected the learning of new vowels.

However, no acoustic study comparing Mandarin speakers' productions of English vowels with those of native English speakers has yet been reported. In this study, the acoustic properties of Mandarin vowels and English vowels will be measured and compared in terms of duration, and F1 and F2 frequencies. The comparisons will be made only with those segments that are comparable to each other, i.e.

corresponding vowels in the two systems will be compared. Therefore, only five Mandarin vowels [i eɪ u ou a] produced by the 15 native Beijing Mandarin speakers in Experiment 1 will be measured and compared to the English counterparts produced by the 15 native English speakers in that experiment. The results will then be used to determine the relationship in terms of quality (F1-F2 values) and temporal differences between these vowels cross-linguistically. Although there are problems in classifying the vowels of two languages in terms of being "identical", "similar" or "new" as discussed in Chapter 1, it is, nevertheless, convenient for the comparisons of those segments that exist in both languages to see how "similar" or close they really are in terms of the acoustic dimensions mentioned above. It is also important to point out that the purpose of the comparisons of the vowels that have counterparts in both languages is to seek some possible explanations for the results of the previous perception and production tests rather than to classify which English vowels are either "similar" or "new" to Mandarin speakers. In simple terms, this comparison procedure serves only as a method rather than a goal in itself.

The productions of the 10 English vowels by both the native Mandarin and English talkers in Experiment 1 will also be measured and compared in terms of duration and F1 and F2 values. It is expected that an examination of the acoustic properties of the Mandarin Chinese vowels and the Mandarin-accented English vowels will shed some light on some of the results observed in the previous two experiments.

## 4.1. METHODS

The same 300 isolated English vowel tokens [i ɪ eɪ ɛ æ u ʊ ou ɒ ʌ] produced by 15 native English and 15 native Mandarin speakers that were examined for intelligibility in Experiment 1 were used as tokens for measurement (see Experiment 1 for more descriptions). Seventy-five isolated Mandarin vowel tokens [i eɪ u ou a] produced by the same 15 native Mandarin speakers were also measured. The vowel tokens were all measured for duration and F1 and F2 values on an Apple Macintosh computer using Signalyze software. These were all isolated vowels (except Mandarin [eɪ] which was produced in /peɪ/ due to the lack of isolated /eɪ/ in Mandarin) digitized earlier for the intelligibility test (see descriptions in Experiment 1). Vowel duration was measured to the nearest 1 ms. on a waveform display by placing boundary cursors at the beginning of the first pitch period and the end of the last. For each duration measurement, a simultaneous wide-band spectrogram was displayed just below the waveform to help with judgment of the starting and end points of the vowel. The frequencies of F1 and F2 were obtained at approximately 30% and 70% of the distance into the vowel. In most cases, FFT analyses was used to obtain the formant values. In general, the wide-band selection was used to measure male voices and the very-wide or extra-wide band selection was used for female voices to accommodate the differences in fundamental frequencies. In doubtful cases or when there was extreme difficulty in observing any candidates for the intended formants, especially with back vowels whose F1 and F2 were close, Wide Cone Kernel and LPC analysis were used as alternatives. The results of the two were then compared to determine

the possible values. It was estimated that in about 5% of cases these alternatives were resorted to. All the tokens were measured and checked at least twice and some of them four times before the final results were determined.

## **4.2. RESULTS**

### **4.2.1. Mandarin vowels [i eɪ u oʊ ə] and their English counterparts**

#### **4.2.1.1. Duration**

The mean durations by vowel (in msec.) for the two groups are given in Table 4-1. There was a tendency for the English vowels to be longer than their Mandarin counterparts. To test the vowel duration differences between the Mandarin [i eɪ u oʊ ə] and English [i eɪ u oʊ ɒ], a two-way mixed design ANOVA was carried out with language (Mandarin or English) as a between group factor and vowel (5 levels) as a within group factor. The results indicated the effect of language [ $F(1,28) = 5.8, p < 0.05$ ] was significant, because the English vowels tended to be longer overall than the Mandarin ones. The effect of vowel [ $F(4,112) = 1.81, p > 0.10$ ] and interaction of language and vowel [ $F(4,112) = 1.17, p > 0.30$ ] were non-significant, suggesting that the patterns of duration differences among the five vowels in the two language groups were not significantly different.

#### **4.2.1.2. Spectral data**

To compare the average formant frequency values between speaker groups, it is important to control for variability caused by

gender differences. To compensate for the gender differences, the F1 and F2 values were normalized using a log mean transformation, a method developed by Nearey, (see e.g. Hindle, 1978). Close attention must be paid to the procedures of vowel frequency normalization for cross language study. In this experiment, two separate normalization procedures were applied for the native English vowels. First, to compare the differences between English and Mandarin-accented vowels, both the native English and Mandarin-accented vowels were normalized using a log mean transformation on the basis of an equal number of 10 vowels. For the purpose of comparing the English and Mandarin vowels, however, a different normalization procedure was used to normalize the native English vowels based only on the five vowels that were to be compared with the five relevant Mandarin vowels. This was because the vowel systems of the two languages differ in distribution and overall number of vowels. A normalization procedure based on the total number of vowels in English (10) being studied in this experiment and five Mandarin vowels would result in "over normalization"<sup>1</sup>

A series of two-tailed *t*-tests were then carried out on the log mean transformed F1 and F2 values of both English and Mandarin vowels to examine the differences between them. For the vowels [ɹu a], the values of measurement (a) (the measurement made at about 30% of the distance into the vowel) were compared. For the two diphthongs,

---

<sup>1</sup> Disner (1986) explained this point using the example of French and Italian vowel systems. French has twice as many front vowels as Italian does, but the back vowels are comparable. Normalizations based on the means of French and Italian vowels would not be comparable with each other because the overall formant means would pull the French front vowels back in the phonetic space. In order to avoid such a distortion in this comparison, a special precaution was taken to normalize the native English vowels differently for the two comparisons.

[eɪ ou], both measurements (a) and (b) (the measurements made at about 30% and 70% of the distances into the vowel) were compared, because these vowels show noticeable movement and therefore cannot be characterized by a single measurement.

The results of the tests, along with the means and standard deviations are summarized in Table 4-2. No significant spectral differences were found between the Mandarin and English vowel [i]. For [eɪ], the results showed that the Mandarin [eɪ] was significantly lower [ $t(28) = 2.48, p < 0.05$ ] and more posterior than the English counterpart [ $t(28) = 3.04, p < 0.01$ ] at measurement (a) but the difference was not significant at measurement (b). The Mandarin [ou] proved to be significantly lower at both measurement (a) [ $t(28) = 2.48, p < 0.05$ ] and (b) [ $t(28) = 2.50, p < 0.05$ ] but more posterior than the English counterpart at only measurement (b) [ $t(28) = 2.65, p < 0.05$ ]. For [u], no significant differences were found in F1 across languages. However, there was a large difference in F2 values, [ $t(28) = 8.79, p < 0.0001$ ], indicating that Mandarin [u] was indeed significantly more posterior than the English counterpart. For the Canadian English [ɒ] and Mandarin [ɑ], significant differences were revealed in both F1 [ $t(28) = 4.07, p < 0.001$ ] and F2 [ $t(28) = 5.85, p < 0.001$ ]. The Mandarin [ɑ] was lower and more anterior than the Canadian English [ɒ].

In Experiment 1, the results of the intelligibility test showed that the rate of correct identification for the native Mandarin speakers' productions of English [ɒ] was significantly lower than that for their productions of [ʌ]. Additionally, the confusion matrix for [ɒ] showed that a considerable number of intended [ɒ]s were heard as [ʌ]s. These factors prompted the comparison between formant values of English [ʌ] and

Mandarin [ɑ]. It had never occurred to the experimenter to relate this pair as possible counterparts before, partly because of the use of different phonetic symbols for them. When the *t*-tests were performed on these two vowels, the results also showed significant differences in F1 [ $t(28) = 3, p < 0.01$ ], and in F2 [ $t(28) = 2.82, p < 0.01$ ] between the English [ʌ] and Mandarin [ɑ]. However, although the Mandarin [ɑ] was significantly different from both English [ɒ] and [ʌ], the difference between [ɑ] and [ʌ] was less extreme, particularly in terms of F2.

Data from this study showed that in fact, the Mandarin [ɑ] is acoustically closer to the Canadian English [ʌ] than to [ɒ] in terms of spectral properties. To test for a vowel duration difference between the Mandarin [ɑ] and English [ʌ], a one way ANOVA was carried out. No effect of language on duration was revealed [ $F(1,28) = 1.89, p > 0.1$ ]. The result of the duration test between these two vowels is also given in Table 4-1. The Mandarin [ɑ] was longer than the English [ʌ] but the difference was non-significant. Therefore, instead of [ɒ], the English [ʌ] could be considered as somewhat 'similar' to Mandarin [ɑ]. The mean F1 and F2 (normalized values) of English [i eɪ ou u ɒ ʌ] and Mandarin [i eɪ ou u ɑ] are plotted in Figure 4-1.

#### 4.2.1.3. Discussion

The comparisons of the durations of English [i eɪ u ou ɒ ʌ] and Mandarin [i eɪ u ou ɑ] indicated that all the English vowels were longer than the Mandarin counterparts with the exception of [ʌ], which is non-significantly shorter than the Mandarin [ɑ]. While there was a tendency for English vowels to be longer than their Mandarin counterparts, it is important to note that the Mandarin vowels in this study were all

produced in sentence final position in the fourth tone (falling tone), the environment and tone that are associated with the shortest vowel durations as demonstrated in earlier studies (Ho, 1976). It is possible that the environment and tone could be a determining factor in the short vowel durations that were obtained in this study.

The cross-language spectral comparisons indicated that there were no differences between the English and Mandarin [i]. These two vowels are almost identical spectrally as reflected in Figure 4-1. Spectral similarity was found between the English and Mandarin [u] in vowel height, but the English [u] was considerably more anterior than the Mandarin counterpart. Although the Mandarin [u] has often been described as noticeably higher and more posterior than English [u] (Norman, 1988), the difference in vowel height was not confirmed by the data in this study. There were also noticeable spectral differences between the two diphthongs [ei ou] cross-linguistically. Both Mandarin [ei ou] were lower and more posterior than their English counterparts. Therefore, the Mandarin [u ei ou] can be considered to be similar but not identical to the English counterparts. The comparisons of English [ɒ] and [ʌ] with the Mandarin [ɑ] showed spectral differences between both pairs. Strictly speaking, Mandarin [ɑ] has no clear counterpart in English as it was found to be significantly different from both the English [ɒ] and [ʌ]. Yet, the results indicated that Mandarin [ɑ] was closer to English [ʌ] in terms of spectral properties. In fact, the Canadian English [ɒ] is more toward the back of the vowel space than the Mandarin [ɑ], which is realized as a low central vowel [ʌ] syllable finally.

## 4.2.2. Comparison of native English and Mandarin-accented vowels [i ɪ eɪ ε æ u ʊ ou ɒ ʌ]

### 4.2.2.1. Duration

Table 4-3 shows the mean durations and standard deviations for the 10 English vowels [i ɪ eɪ ε æ u ʊ ou ɒ ʌ] produced by the 15 native speakers of English and 15 native speakers of Mandarin. A two-way ANOVA test revealed no effect of first language [ $F(28) = 0.114, p > 0.73$ ]. There was an effect of vowel [ $F(9) = 25.15, p < 0.0001$ ] and an interaction of first language and vowel [ $F(9) = 2.924, p < 0.01$ ]. However, tests of simple main effects showed that only the Mandarin-accented [i] was significantly longer than the native English [i]. For the remaining nine vowels, duration differences were non-significant.

The two groups showed different durational patterns in their productions of both the front and back vowels. For the English speakers, the front vowels showed the pattern of  $eɪ > i > æ > ε > ɪ$  while the Mandarin-accented vowels had the pattern of  $i > eɪ > æ > ε > ɪ$ . The only difference was in the ordering of [i] and [eɪ] because the Mandarin speakers produced a very long [i]. For the back vowels, the native English speakers showed a pattern of  $u > ou > ɒ > ʊ > ʌ$  and the Mandarin speakers showed the pattern of  $ou > u > ʊ > ɒ > ʌ$ . To test for within group differences in vowel duration, a one-way repeated measures ANOVA was carried out for both groups. The results revealed an effect of vowel for both the native English speakers [ $F(9) = 22.95, p < 0.0001$ ] and for the Mandarin speakers [ $F(9) = 11.99, p < 0.0001$ ]. Post-hoc Tukey HSD tests revealed the native English group's productions [eɪ i u ou æ ɒ] were significantly longer than their [ʊ ε ʌ ɪ]. The results indicated that the

native English speakers showed a tendency to produce longer tense vowels than lax vowels. However, these results differ somewhat from the expected pattern of longer low vowels than high vowels. Returning to the Mandarin speakers' vowel duration differences, the post-hoc Tukey HSD test showed a different pattern from that of the native English group. The Mandarin speakers' [ɒ] was significantly shorter than their [i eɪ u oʊ]; their [æ] and [ʊ] were not significantly shorter than the tense vowels [eɪ u oʊ]; and their [i] was longer than all the vowels except [eɪ u oʊ].

The duration ratios from the two groups on the 3 pairs of tense and lax vowels [i]-[ɪ], [eɪ]-[ɛ], and [u]-[ʊ] were calculated and compared to see if the native and non-native speakers applied a different strategy in their productions of these vowel pairs. The tense and lax vowel duration ratios for the Mandarin-accented vowel pairs were 2 for [i]-[ɪ]; 1.6 for [eɪ]-[ɛ]; and 1.3 [u]-[ʊ]. For the native English vowel pairs, the ratios were 1.5, 1.4, and 1.3 respectively. The Mandarin group seemed to have exaggerated the [i]-[ɪ] difference. Because the Mandarin speakers produced a much longer [i] than the English speakers, and their [ɪ] was shorter, the duration ratio of [i] and [ɪ] was considerably larger for the Mandarin speakers than for the native English speakers.

#### 4.2.2.2. Spectral Data

A series of two-tailed *t*-tests were carried out on the log mean transformed F1 and F2 values of the English and Mandarin-accented vowels to examine the spectral differences between them. For the vowels [i ɪ e æ u ʊ ɒ ʌ], the values based on measurement (a) (the measurement made at about 30% of the distance into the vowel) were

compared. For the two diphthongs, [eɪ oʊ] both measurements (a) and (b) (the measurements made at about 30% and 70% of the distances into the vowel) were compared. The results of *t*-tests along with the table of means, and standard deviations are summarized in Table 4-4. Significant differences were revealed in F1 for [ɪ] [*t*(28)= 2.36, *p* < 0.05]; for [eɪ] at measurement (a) [*t*(28)= 2.75, *p* < 0.05]; for [ʊ] [*t*(28)= 2.35, *p* < 0.05]; for [ʌ] [*t*(28)= 2.21, *p* < 0.01), and for [oʊ] [*t*(28)= 2.89, *p* < 0.01] at measurement (a). In general, more dramatic differences were revealed in F2 than in F1. Differences in F2 were found for [ɪ] [*t*(28)= 3.73, *p* < 0.001]; for [æ] [*t*(28)= 2.86, *p* < 0.01]; for [u] [*t*(28)= 6.71, *p* < 0.001]; and for [ʊ] [*t*(28)= 5.78, *p* < 0.001]. Less but still significant differences in F2 between the two groups were also found for [ɒ] [*t*(28)= 2.38, *p* < 0.05] and for [ʌ] [*t*(28)= 2.19, *p* < 0.05]. Only [i] and [ɛ] were found to have no significant differences in both vowel height and advancement.

In order to provide a clear view of differences in vowel qualities between the two groups, the average, log mean transformed F1 and F2 values of the English vowels, the Mandarin-accented vowels, and the Mandarin vowels were plotted in Figure 4-2. There was no evidence of spectral differences between the English, Mandarin-accented and Mandarin [i]. There was also unexpected spectral similarity between the English and Mandarin-accented [ɛ]. However, as seen in Figure 4-2, there were observable phonetic distances between English, Mandarin-accented and Mandarin [oʊ], [eɪ], [ɒ] and [ɑ]. Similarly, differences among English and Mandarin-accented [æ], [ʌ], and [ɒ] (vowel height) were also obvious. Furthermore, Figure 4-2 depicts the dramatic distance between the English and Mandarin-accented [ɪ] and [ʊ] in both vowel height and advancement as the results of the *t*-tests showed. The Mandarin-

accented [ɪ] was [i]-like, as it was both higher and more anterior than the English [ɪ]. It lay between the native English [i] and [ɪ] and in fact was much closer to [i] than [ɪ]. The Mandarin-accented [u] was higher and considerably more posterior than the English [u]. It fell between the English [u] and the Mandarin-accented [u] in the vowel space in Figure 4-2. The difference in F2 between the Mandarin-accented [u] the native English [u] was also dramatic. As expected, the Mandarin-accented [u] was more posterior than the English [u], much closer to the Mandarin [u] in F2. It fell between the English and Mandarin [u].

The lack of spectral differences between the native English and Mandarin-accented [ɛ] was surprising, as the accented [ɛ] was significantly less well identified than the [ɛ] produced by the native English speakers in Experiment 1. There might be factors that affected the intelligibility scores for the front vowels that were not reflected by the overall F1 and F2 values of the vowels themselves. For example, the formant movement patterns could have played a role in the listeners' perceptions. In addition, between-group differences in vowel spacing or distribution between the two neighboring vowels, [ɛ] and [æ], could also affect perception. As formant movement information was not included in the present study, no data were available for further analysis. In order to explore the other possible factor (i.e. differences in vowel distribution patterns for the front vowels between the two groups), the data from English and Mandarin-accented front vowels were all submitted to statistical tests. Separate one-way ANOVAs were carried out to explore the differences in both F1 and F2 of the front vowels. For the English front vowels [i ɪ eɪ ɛ æ] produced by the native English speakers only, the results showed an effect of F1 [ $F(4) = 210.74, p <$

0.0001], as well as an effect of F2 [ $F(4) = 158.72, p < 0.0001$ ]. Post hoc Tukey (HSD) tests revealed that both F1 and F2 values for the front vowels were significantly different from each other with the exception of [eɪ] (measurement 'a') and [ɪ]. The results indicated that native English front vowels were all significantly separated in F1 and F2 except [eɪ] (at measurement 'a') and [ɪ].

Parallel one-way separate ANOVAs were also carried out to examine the Mandarin-accented English front vowels. The results showed an effect of both F1 [ $F(4) = 108.15, p < 0.0001$ ] and F2 [ $F(4) = 44.20, p < 0.0001$ ] as well. For F1 values, a post hoc Tukey test revealed similar results to those for the native English vowels. However, the Tukey (HSD) tests for F2 values revealed a different result on the Mandarin-accented front vowels. No significant differences were found between the [æ]-[ɛ], and [ɪ]-[i] pairs in the front-back dimension. These two pairs of front vowels produced by the native Mandarin speakers were frequently confused in the intelligibility test in Experiment 1. An F1 and F2 plot of the four front vowels [i, ɛ, æ] by Mandarin speakers and by native English speakers is given in Figure 4-3. ([eɪ] was not included in the plot because of formant movement). Obvious spectral overlaps can be seen between [i] and [ɪ], [ɛ] and [æ] produced by the native Mandarin speakers in Figure 4-3. Much less spectral overlap was found for the vowel tokens produced by the native English speakers.

Similarly, as seen in Figure 4-2, obvious distribution differences were found between the English and Mandarin-accented [u]-[ʊ], and [ou]-[ɒ] pairs. Although the Mandarin-accented [ou] and [ɒ] were relatively close to the corresponding native English [ou] and [ɒ] in both F1 and F2 as seen in Figure 4-2, the accented [ou] and [ɒ] were crowded together,

while the English [ou] and [ɒ] were further apart in F1. In fact, the Mandarin-accented [ou] and [ɒ] fell exactly between the native English [ou] and [ɒ]. This difference in distribution may well be the cause of the Mandarin-accented [ou] and [ɒ] confusion in the intelligibility test in Experiment 1. Furthermore, the native English [u] and [ou] showed much more separation in F2 than the accented [u] and [ou]. The F1 and F2 plots of the back vowel pairs [u]-[ɒ], and [ou]-[ɒ] produced by Mandarin speakers and by native English speakers are given in Figure 4-4 and Figure 4-5 respectively. While the English [ɒ] and [ʌ] were closer in distribution than the Mandarin-accented [ɒ] and [ʌ] in terms of vowel height, the result was the reverse for their advancement. The native English [ɒ] and [ʌ] were well separated in F2 as reflected in Figure 4-2.

#### 4.2.2.3. Discussion

The comparison of vowel durations of Mandarin-accented and native English vowels indicated that the Mandarin-accented [i] was much longer than the native English speakers' [i] and thus contributed to a larger [i]-[ɪ] duration ratio for the [i]-[ɪ] tense-lax pair. The fact that the Mandarin speakers exaggerated the duration differences in this tense-lax vowel pair, without apparently providing clear spectral differences, was found to have a perceptual basis in past studies. In a perception test on the same English front [i]-[ɪ] tense-lax pair using synthesized vowel tokens, Bohn (1995) reported that native Mandarin speakers relied exclusively on duration cues in differentiating English vowel pair [i]-[ɪ]. Bohn concluded that the Mandarin speakers relied on duration cues for the perception of this vowel pair because their previous linguistic experience did not sensitize them to the spectral

differences. The listeners' L1 vowel system has only one vowel [i] in the vowel space where English has both [i] and [ɪ]. No parallel production test was conducted in that study. The production data in the present study indicated that the Mandarin speakers really did exaggerate the duration differences between the [i] and [ɪ]. Their reliance on the duration cue appeared to be connected with their inability to differentiate this vowel pair spectrally. The strategy used by some Mandarin speakers may have been to produce the L1 vowel [i] spectrally and simply shorten it to make it the English [ɪ] rather than to establish a spectrally different vowel category in that area of the vowel space.

It is interesting to note that no significant duration differences were found between the Mandarin-accented tense-lax back vowel pair [u] and [ʊ]. Mandarin speakers did not seem to use the same strategy in differentiating this pair, but for the most part, simply substituted it with the neighboring [u] and [oʊ]. The different strategies the Mandarin speakers employed in their productions of the tense and lax vowel pairs indicated that their reliance on duration cues seemed to be limited to the [i] and [ɪ] pair only.

The spectral comparisons between the English and Mandarin-accented vowels indicated that the majority of the Mandarin-accented vowels were significantly different from the native English ones when the absolute mean F1 and F2 values were compared. Of the 10 vowels, only [i] and [ɛ] were not significantly different spectrally. The data suggested that with the exception of [i], in general, the Mandarin speakers had not learned to produce most of the English vowels in a perfectly native-like way. Although some Mandarin-accented vowels,

including [eɪ ou] were almost as intelligible as those produced by the native English speakers, they deviated significantly in F1 measurement (a) from the native English ones. The Mandarin-accented [eɪ ou] approximated the native English [eɪ ou] to some extent as shown in Figure 4-2. Yet, the differences between the native and Mandarin-accented and English [eɪ ou] remained significant at measurement (a) in F1. This finding seems to support Flege's Speech Learning Model which states that equivalence classification will block L1 speakers from establishing categories for some similar L2 phones. The same situation also applied to the Mandarin-accented [u] which was spectrally different from the English [u] in F2. The Mandarin-accented [u] was spectrally close to the Mandarin [u] but showed a tendency toward slight approximation to the native English [u].

Significant spectral differences were found for both [ʌ] and for [ɒ] in F2 across groups. As the Mandarin [ɑ] is relatively closer to the English [ʌ] than to [ɒ] spectrally, the presence of Mandarin [ɑ] may affect Mandarin speakers' productions of the English vowel [ʌ]. This prediction was supported by the fact that a spectral overlap was found between Mandarin [ɑ] and Mandarin-accented [ʌ]. It is likely that Mandarin speakers made use of their L1 [ɑ] in their production of L2 [ʌ].

On the other hand, the Mandarin speakers did not seem to have established new categories for the English vowels that do not have apparent L1 counterparts. The accented English [æ], [ʊ], and [ɪ] all proved to be significantly different from the native English vowels spectrally. These findings appear to contradict Flege's Speech Learning Model in that the greater phonetic distance seen between the L1 and L2 vowels did not make it easier for these Mandarin speakers' to learn to produce

• these "new" vowels more successfully than the similar vowels that are spectrally closer to the L1 vowels. However, the influence of L1 vowels on the production of L2 vowels was obvious. For example, the accented [ɪ] and [ʊ] were both "pulled" toward the Mandarin [i] and [u], though both vowels showed some approximation toward the English [ɪ] and [ʊ] in vowel height and advancement as seen in Figure 4-2.

The overall spectral comparisons indeed showed differences between the native English and Mandarin-accented vowels in general. In addition, there were obvious and consistent differences in vowel distributions especially for front vowels between the two speaking groups. While the native English front vowels were well separated in both vowel height and advancement, the Mandarin-accented ones showed spectral overlap between [æ]-[ɛ], and [ɪ]-[i] pairs in F2 values. These spectral overlaps between the vowels in an area that is already inherently high in density may be the cause of the poor intelligibility test scores for some of the Mandarin-accented vowels. A good example is the front vowel [ɛ]. The mean F1 and F2 values of this vowel for the two groups were almost identical, and the spectral similarity was obvious in the vowel figure. However, the Mandarin [ɛ] and [æ] overlapped so much in F2 that the two vowels had no front-back differences. For the English [ɛ] and [æ], no such F2 overlap was found. Similarly, distribution differences between the two groups were also found for the back vowels [ʊ]-[u], and [oʊ]-[ɒ]. All these differences in overlapping distributions of vowels affected the intelligibility of the Mandarin-accented vowels. The relationship between the acoustic properties and intelligibility will be discussed in the next chapter.

#### 4.2.2.4. Individual Data

The comparisons of acoustic measurements indicated that there were temporal, spectral, and vowel distribution differences between the English and Mandarin-accented vowels. To what extent did these differences affect the intelligibility of the Mandarin speakers' English vowel productions that were tested in Experiment 1? The average correct identification scores do reflect the acoustic differences revealed in Experiment 3. On average, [u] and [ɪ] were the least well identified back and front vowels respectively in Experiment 1. Similarly, both vowels deviated acoustically from the native norm as shown in Experiment 3. The correlation between the poor identification and acoustic mismatch was obvious when the average numbers were compared. However, the use of average scores may sometimes conceal important differences in individual performance.

To investigate this potential problem, the individual speakers' identification scores and the results of the acoustic measurements of [u] and [ɪ], the two least intelligible vowels in Experiment 1, were examined in detail. First, all the individual speakers whose productions of [u] and [ɪ] were 0% correctly identified in Experiment 1 were selected for an examination of their acoustic properties. Three Mandarin speakers' [ɪ]s and six speakers' [u]s had 0% correct identification scores. The F1 and F2 frequency values of [u] and [ɪ] produced by these speakers were compared first with the mean F1 and F2 values of the Mandarin-accented [u] and [ɪ] and then with the mean of the native English [u] and [ɪ]. It was expected that those vowel tokens would have F1 and F2 frequency values that deviated considerably from the average values. It was also expected that the [ɪ]s that were heard as [i]s would have below

average F1 and above average F2 values that were more [i]-like in acoustic terms.

The expected differences in vowel spectrum were all confirmed. For example, all the misidentified [ɪ]s produced by the three individual speakers were heard as [i]s. Just as expected, all these targeted [ɪ]s had [i]-like F1 values that were lower than not only the average of the native English speakers' [ɪ]s but also the Mandarin-accented [ɪ]s. All their [ɪ]s had above average F2 values. The deviant [ɪ]s produced by these three speakers fell within the range of the English [i]s acoustically and were all heard as [i]s. Similar results were also found for the misidentified [ʊ]s. The [ʊ]s that had below average F1 values were heard mostly as [u]s. Those that had above average F1 values were heard as [oʊ]s. The deviation of both F1 and F2 values in the Mandarin-accented [ʊ]s from the native English [ʊ] could be found across the board. This was also expected because the average correct identification rate for [ʊ] was only 15%.

The correlation between the low identification scores and spectral deviations in [ɪ] and [ʊ] demonstrates that the misidentification of these tokens had an acoustic basis that could be readily described in terms of F1 and F2 frequencies. The individual speaker differences observed in the intelligibility test and the acoustic measurement study proved to match.

### 4.3. CONCLUSION

The analysis of acoustic data presented here indicated that the majority of the Mandarin-accented vowels were significantly different

from the native English ones in absolute mean F1 or F2 values. In general, the Mandarin speakers in this study had not learned to produce most of the English vowels in a native-like manner. The acoustic differences between the Mandarin-accented and native English vowels are explicable in terms of similarities and differences between the Mandarin and English vowel systems. The majority of the Mandarin speakers did not seem to have established separate categories for the English vowels in their productions. Rather, the results suggested that some speakers substituted similar Mandarin vowels such as [i] and [u] for the English counterparts in their productions. As shown in Figure 4-2, the Mandarin-accented vowels tended to be "pulled" toward the Mandarin vowels spectrally. However, most accented vowels, especially the ones that lack Mandarin counterparts, showed some degree of approximation towards the native English vowel targets in terms of acoustic properties. Therefore, not every Mandarin speaker substituted the new English vowels with the closest vowels in the native stock.

The other findings from the comparison of acoustic properties were the obvious and consistent differences in vowel distributions, especially for front vowels, between the two speaking groups. The Mandarin-accented tense-lax vowel pairs demonstrated significant spectral overlap in F2, which caused confusion in the identification task.

The Mandarin speakers' reliance on duration cues for differentiation of the tense-lax vowel pairs was found to be related to [i] and [ɪ] only in production. This was because the Mandarin-accented [i] was much longer in duration than [ɪ]. It is possible that some Mandarin speakers employed more than one strategy to distinguish the English

tense and lax vowels. More studies with carefully controlled synthetic vowel tokens are needed to further explore the issue.

With regard to Flege's Speech Learning Model, the results of the present study supported the prediction that "equivalence classification" will prevent L2 learners from establishing a new category for "similar" phones. In general, the Mandarin speakers did not produce the English [eɪ ou u] in terms of acoustic properties in a native-like way. The spectral difference between the Mandarin-accented and native English [u] was especially clear. However, data from the acoustic analysis showed that the majority of "new" vowels were actually less well produced in terms of acoustic properties (formant frequency values). At least in the present study, acoustic dissimilarity between L1 and L2 vowels did not appear to facilitate the Mandarin speakers in their production of these "new" vowels as the SLM would predict. It is important to point out that there may be a difference between the experienced and inexperienced L2 speakers in the acquisition of "new" categories. The results of previous studies showed that experienced German English learners, not the inexperienced, were more successful in identifying and producing the English /æ/, a "new" vowel for German speakers (Bohn & Flege, 1990, 1992). Further studies with more controlled L2 experience factors are definitely needed.

**TABLE 4-1**

Mean durations (msec) and standard deviations of isolated English vowels [i ei u ou ɒ ʌ] and Mandarin vowels [i ei u ou ɑ]

Vowel	English	Std.	Mandarin	SD
i	270*	(55)	224*	(45)
ei	275*	(51)	211*	(118)
u	261	(52)	241	(43)
ou	258	(52)	229	(40)
ɒ /ɑ	243	(55)	204	(40)
ʌ/ɑ	181	(53)	204	(40)

\* Indicates that the differences between Mandarin and English vowels were significant at  $p < 0.05$

TABLE 4-2

Mean normalized F1 and F2 frequencies (Hz) and standard deviations of English [ɪ u ʌ eɪ ou] and Mandarin [ɪ u a ei ou]. (The values are based on measurements at 30% of the distance into the vowel and 70% as specified for isolated vowels.)

Vowel	F1 (Hz)		F2 (Hz)		t	Mandarin	English	t	Mandarin	English	t
	English	Mandarin	English	Mandarin							
ɪ	323 (32)	304 (32)	2643 (174)	2663 (295)	1.64						0.24
u	352 (32)	362 (57)	1155 (147)	754 (98)	0.58						8.79**
ɒ / ɑ	725 (65)	832 (78)	1144 (73)	1324 (94)	4.07**						5.85**
ʌ / ɑ	752 (68)	832 (78)	1412 (76)	1324 (94)	3.00**						2.82**
eɪ at 30%	481 (42)	530 (63)	2285 (145)	2090 (201)	2.48*						3.04**
eɪ at 70%	393 (55)	435 (71)	2456 (177)	2464 (229)	1.80						0.11
ou at 30%	527 (29)	561 (45)	1019 (73)	973 (64)	2.48*						1.82
ou at 70%	445 (47)	486 (42)	924 (89)	844 (75)	2.50*						2.65*

\* Indicates that the differences in frequencies were significant at  $p < 0.05$ ; \*\* at  $p < 0.01$

**TABLE 4-3**

Mean durations and standard deviations for English [i ɪ eɪ ε æ u ʊ ou ɒ ʌ] produced by English and Mandarin speakers

Vowel	English	Std.	M-accented	SD
i	270*	(55)	322*	(87)
ɪ	177	(51)	165	(75)
eɪ	275	(51)	287	(88)
ε	188	(49)	185	(88)
æ	248	(52)	222	(69)
u	261	(52)	291	(89)
ʊ	199	(49)	232	(93)
ou	258	(52)	299	(76)
ɒ	243	(55)	196	(98)
ʌ	181	(53)	165	(80)

\* Indicates that the duration differences between the English and M-accented vowels were significant at  $p < 0.05$

TABLE 4-4

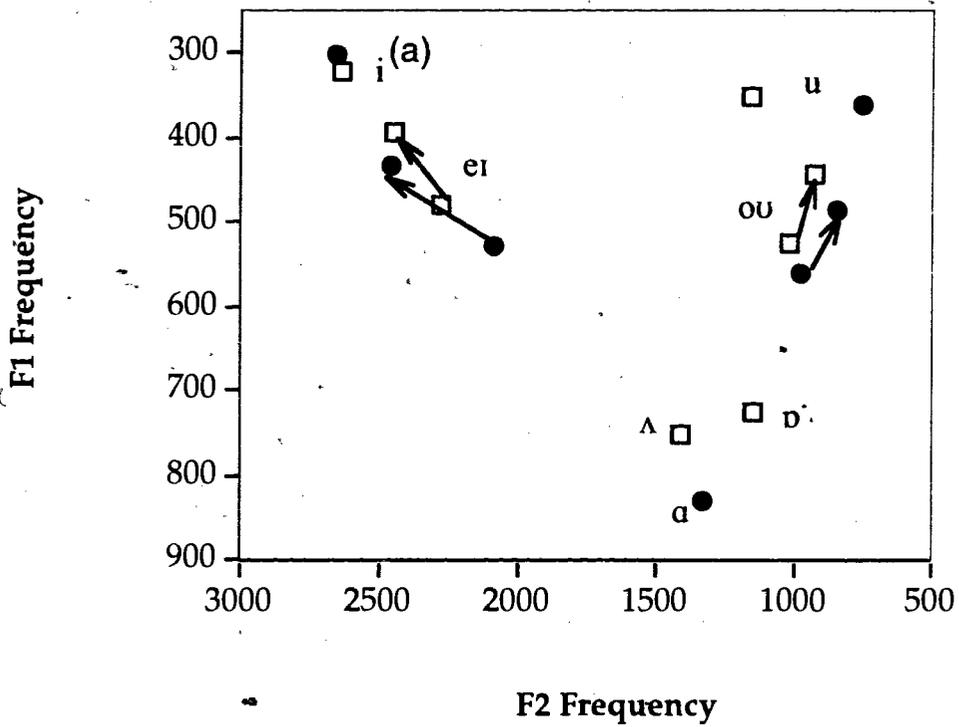
Mean normalized F1 and F2 frequencies (Hz) and standard deviations for English [i eɪ ε æ u ou ʊ ʌ] produced by both groups. (The values are based on measurement (a) at 30% for most vowels and at measurement (a) and (b) 70% for the two diphthongs [eɪ ou].)

Vowel	English		M- accented		English		M- accented		t
	F1	F2	F1	F2	F1	F2	F1	F2	
i	323 (32)	313 (36)	313 (36)	0.82	2645 (204)	2642 (244)	2642 (244)	0.04	
ɪ	467 (43)	408 (85)	408 (85)	2.36*	2187 (169)	2495 (272)	2495 (272)	3.73***	
eɪ at 30 %	482 (44)	548 (83)	548 (83)	2.75*	2285 (137)	2235 (192)	2235 (192)	0.82	
eɪ at 70 %	393 (55)	417 (83)	417 (83)	0.93	2457 (178)	2538 (257)	2538 (257)	1.01	
ε	665 (44)	671 (44)	671 (44)	0.29	1985 (133)	2024 (200)	2024 (200)	0.63	
æ	832 (82)	794 (75)	794 (75)	1.32	1688 (102)	1923 (301)	1923 (301)	2.86**	
u	352 (32)	369 (43)	369 (43)	1.16	1156 (157)	822 (113)	822 (113)	6.71***	
ʊ	529 (64)	457 (101)	457 (101)	2.35*	1328 (168)	971 (170)	971 (170)	5.78***	
ou at 30 %	527 (35)	581 (64)	581 (64)	2.89**	1019 (80)	1040 (91)	1040 (91)	0.66	
ou at 70 %	445 (46)	481 (75)	481 (75)	1.59	924 (90)	938 (109)	938 (109)	0.38	
ɑ	725 (59)	677 (77)	677 (77)	1.91	1145 (75)	1072 (92)	1072 (92)	2.38*	
ʌ	752 (68)	831 (121)	831 (121)	2.21*	1412 (76)	1322 (140)	1322 (140)	2.19*	

\* Indicates that the differences in frequencies were at  $p < 0.05$ , \*\* at  $p < 0.01$ , \*\*\* at  $p < 0.001$

FIGURE 4-1

Mean F1 and F2 (normalized values) of English [i eɪ u ou ɒ ʌ] produced by 15 English speakers and Mandarin [i eɪ u ou a] produced by 15 Mandarin speakers.

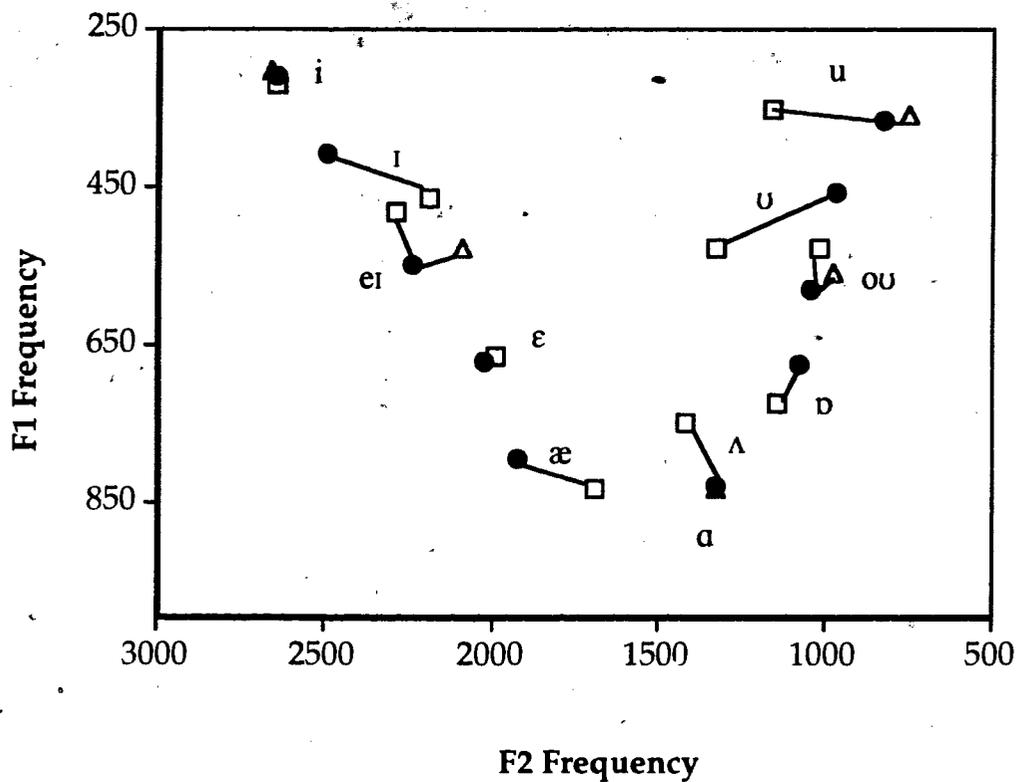


□ English

● Mandarin

FIGURE 4-2

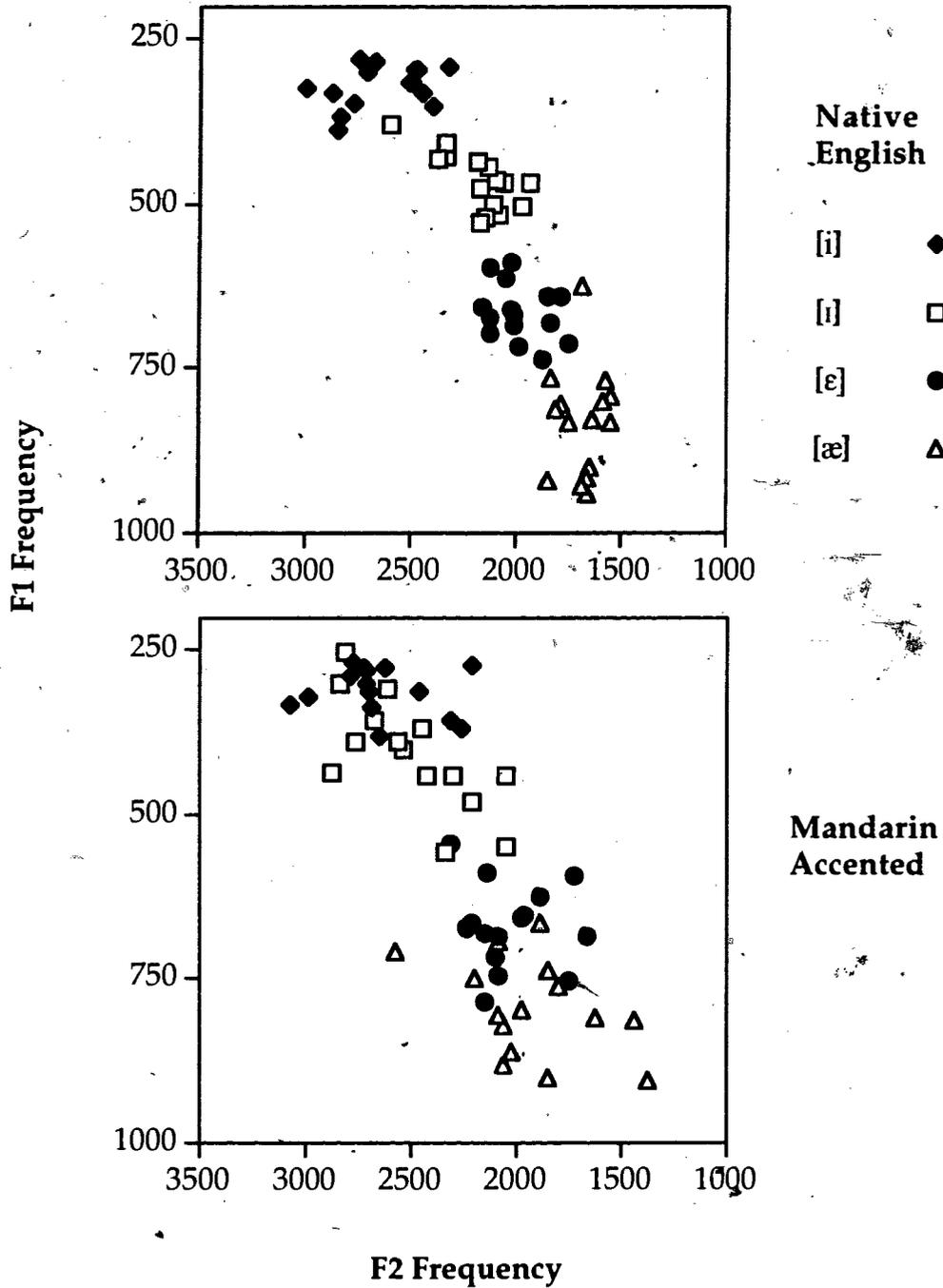
Mean F1 and F2 (normalized values) of English [i ɪ eɪ ε æ u ʊ ou ɒ ʌ] by English and Mandarin speakers and Mandarin [i eɪ u ou ɑ] by Mandarin speakers.



- M-accented
- English
- △ Mandarin

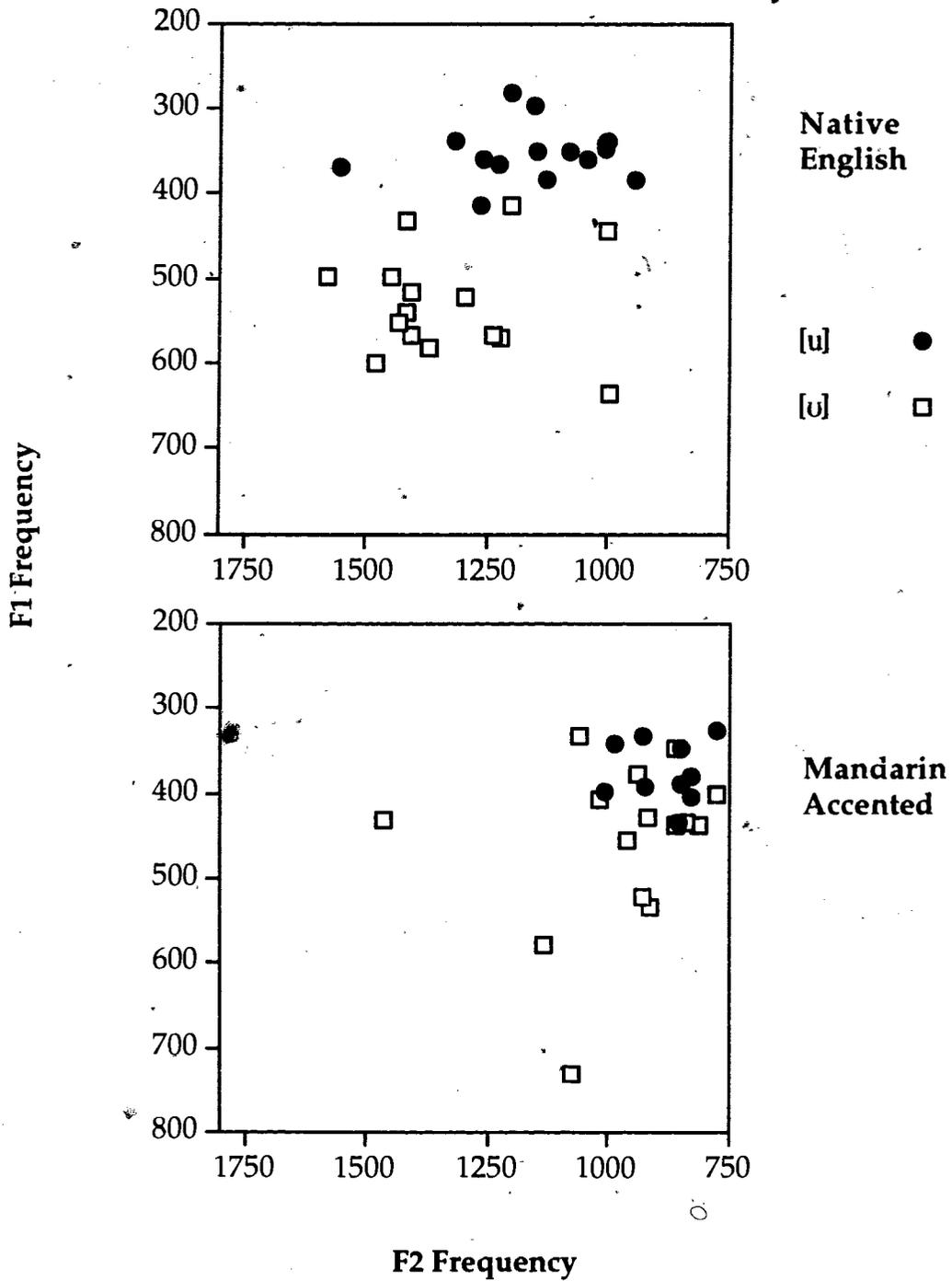
**FIGURE 4-3**

F1 & F2 (normalized values) plot of English [i i ε æ] produced by 15 English and 15 Mandarin speakers



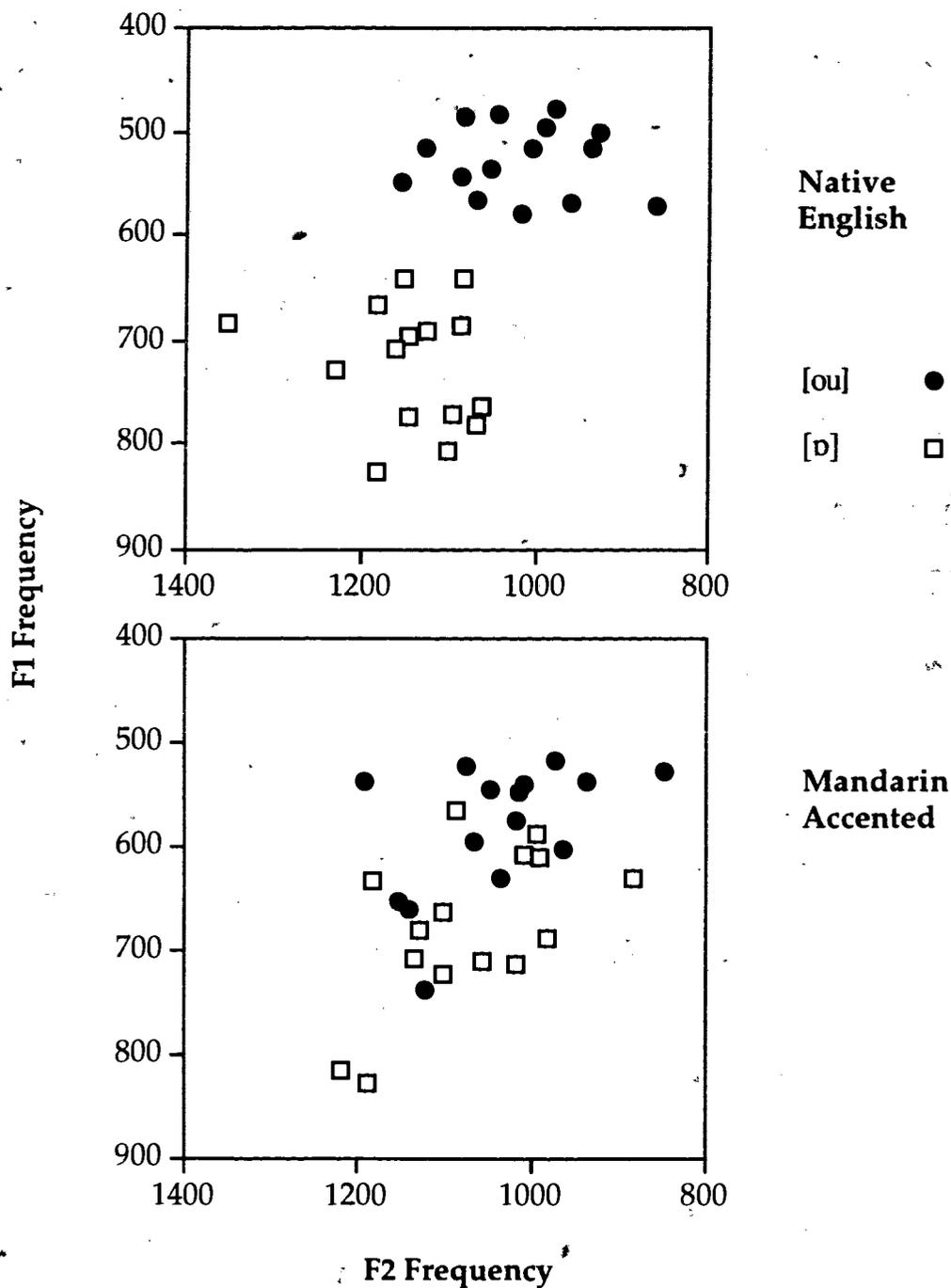
**FIGURE 4-4**

F1 & F2 (normalized values) plot of [u u] produced by 15 English speakers (top) and by 15 Mandarin speakers (bottom)



**FIGURE 4-5**

F1 & F2 (normalized values) of English [ou] at measurement (a) and [ɒ] produced by English and Mandarin speakers



## CHAPTER V

### GENERAL DISCUSSION

In this chapter, the relationship between the three experiments regarding the perception and production of English vowels, and the acoustic properties of English, Mandarin, and Mandarin-accented vowels will be examined. A key issue to be explored is whether the acoustic properties can account for the success or failure of the Mandarin speakers' performance in their production and perception of the English vowels.

The general tendency observed in Experiment 1 was that Mandarin speakers' productions of English vowels having Mandarin counterparts were significantly more intelligible than those lacking Mandarin analogs. This appeared to have an acoustic basis as indicated by the results of Experiment 3. Even though the majority of Mandarin speakers had not learned to produce the English vowels in a native-like manner, in terms of acoustic properties as shown in Experiment 3, their productions of "similar" vowels were almost as intelligible as the native English ones. (See Table 2-4 and Table 2-5.) This was because, in general, the Mandarin vowels that have English counterparts were acoustically similar to the corresponding English vowels and the Mandarin-accented vowels, to some degree, approximated the English vowel targets.

The close spectral resemblance among the English, Mandarin, and Mandarin-accented [i]s was especially obvious as reflected in Figure 4-2. In fact, the Mandarin accented [i] was correctly identified 96% of the time by the native English listeners. The success in the production of

this almost "identical" vowel can be explained by the fact that not much, if any, learning of [i] was necessary. Yet, it is interesting to note that the Mandarin-accented [i] was significantly longer than the native English [i], even though the native Mandarin [i] was significantly shorter than the English [i] in this study. The Mandarin speakers' exaggeration of the English [i] in duration was apparently a strategy used to differentiate the English [i]-[ɪ] pair. The speakers seemed to have difficulty separating the two vowels spectrally.

The unexpected complete spectral overlap of Mandarin-accented English [ʌ] and Mandarin [ɑ] observed in Experiment 3 provides an explanation for the high identification scores for the Mandarin-accented [ʌ] (76%). It is possible that some Mandarin speakers simply substituted the L1 [ɑ] in their production of the English [ʌ]. Unlike [i], which is spectrally identical to the English [i], the Mandarin [ɑ] and English [ʌ] were not identical spectrally. However, the cross-language comparison showed that the phonetic distance between Mandarin [ɑ] and English [ʌ] was smaller than the distance between the commonly assumed [ɑ] and [ɒ] counterparts.

Similarly, although Mandarin, Mandarin-accented, and English [eɪ] and [oʊ] were not identical, they were close spectrally. The correct identification score for the Mandarin-accented [eɪ] was identical to that for the native English [eɪ] (96% vs. 96%) and for [oʊ] it was very close (95% vs. 99%).

The Mandarin speakers' success in their production of the English vowels [i eɪ oʊ ʌ] in terms of intelligibility did not differ significantly from that of the native English speakers. The consistency between the acoustic properties and high intelligibility scores of Mandarin-accented

vowels suggested that the Mandarin speakers had either made use of their L1 vowels in their production of these English counterparts (like [i] and [ʌ]) or approximated the target English vowels (like [eɪ] and [oʊ]).

Discrepancies between the intelligibility scores and the acoustic properties of the Mandarin-accented vowels were also observed. A good example was the case of [u]. The highly intelligible Mandarin-accented [u] which had a correct identification rate of 97% was dramatically different from the native English speakers' [u] in F2. The results of the cross-language acoustic comparison indicated that the Mandarin-accented [u] was much more posterior than the native English [u]. Even though they are analogous vowels, represented by the same phonetic symbol, the data indicated that the Canadian English [u] tends to be much less posterior (mean F2 = 1156 Hz) than a typical high back rounded Mandarin [u] (mean F2 = 754 Hz) as indicated by the data in this study. Similar acoustic properties for the Canadian English [u] (mean F2 = 1174 Hz) were found in a study of 10 native Canadian English speakers (five male and five female) by Nearey and Assmann (1986). The spectral differences characterizing Mandarin-accented English [u] did not affect its intelligibility in Experiment 1. The high rate of correct identification for the Mandarin-accented English [u] can basically be explained by the English vowel system. Since [u] is the highest back vowel in English, anything that is higher or more posterior and is [u]-like in this area of the vowel space is likely to be identified as [u] in an identification test like this. Therefore, the Mandarin accented [u] that was similar in vowel height but further back than the Canadian English [u] was consistently identified as [u]. That the phonetic distance between the native and accented [u] did not affect the native English listeners'

identification scores is perhaps a special case for the identification task. An accent-rating task might yield a different result. The accented [u] was obviously influenced by the Mandarin [u] and was spectrally very close to the Mandarin [u]. The majority of individual speakers may have substituted the Mandarin [u] for the English counterpart.

To summarize, the high identification scores for the Mandarin speakers' productions of English vowels that have English counterparts had an acoustic basis. The phonetic resemblance (with the exception of [u]) seemed to be the main reason for the listeners' success in identifying these vowels.

The relationship between the phonetic properties and intelligibility scores was more complicated for Mandarin-accented vowels that do not have obvious Mandarin analogs. First, as expected, the deviation from the native norms in acoustic properties of the Mandarin-accented vowels appeared to be the major cause of the low identification scores. The Mandarin-accented [u] is a typical example. As can be seen in Figure 4-2, the Mandarin-accented [u] was both higher and more posterior than the English [u]. It lay between the English [u] and [ou], and was spectrally closer to the English [u] and [ou] than to [u]. Most of the misidentified Mandarin-accented [u]s were, therefore, heard as [u]s and [ou]s. Although [u] was the least intelligible of the ten vowels produced by the native Mandarin speakers, acoustically, it showed some approximation toward the target English [u] in both F1 and F2 values (See Figure 4-2). The approximation, however, did not appear to be sufficient to warrant a distinction between a clear [u] and [u] to the native English listeners. The intended [u]s that were higher and more posterior would be heard naturally as [u]s and those that were lower

and more posterior as [ou]s. As analyzed earlier in the individual data for this particular vowel, there were obvious cases of substitution of the L1 [u] and [ou] for the L2 [ʊ]. In fact, in this study, the majority of Mandarin speakers failed to produce an intelligible [ʊ].

Another poorly-identified Mandarin-accented vowel was [ɪ], which deviated significantly from the native English [ɪ]. The accented [ɪ] was spectrally very close to the neighboring vowel [i] both in height and advancement. It fell between the English [i] and [ɪ] and was closer to [i] than to [ɪ]. The attempted [ɪ]s tended to be heard as [i]s. In the preceding chapter, the totally unintelligible [ɪ]s produced by three Mandarin speakers all proved to be more [i]-like spectrally. The tendency toward approximation of the English [ɪ] was also obvious for some speakers, as shown in Figure 4-3. Three speakers' [ɪ]s had a high intelligibility score of above 80%. These speakers appeared to have learned to produce this vowel successfully.

Another factor affecting the intelligibility of the accented vowels was the difference in vowel distribution or spacing. As discussed in the previous chapter, in terms of absolute average F1 and F2 values, no significant difference was discovered between the native English and Mandarin-accented [ɛ]. However, the spectral overlap between the Mandarin-accented [ɛ] and [æ] was obvious, as shown in Figure 4-3. No such overlap of F2 was found between the native English pair [ɛ] and [æ]. The overlapping in distribution of the accented [ɛ] and [æ] pair in F2 could be the major cause for the misidentification of [ɛ] for [æ], and [æ] for [ɛ]. The similar overlapping distribution between the other accented front vowel pair [i] and [ɪ] showed a parallel tendency toward similar F2 values for the two vowels. The Mandarin speakers' efforts in

approximating the target vowels [ɪ ε æ] were noticeable. Yet, the density of the vowel category distribution in this particular area may require well-defined boundaries for intelligibility. The non-native speakers, whose L1 lacks these vowels in this area of the vowel space, were less sensitive to the spectral differences between these pairs. The density of the vowel distribution in the L2 system increased the level of difficulty for the learners. They were not sensitive to the spectral differences between these "new" vowels because of their L1 experience. For example, Mandarin speakers will probably find it much easier to produce either [ɛ] or [æ] than to produce both of them.

The same situation may apply in the case of the back vowel [ɒ] that was mostly misidentified as the neighboring [oʊ] and [ʌ]. The spectral overlap between the accented [oʊ] and [ɒ] was clearly reflected in Figure 4-5. Another reason for the misidentified [ɒ] could be that some Mandarin speakers, before they came to Canada, might have been exposed to Received Pronunciation, a dialect that distinguishes [ɔ] from [ɑ]. The effect could be driving the intended [ɒ] up to [ɔ] which is closer to [oʊ] in vowel height. The experimenter, who is herself a native Mandarin speaker and has learned the RP accent in China, had the impression of hearing [ɔ] for some of the [ɒ] tokens which were analyzed.

The comparisons among the three experiments shed some light on many of the issues that were discussed in the previous chapters. The discrepancies between the Mandarin speakers' performance in the production and perception of the English vowel [eɪ] remained unsolved when the acoustic properties of their [eɪ] were examined in Experiment 3. Spectrally, the Mandarin [eɪ] was found to be lower and more

posterior than the English [eɪ] at measurement (a). The Mandarin-accented [eɪ] was closer to the English [eɪ] in advancement but was also lower. If there is a correlation between the high intelligibility scores and near native acoustic properties of the vowel [eɪ], then why did the Mandarin speakers have difficulty in identifying native English speakers' productions? This naturally brings up the question about the complex relationship between perception and production.

Before exploring the correlation between perception and production in terms of acoustic properties, it is important to point out the limitations of such an attempt. First, the perception study conducted in Experiment 2 was limited to only the five front vowels. Second, as mentioned earlier, the tokens used in the two studies were not exactly the same. In spite of these problems, it is still worthwhile to consider some of the issues:

The first question to be asked is whether the Mandarin speakers' identifications of the five English front vowels were related to their own productions of these vowels. For example, if a particular Mandarin speaker had a problem in producing the distinction between the vowel pair [i] and [ɪ] in Experiment 1, was there a parallel problem in the perception of this vowel pair? To answer this question, the individual speaker's scores for [ɪ] in Experiment 1 and Experiment 2 were compared. Three of the Mandarin speakers' productions of [ɪ] were 0% correctly identified in Experiment 1. However, two of these speakers had perception scores that were above average for the Mandarin listeners in Experiment 2. The other listener had a perception score that was below average, but was not the lowest among the 15 listeners. Similarly, the listener who had the highest perception score for [ɪ] (93%)

among the 15 listeners had produced this same vowel less successfully. The intelligibility score for his production of [ɪ] was below average (45%). In fact, as discussed in Chapter 3, the Mandarin speakers generally performed better in their perception of [ɪ] than in their production. This implies that most speakers' perception appeared to precede production. The failure in the production of this vowel by some speakers indicated that it may take time to develop a perceptual representation for a certain sound. It is also possible that these speakers have actually established a perceptual category for the English [ɪ], but that it differs from that of native English speakers. The Mandarin speakers may have simply relied on the duration cue to differentiate the English [ɪ] from the neighboring [i] as indicated in the results of Experiment 3. If the latter prediction is true, then it is likely that Mandarin listeners may discriminate the English [ɪ] from the neighboring [i] on the basis of the temporal cues. On the other hand, the intended [ɪ] produced by Mandarin speakers that was [i]-like spectrally would naturally be heard as [i] by the native English listeners. The duration cues employed by the Mandarin speakers in their production would not make much difference to English speakers who rely mostly on spectral cues in the [i]-[ɪ] distinction.

A different picture was found for the other troublesome pair of front vowels, [ɛ] and [æ]. A correlation between scores in both perception and production was found for [æ]. For example, the three speakers who had perfect intelligibility scores on their productions of [æ] were the same three who had the highest perception scores for this vowel. The situation was different for [ɛ]. There was no such consistency in individual performance between perception and production. For

instance, the speaker who performed best in the perception of [ɛ] (91%) produced an [ɛ] that received a below average score (45%). However, even for [æ], a more careful comparison of the individual production scores with that of the neighboring [ɛ] actually complicated the situation. In fact, as shown by the vowel distribution pattern test in the preceding chapter, the spectral overlapping of the accented [ɛ] and [æ] affected the intelligibility of these two vowels. The confusion patterns for the misidentified [ɛ] and [æ] confirmed the trouble Mandarin speakers had in differentiating this vowel pair. There was clear evidence of certain speakers' extreme differences in success in the productions of these two vowels. For example, the speaker whose production score for [æ] was 0% had a high production score of 95% for [ɛ]. One third of the Mandarin speakers were found to have extremely high scores on one of the vowels and low scores on the other in their production of [ɛ] and [æ]. This suggested that the success in producing either [æ] or [ɛ] should be treated with caution as it is likely that the success of one may mean the failure of the other.

No such extreme differences were ever found for even one of the 15 Mandarin speakers in their perception of the same vowel pair [ɛ] and [æ] produced by the native English speakers. It is possible that some Mandarin speakers had noticed the acoustic differences in this pair of vowels in their perception of the native English speakers' productions but had not yet learned how to produce a good and clearly separated [æ] and [ɛ] pair. If this is the case, then, it appears to be easier for the Mandarin speakers to separate the vowels [æ] and [ɛ] in perception than in production.

Very few Mandarin speakers demonstrated success in production and perception of both [ɛ] and [æ]. In contrast, more speakers were found to be successful in perceiving or producing either [ɛ] or [æ]. It is possible for L2 learners to be able to distinguish two non-native phones first before they learn to produce them. However, it is not clear whether the Mandarin speakers used the same strategy for both their perception and production. Furthermore, it is not known whether their strategy was the same or different from that of the native English speakers. As the Mandarin speakers did not produce a clear duration difference between the [ɛ] and [æ] pair as they did with [i] and [ɪ], it is not likely that they relied on durational cues to produce the [ɛ] and [æ] contrast. More controlled perception studies with carefully designed synthetic vowel tokens are needed.

Returning to the problem of [eɪ], it is possible that the Mandarin speakers used strategies that are different from those of the native English speakers' in their perception and production of this vowel, just as they did with the other front vowel [ɪ]. However, their poor performance on [eɪ] could also be the result of other factors that should be investigated with care. As discussed in Chapter 3, the English [eɪ] has never been reported as a problem for Mandarin learners of English either in previous studies or in textbooks for pedagogical purposes. Another factor could be the distraction caused by the nature of the forced-choice listening task. The squeezing of three additional categories [ɪ ɛ æ] into the area of the vowel space where Mandarin contrasts only [i] and [eɪ] may have affected the Mandarin listeners' identification of [eɪ]. It would probably be much easier for them to identify only [i] and [eɪ] in the perception test. Therefore, the problem found in this study cannot

simply be treated as a case of intelligible production of a certain L2 vowel without an accurate perceptual basis. Obviously, more studies are needed in order to have a better understanding of the perceptual problem regarding [eɪ] which was found in the present study.

To summarize, in general, there is a relationship between the Mandarin speakers' degree of success and failure in their production of the 10 English vowels and the acoustic properties of these vowels. To some extent, the English vowel system, especially the vowel distribution pattern, explains some of the mismatches between the acoustic properties and the identification scores for some vowels. (e.g. the case of [u]). The relationship between perception and production of the English front vowels is not straightforward. There were obvious differences between the Mandarin speakers' performance in their perception and production of the five English front vowels. This was particularly evident for the front vowel [eɪ]. Examples of better perception than production were found with [ɪ] and [æ] for many speakers. Furthermore, the evidence suggested that many Mandarin speakers were able to differentiate the English [ɪ] from [i] in their perception, even though their production of this vowel was still problematic in terms of intelligibility.

## CHAPTER VI

### CONCLUSION

Until now there have been very few studies of Mandarin speakers' production and perception of English vowels. In fact, there have been practically no acoustic studies of Mandarin-accented English vowels produced by a relatively large group of speakers. Furthermore, no studies of Mandarin vowels have compared them with English vowels in terms of acoustic properties. This study provides important new data regarding these unexplored areas.

As discussed in Chapter 1, few studies of L2 vowels have controlled the dialectal factor as carefully as the present study. Ignoring this factor could distort the results observed in both perception and production. Because the subdialects of Mandarin Chinese are widely different, the importance of such dialectal control cannot be over-emphasized.

Blankenship (1991) pointed out that most previous cross-language vowel studies have compared no more than four vowels. Her study of L2 vowel perception included five English front vowels which were perceived by native Spanish speakers. Munro, Flege and MacKay (1996) looked at most of the English vowels, but they did not obtain measurements of native Italian vowels from the same speakers. The present study examined 10 English and Mandarin-accented vowels. These were included in both intelligibility tests and acoustic analyses. In addition, five Mandarin vowels produced by the same Mandarin speakers were included in the acoustic analyses.

One of the benefits of a study involving a large number of vowels is the scope it provides for examining cross-linguistic differences in vowel distributions and the effect on perception and production of L2 vowels. The results of the present study implied that it is more difficult to learn to perceive and produce a set of L2 vowels in a high density area of the vowel space. This is especially true for those learners whose L1 vowel system is relatively less crowded in this area. The results also show that non-native speakers who are 'desensitized' (See Bohn, 1995) to the densely distributed vowels in the whole area of the vowel space may sometimes successfully produce a certain vowel in this area, but may not have learned how to distinguish the whole set of vowels successfully. In the case of the English front vowels, the spectrally overlapping Mandarin-accented English front vowel pairs [i]-[ɪ] and [ɛ]-[æ] did not meet the dispersion requirement in F2 and were therefore relatively poorly discriminated in the identification task. For instance, the ambiguous [ɛ] and [æ] tokens produced by the Mandarin speakers were problematic for the native English listeners. It is important to point out, however, that the production test was conducted in an experimental setting with isolated vowels, a situation that differs significantly from the rich context typical of daily communication.

The production study aimed at the most basic and important issue of which vowels are problematic for Mandarin speakers in learning to speak English. The results indicated that Mandarin speakers' productions of English vowels that have Mandarin analogs were significantly more intelligible than those that lack Mandarin vowel counterparts. The acoustically deviant Mandarin-accented vowels had low intelligibility scores because of the influence of the L1 vowel

system. This was indicated by the acoustic measurements of the Mandarin vowels. The findings of the empirical studies can be understood in terms of the effect of L1 on L2 production. Complete transfer did not seem to happen to all the speakers or to all the vowels. Instead, the Mandarin speakers showed different degrees of approximation toward the target L2 vowels in their productions.

Previous studies concerning L2 vowel perception have mostly been carried out without involving production tests, or vice versa. The present study avoided this drawback by including both perception and production tests. With the limitation of five front vowel tokens in the perception test, only preliminary conclusions could be drawn. Nevertheless, certain tendencies were observed through the comparisons of data from the perception and production tests. For the English front vowels that lack Mandarin counterparts, [ɪ ɛ æ], there was a tendency for better perception than production. In contrast, for [i eɪ], vowels that have L1 analogs, the opposite tendency, better production than perception, was observed. The discrepancy between perception and production of the English front vowels by Mandarin speakers suggested that learners may have used different cues or strategies in their perception and production of the L2 vowels. Furthermore, the assumption that production had a perceptual basis was not fully supported by the results of the present study.

Sheldon & Strange (1982) and Blankenship (1991) reported that learners sometimes produced certain L2 consonants and vowels without being able to perceive them accurately. The results of those studies cannot be directly compared to Mandarin speakers' perception and production of the English [eɪ] observed in the present study. Both the

conditions and tasks were different. The identifications of five vowel categories, as opposed to two, may have complicated the situation by causing more distraction. No firm conclusion of possible accurate production without accurate perception of L2 segments can be drawn from the findings regarding [e<sub>1</sub>] in the present study. The relationship between perception and production of L2 vowels is a complex issue, and it is clear that further studies are needed.

Another drawback of some of the previous studies has been the tendency to overlook speaker variation. Averaging data across a group of speakers may conceal extreme individual differences. By examining the individual speaker variations in the results of each of the three experiments, such problems were avoided. As expected, the least intelligible vowels were found to be produced by those individual speakers whose vowel productions deviated the most from the native norms in terms of acoustic properties.

With respect to second language segmental learning, two theories have been examined. The traditional Contrastive Analysis Hypothesis (CAH) predicted the difficulties in learning L2 segments by comparing the inventories of the first and second language sound system. The results of the production study partially supported this theory. In general, the better identified Mandarin-accented vowels were those that had Mandarin counterparts. Vowels that did not have Mandarin analogs were significantly less intelligible to the native English listeners and more difficult for the Mandarin speakers to produce. This confirmed the fact that missing categories appeared to cause problems for learners.

The CAH failed, however, to explain some of the more subtle problems encountered in the data. First, the theory cannot explain the differences in the degrees of difficulty Mandarin speakers have in their production of vowels that lack L1 counterparts. Within the same group of 'missing categories', there are certain sounds that appear to be more difficult for certain L2 speakers. For example, the Contrastive Analysis Hypothesis does not predict that [ʊ] would be more difficult than [ɪ ɛ æ] for the Mandarin speakers to learn as this study showed. Second, by comparing only the individual vowel categories across languages and ignoring the differences in distributions of the vowel inventories between the two systems, the CAH failed to predict the potential problems for L2 speakers to learn a set of densely distributed L2 vowels in specific areas of the vowel space. For example, Mandarin speakers had difficulty in differentiating [ɛ]-[æ] and [i]-[ɪ] pairs. Another problem with CAH is that it cannot account for individual speaker variation in the perception and production of L2 sounds. Learner differences are an important factor that cannot be ignored in speech learning. The fact that certain L2 phones are often more difficult to learn for certain L2 speakers does not mean they are a problem for all or most speakers.

This study also reviewed Flege's Speech Learning Model (SLM). According to the SLM hypothesis, whether an L2 speaker will learn an L2 phone depends largely on the establishment of a phonetic category for that sound. An earlier version of the model stated that equivalence classification will prevent adult L2 learners from establishing a phonetic category for "similar" but not "new" L2 phones. The newer version of the SLM modifies the concept of new versus similar phones to more

general phonetic dissimilarities between the L1 and L2 phones. The greater the differences between the L1 and L2 sounds, the easier it will be for the speaker to perceive the differences and eventually to produce the L2 sound.

The acoustic data of similar vowels in the present study suggest that equivalence classification did seem to prevent the Mandarin speakers from producing spectrally native-like [u ou eɪ], despite the fact that the Mandarin-accented [u ou eɪ] were almost as intelligible as their native English vowel counterparts. The Mandarin-accented "new" vowels [ɪ ɛ æ ʊ] showed consistent deviation from the native norms in terms of acoustic properties, and were significantly less intelligible than their native English vowel counterparts. The results of the Mandarin speakers' productions of [ɪ ɛ æ ʊ] did not seem to support the SLM in that phonetic distance between L1 and L2 vowels did not at least facilitate the production of these L2 phones.

Finally, the findings of this study about native Mandarin speakers' English vowel perception, production and their acoustic properties suggested that the perception and production of L2 segments by adult learners is a highly complex issue. No single theory or hypothesis can fully account for the phenomena observed here. More systematic studies are needed to address the issues.

## Appendix 1

### English List

1.	Now I say beat	bea	ea	beat
2.	Now I say bit	bi	i	bit
3.	Now I say bait	bai	ai	bait
4.	Now I say bet	be	e	bet
5.	Now I say bat	ba	a	bat
6.	Now I say boot	boo	oo	boot
7.	Now I say buwt (like book)	buw	uw	buwt
8.	Now I say boat	boa	oa	boat
9.	Now I say bot (like pot)	bo	o	bot
10.	Now I say but	bu	u	but
11.	Now I say Bert	ber	er	Bert
12.	Now I say bite	bi	i	bite
13.	Now I say bout	bou	ou	bout
14.	Now I say book	boo	oo	book
15.	Now I say beat	bea	ea	beat
16.	Now I say bit	bi	i	bit
17.	Now I say bait	bai	ai	bait

Appendix 2 Mandarin List

1.	我说	必	易	必
2.	我说	爸	啊	爸
3.	我说	布	雾	布
4.	我说	叙	遇	序
5.	我说	簸	哦	簸
6.	我说	个	恶	个
7.	我说	四		四
8.	我说	市		市
9.	我说	二		二
10.	我说	败	爱	败
11.	我说	抱	澳	抱
12.	我说	背		背
13.	我说	够	恆	够
14.	我说	过	卧	过
15.	我说	下	亚	下
16.	我说	别 (扭)	夜	别
17.	我说	挂	袜	挂
18.	我说	秀	又	秀
19.	我说	贵	为	贵

20.	我说	血(肉)	悦	血
21.	我说	鯮	要	鯮
22.	我说	怪	外	怪
23.	我说	殡	印	殡
24.	我说	并	硬	并
25.	我说	办	案	办
26.	我说	谤	盎	谤
27.	我说	笨	摠	笨
28.	我说	蹦		蹦
29.	我说	棍	问	棍
30.	我说	训	运	训
31.	我说	共		共
32.	我说	变	燕	变
33.	我说	项	样	项
34.	我说	贯	万	贯
35.	我说	逛	旺	逛
36.	我说	炫	院	炫
37.	我说	匈	用	匈



13. Age at first exposure to English: \_\_\_\_\_

14. Parents' native languages:

Mother: \_\_\_\_\_

Father: \_\_\_\_\_

15. Languages currently spoken at home:

-----

16. In what countries and for how long have you studied English formally (i.e., in school)?

-----

\_\_\_\_\_ yrs.

-----

\_\_\_\_\_ yrs.

-----

\_\_\_\_\_ yrs.

17. TOEFL score (optional): \_\_\_\_\_ TSE score (optional): \_\_\_\_\_

18. How would you rate your pronunciation?

1 2 3 4 5 6 7 8 9

(native-like)

(strong accent)

19. How often do you use English in your day-to-day affairs?

1 2 3 4 5 6 7 8 9

(never)

(use ONLY English)

20. Additional Information:

## REFERENCES

- Assmann, P., Nearey, T., & Hogan, J. (1982). Vowel identification: Orthographic, perceptual, and acoustic aspects. *Journal of the Acoustic Society of America*, 71, 975-989.
- Best, C. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge: MIT Press.
- Blankenship, B. (1991). Vowel perception in second language. *UCLA Working Papers in Phonetics*, 80, 38-64
- Bohn, O.-S. (1995). Cross-language speech perception in adults: first language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 279-304). Baltimore: York Press.
- Bohn, O.-S., Flege, J. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11, 303-328.
- Bohn, O.-S., Flege, J. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131-158.
- Cebrian, J. (1996). *Voicing in final obstruents and the transfer of phonological rules in Catalan-English interlanguage*. Unpublished master's thesis, University of Toronto.
- Cheng, R. (1966). Mandarin phonological structure. *Journal of Linguistics*, 2, 135-262.
- Chao, Y. (1968). *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Dow, F. (1972). *An outline of Mandarin phonetics*. Canberra: Australian National University Press.

- Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, 27, 315-330.
- Flege, J. (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. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustic Society of America*, 84, 70-79.
- Flege, J. (1991a.). Perception and production: The relevance of phonetic input to L2 phonological learning. In Heubner, T., & Ferguson, C. (Eds.), *Crosscurrents in second language acquisition and linguistic theory*. (pp. 249-289). Philadelphia: John Benjamins.
- Flege, J. (1991b.). The interlingual identification of Spanish and English vowels: Orthographic evidence. *The Quarterly Journal of Experimental Psychology*, 43, 701-731.
- Flege, J. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Baltimore: York Press.
- Flege, J. & Bohn, O.-S. (1989). The perception of English vowels by native speakers of Spanish. *Journal of the Acoustic Society of America*, 85, Suppl. 1, S85
- Flege, J., Munro, M. & MacKay, I. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustic Society of America*, 97, 3125-3134.
- Flege, J. Naoyuki, T. & Mann, V. (1995). Japanese adults can learn to produce English /ɪ/ and /I/ accurately. *Language and Speech*, 38, 25-55.
- Hindle, D. (1978). Approaches to vowel normalization in the study of natural speech. In Sankoff, D., *Linguistic variation: Models & methods*. (pp. 161-171). New York: Academic Press.

- Ho, A. (1976). The acoustic variation of Mandarin tones. *Phonetica*, 33, 353-367.
- Howie, J. (1976). *Acoustical studies of Mandarin vowels and tones*. London: Cambridge University Press.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.
- Lenneberg, E. (1967). *Biological Foundations of Language*. New York: Wiley.
- Li, C., Thompson, S. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Long, M. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.
- 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., & Derwing, T. (1996). The effects of speaking rate on the comprehensibility of native and foreign-accented speech. *Canadian Acoustics*, 24(3), 21.
- 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. (1996). *Speech playback software*. Unpublished computer software, Linguistics Department, Simon Fraser University.
- Nearey, T., Assmann, P. (1986). Modeling the role of inherent spectral change in vowel identification. *Journal of the Acoustic Society of America*, 80, 1297-1308.
- Norman, J. (1988). *Chinese*. Cambridge: Cambridge University Press.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.

Patkowski, M. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, 11, 73-89.

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.

Shih, C. (1995). Study of vowel variations for a Mandarin speech synthesizer. *Proceedings of Eurospeech 1995*. (pp. 1807-1810.) Madrid.

Svantesson, J.-O. (1984). Vowels and diphthongs in standard Chinese. *Lund University Working Papers*, 27, 209-235.

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.

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning*, 41, 177-204.

Werker, J. & Polker, L (1993). Developmental changes in speech perception: new challenges and new directions. *Journal of Phonetics* 21, 83-101.

Wu, Y. (1994). *Mandarin segmental phonology* Unpublished doctoral dissertation, University of Toronto.