EFFECT OF L2 PHONETIC LEARNING ON L1 VOWELS

by

Haisheng Jiang
B.A., Liaoning Normal University, 1988
M.A., Jilin University, 1993

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

DOCTOR OF PHILOSOPHY

In the
Department of Linguistics

SIMON FRASER UNIVERSITY

Summer 2008

APPROVAL

Name: Haisheng Jiang

Degree: Doctor of Philosophy

Title of Thesis: Effect of L2 Phonetic Learning on L1 Vowels

Examining Committee:

Chair: **Dr. Nancy Hedberg**
Associate Professor of Linguistics

Dr. Zita McRobbie
Senior Supervisor
Associate Professor of Linguistics

Dr. Murray Munro
Supervisor
Professor of Linguistics

Dr. Yue Wang
Supervisor
Assistant Professor of Linguistics

Dr. Christian Guilbault
Internal Examiner
Assistant Professor of French

Dr. Hua Lin
External Examiner
Associate Professor of Linguistics, University of Victoria

Date Defended/Approved: August 6, 2008

SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

This research examines the effect of L2 phonetic learning on L1 vowel production. Mandarin-English bilinguals differing in amount of L1 use produced Mandarin and English vowels. An acoustic analysis showed that both the Mandarin-English bilinguals of high L1 use and those of low L1 use deviated from the norm of Mandarin vowel /i/. The Mandarin-English bilinguals of low L1 use who successfully acquired English vowel /aj/ deviated from the norm of Mandarin vowel /aj/, indicating a carry-over effect of L2 vowel on L1 vowel production.

In a perception test, Mandarin vowel production by the Mandarin-English bilinguals was presented to Mandarin as well as English listeners for goodness rating. The results showed that both Mandarin-English bilinguals of high L1 use and those of low L1 use differed significantly from Mandarin monolinguals in the production of /y/, a vowel with no counterpart in English. An analysis of inter-speaker variability indicated that some individual Mandarin-English bilinguals, including both speakers of high L1 use and low L1 use, were accented in the production of /y/, /aj/ and /aʊ/. Possible acoustic properties contributing to their accentedness included lower second formant frequency, larger first or second formant frequency movement, extremely short or long duration, and tone deviation. L2 English learning led to some Mandarin-English bilinguals carrying some English characteristics in their L1 Mandarin vowel production.

In a follow-up perception test, the correlation between the ratings assigned to the Mandarin-English bilinguals' production of Mandarin vowel /y/ and the ratings assigned to their production of English vowel /ɪ/ and /ɛ/ was examined. No inverse correlation was revealed, indicating that good L2 vowel production does not necessarily lead to poor L1 vowel production, and vice versa.

This research suggests that the L1 phonetic system established in childhood does not remain static; instead, it may undergo reorganization when the L1 and L2 phonetic systems coexisting in a common phonological space interact.

Keywords: L2 influence; L1 Mandarin vowel; accentedness; L1 use; cross-language similarity; L2 vowel acquisition

Subject Terms: second language learning; influence; bilingualism; languages in contact

To my beloved wife,

Zhaoqiu (Betty),

and

twin sons,

Ningjian (Rex) and Ningkang (Ryan)

ACKNOWLEDGEMENTS

My sincere appreciation goes to all those who have helped me in the process of completing this work.

I am grateful to my senior supervisor, Dr. Zita McRobbie, who led me into the field of experimental phonetics and guided me, with great patience, through the journey of my PhD study. She has taken good care of me in everything: from academic progress to non-academic issues such as financial support, accommodation and child education.

I am indebted to my committee members, Dr. Murray Munro and Dr. Yue Wang. My PhD work has greatly benefited from the countless discussions I had with them. They were always available for help even with their very tight schedules. From them I have learned not only what speech science is but also what professionalism is.

I thank my external examiner, Dr. Hua Lin and my internal examiner, Dr. Christian Guilbault, for their questions, comments and suggestions, which greatly improved the quality of this work.

I would like to thank Dr. Zita McRobbie and Dr. Yue Wang for the financial support they provided through research assistantships. I also thank Dr. Alexei Kochetov for his generosity in using his SSHRC grant to help cover the cost of my experiments as well as the Linguistics Department and the Language

Training Institute for their help via graduate fellowships and teaching assistantships.

I thank fellow graduate students, Susan Morton, Lorna Fadden, Les Zsoldos and Nicole Carter for their support and friendship. My special thanks goes to Susan Morton who did an excellent job in editing this thesis. Without her work, this thesis would not be as readable.

My sincere appreciation goes to my old friend and colleague, Xinchun (Jean) Wang, who is always available for help. Her support and encouragement have kept me motivated to complete my PhD. My thanks also goes to Rong (Frank) Liu who "babysat" my computer and helped with the *Praat* script.

I also thank all those who participated in my experiments, without whom this work would be impossible.

Finally, I owe the successful completion of this work to my beloved wife, Zhaoqiu (Betty) and twin sons, Ningjian (Rex) and Ningkang (Ryan). To help me pursue my PhD study, my wife gave up her associate professorship in China and moved to Canada to join me. What a sacrifice! My twin sons, who have kept asking "why does doing a PhD take so long", have always had patience with me and have believed that their dad will make it. Thank you, boys!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

It is generally believed that the two phonetic systems coexisting in a bilingual's mind interact with each other bidirectionally (e.g., Flege, 1995; Grosjean, 1989; Pavlenko, 2000; Yeni-Komshian, Flege, & Liu, 2000; Guion, 2003; Baker & Trofimovich, 2005). The first language (L1) phonetic system influences the second language (L2) phonetic system, and vice versa. The former has been well documented (e.g., Weinreich, 1953; Lado, 1957; Flege, 1987, 1995, 2003; Flege, Bohn & Jang, 1997; Flege, Schirru & MacKay, 2003; and many others). The latter, however, has not been studied equally well. Many issues regarding the influence of the L2 phonetic system on the L1 phonetic system remain unresolved. For example, if the L1 phonetic system changes, what is the direction of the change? Does an L1 phonetic segment undergoing modification become more similar or dissimilar to its L2 counterpart? Do factors such as cross-language similarity, amount of L1 use and formation of L2 phonetic categories play a role in the L2 influence on L1? Is the formation of L2 phonetic categories related to L1 phonetic proficiency?

The aim of this dissertation is to address the above-mentioned issues by examining the influence of the L2 vowel system on the L1 vowel system among a group of Mandarin-English bilingual speakers living in Vancouver, Canada. First, I present a review of the previous studies and directions for the present study. It is followed by production tests whose major goal is to examine if the Mandarin-

English bilinguals differ from the Mandarin monolinguals in Mandarin vowel production. Then, I discuss the results of perception tests that examine the Mandarin-English bilinguals' accent in Mandarin vowel production and the correlation between the formation of the L2 English vowel categories and the proficiency of the L1 Mandarin vowels. Finally, a general discussion of the findings concludes the dissertation.

1.1. Previous studies

1.1.1. Major models of cross-language speech perception and production

Three major models of cross-language speech perception and production have been proposed. They are Kuhl and colleagues' *Native Language Magnet* model (NLM) (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Iverson & Kuhl, 1995; Kuhl & Iverson, 1995), Best's *Perceptual Assimilation Model* (PAM) (Best 1994, 1995), and Flege's *Speech Learning Model* (SLM) (Flege, 1981, 1987, 1991, 1992, 1995, 2003).

The NLM maintains that speech prototypes, defined as sounds that are identified as ideal representatives of a phonetic category by adult speakers of a given language, function as perceptual magnets attracting their perceptual variants toward themselves. With regard to the perception of foreign language sounds, the NLM predicts that the native language magnets will pull a similar foreign sound toward a single native prototype. The closer the foreign sound is to the native prototype, the more similar to the native-language it will be perceived.

As a result, L2 learners will not be able to distinguish the similar foreign sound from the native-language sound.

The PAM is primarily concerned with how L2 listeners incorporate L2 sounds into their L1 phonetic system in perception. L2 listeners can perceive the L2 contrasts either as speech sounds or as non-speech sounds. The L2 contrasts that are perceived as speech sounds are further classified into those that can assimilate to native phonetic categories and those that can not. For the L2 contrasts that can assimilate to native phonetic categories, a set of assimilation patterns and degrees of discrimination are proposed.

The SLM is similar to the NLM and the PAM regarding the perceptual assimilation of L2 sounds to similar L1 sounds. According to Flege's principle of *equivalence classification* (1987, 1991, 1992, 1995), L2 sounds that are similar to an L1 sound are identified as instances of the L1 sound. In other words, they are treated as belonging to the same category, even though some phonetic differences between the L2 and L1 sounds exist. As a result, a merged category covering both the L2 sound and the L1 sound will emerge. Accurate production and perception of such L2 sounds are predicted to be difficult for the L2 learner.

Although the three models offer similar accounts in terms of the assimilation of L2 sounds to similar L1 sounds, the SLM differs from the NLM and the PAM in one important aspect. That is, while the NLM and the PAM are perception-oriented and therefore do not make predictions about the production of L1 and L2 speech sounds when cross-language perceptual assimilation occurs, the SLM makes predictions on both perception and production.

A general assumption that the SLM makes is that the L1 phonetic system and the L2 phonetic system coexisting in a bilingual's mind *"remain adaptive over the life span,"* and that the two phonetic systems reorganize through the addition of new L2 phonetic categories or the modification of existing L1 phonetic categories. This assumption implies that, regardless of the age of the language learner, both his L1 phonetic system and L2 phonetic system may undergo reorganization due to the interaction between the two phonetic systems coexisting in "a common phonological space" (Flege, 1995). The interaction between the two phonetic systems is bidirectional in nature (Flege, 1995). That is, the L1 phonetic system influences the L2 phonetic system, and vice versa. With regard to the latter, the SLM makes a specific prediction that the L1 phonetic categories established in childhood do not remain static; instead, they may undergo modification when similar L1 and L2 sounds interact in the process of L2 learning.

According to Flege (2003), the L1 and L2 phonetic categories interact through mechanisms called "category assimilation" and "category dissimilation". The mechanism of "category assimilation" is similar to his principle of "equivalence classification" noted earlier. It claims that if a new category has not been established for an L2 sound that is similar, but not identical, to an L1 speech sound, "category assimilation" may occur. When "category assimilation" occurs, an experienced L2 learner may assimilate the phonetic properties of the L2 sound into the L1 phonetic category, thus developing a category that merges the phonetic properties of the L1 and L2 categories. As a result, pronunciation of the L2 sound will resemble the corresponding L1 sound and pronunciation of the L1 sound will be L2-like. For example, Flege (1987) examined the voice onset time (VOT) of the /t/ production in French and English by experienced French-English bilinguals and English-French bilinguals. He found that the French-English bilinguals produced English /t/ with a shorter VOT than the English monolinguals, but with a longer VOT than the French monolinguals. The English-French bilinguals, on the other hand, produced French /t/ with a longer VOT than the French monolinguals, but with a shorter VOT as compared with the English monolinguals. This indicates that neither the French-English bilinguals nor the English-French bilinguals produced the L2 /t/ in a native-like fashion. Instead, he proposed that a merged category was formed for the two groups of bilingual speakers. The study also found that, as compared with the French monolinguals, the French-English bilinguals produced French /t/ with a longer VOT value, indicating the influence of English VOT on French VOT. Conversely, as compared with the English monolinguals, the English-French bilinguals' production of English /t/ was shorter, showing the influence of French VOT on English VOT. These findings illustrate the bidirectional nature of interference between L1 and L2. In the process of L2 phonetic learning, not only does the L1 phonetic system influence the L2 phonetic system, but also the L2 phonetic system influences the L1 phonetic system.

Another aspect of the SLM, on the other hand, is that if a new category has been established for an L2 sound, "category dissimilation" may occur. When "category dissimilation" occurs, a newly established L2 category and the nearest

L1 speech category may shift away from one another to maintain the phonetic contrast in a common phonological space. As a result, the L2 category and the L1 category will differ from the categories of monolinguals in either language. In fact, though, not much evidence (except Flege & Eefting, 1987, 1988) is available to support the predictions of “category dissimilation”. Flege & Eefting (1987) obtained evidence of “category dissimilation” when they examined the VOT in the production of Spanish /p, t, k/ by Spanish-English bilinguals. They found that the Spanish-English bilingual children and adults, both having exposure to English early in life, produced Spanish /p, t, k/ with a shorter VOT than the age-matched Spanish monolinguals. Flege & Eefting (1988) later argued that these early Spanish-English bilinguals had established new categories for English /p, t, k/. It was hypothesised that in order to maintain a phonetic distinction between the category of Spanish /p, t, k/ and that of English /p, t, k/ established later in life, the Spanish-English bilinguals shortened the VOT of L1 Spanish /p, t, k/ (Flege, 2007).

Be it “category assimilation” or “category dissimilation,” phonetic categories in a bilingual’s L1 do not remain static. Instead, they may differ from those of monolinguals. Depending on whether or not a similar L2 phonetic category has been established, an L1 phonetic category becomes either similar or dissimilar to its L2 counterpart.

To summarize, the NLM, the PAM and the SLM all make predictions on the perception of L1 and L2 sounds. However, only the SLM makes predictions on both perception and production of L1 and L2 sounds. The SLM posits that the

L1 phonetic system and the L2 phonetic system interact bidirectionally. With regard to the influence of the L2 phonetic system on the L1 phonetic system, the SLM predicts that the phonetic categories established in childhood do not remain static; instead, they will undergo modification through the mechanisms of "category assimilation" and "category dissimilation". Since the focus of the present study is the influence of L2 phonetic learning on L1 vowel production, the SLM is the most relevant and is therefore adopted as the theoretical framework of the present study.

1.1.2. Factors related to L2 influence on L1

Previous studies of L2 influence on an L1 at the phonetic level[1] (Baker & Trofimovich, 2005; Guion et al., 2000; Guion, 2003; Flege et al., 2003; Harada, 2003; Yeni-Komshian et al., 2000; Peng, 1993; Major, 1992; among others) have shown that a variety of factors are related to L2 influence on L1. These factors include, but are not limited to, the acquisition of L2 vowel or consonant categories, the age of L2 learning, the length of residence, the amount or extent of L1 use, and the pronunciation proficiency in an L2.

1.1.2.1. Acquisition of L2 phonetic categories

It is believed that the acquisition of L2 phonetic categories may account for the influence of an L2 on an L1 at the phonetic level (Flege, 1987, 1995, 2003; Guion, 2003). According to Flege et al. (2003), the closer to an L2 phonetic norm a bilingual gets, the more his/her production of the similar (but not identical) L1

---

[1] Previous studies at the syntactic, semantic and pragmatic level are not the concern of the present study.

speech sound tends to deviate from L1 phonetic norms. Flege's (1987) study of VOT and Guion's (2003) study of vowels provided evidence for the above claim. Flege (1987) found that, when compared with French monolinguals, French-English bilinguals produced French /t/ with a longer VOT value, indicating the influence of English VOT on French VOT. Conversely, when compared with English monolinguals, English-French bilinguals produced English /t/ with a shorter VOT value, indicating the influence of French VOT on English VOT. Guion (2003), in her examination of the vowel systems of Quichua-Spanish bilinguals, found that, as compared with the Quichua-Spanish bilinguals who equated Spanish /i/, /e/ and Quichua /ɪ/, those who equated Spanish /i/ with Quichua /ɪ/ but differentiated them from Spanish /e/, produced a higher Quichua /ɪ/. It was proposed that the acquisition of Spanish /e/ conditioned a raising of Quichua /ɪ/. A similar pattern was observed in back vowels. When compared with the bilinguals who equated Spanish /o/, /u/ and Quichua /ʊ/, those who equated Spanish /u/ with Quichua /ʊ/ but differentiated them from Spanish /o/ produced a higher Quichua /ʊ/. Acquisition of Spanish /o/ was believed to condition the raising of Quichua /ʊ/. In most cases, the bilinguals who equated Spanish /i/ and Quichua /ɪ/ but differentiated them from Spanish /e/ also equated Spanish /u/ with Quichua /ʊ/ but differentiated them from Spanish /o/. In addition, the bilinguals' production of Quichua low vowel /a/ was found to be related to the successful acquisition of Spanish /a/. The bilinguals who had acquired distinct front and back vowels had also acquired a distinct Spanish /a/. They produced their Quichua /a/ higher as compared with those who had not acquired any of the Spanish vowels. In sum, the successful acquisition of L2 Spanish vowels triggered the reorganization of the corresponding L1 Quichua vowels. Both Flege's study and Guion's study show that a bilingual's L1 phonetic categories diverge from the phonetic norms of L1 monolinguals as a result of acquiring the L2 phonetic categories.

1.1.2.2. Age of L2 learning

Age of L2 learning, often indexed by age of arrival (AOA) in an L2 speaking country, may be the most important factor determining the degree to which the L2 phonetic system influences the L1 phonetic system. Previous studies (Baker & Trofimovich, 2005; Harada, 2003; Yeni-Komshian et al., 2000) have shown that, other factors being equal, the earlier one learns an L2, the more likely his/her L1 is influenced by his/her L2.

Baker & Trofimovich (2005), for instance, examined the interaction of native and second language vowel systems in early and late Korean-English bilinguals. Acoustic analysis showed that the early bilinguals, who had an age of arrival (AOA) ranging from 7 to 13 years and a length of residence (LOR) ranging from 5 to 15 years, differed from age-matched Korean monolinguals in the production of Korean /i/, /u/, /ɛ/. Specifically, when compared with the age-matched Korean monolinguals, these early bilinguals produced Korean /i/, /u/, /ɛ/ higher in the vowel space. In addition, they produced Korean /u/ more anterior in the vowel space. Baker & Trofimovich (2005) concluded that these Korean vowels were “colored” by English vowels after extensive experience with English. However, it is not clear how English vowels “colored” these Korean vowels. Their analysis did not show an acoustic difference between Korean /i/ and /u/ and English /i/ and /u/. Nor did it show that English /ɛ/ was higher and more anterior than Korean /ɛ/. In fact, English /ɛ/ was lower and more posterior than Korean /ɛ/. When interpreting the L2 influence on L1 among the early bilinguals, they claimed that the influence was more likely to occur in early bilinguals than in late bilinguals, because the young learners' L1 was still developing, thus more susceptible to restructuring.

The role of AOA in L2 influence on L1 is also shown in Yeni-Komshian et al. (2000) and Harada (2003). Yeni-Komshian et al. studied Korean-English bilinguals' global pronunciation proficiency in Korean and English and found that the global pronunciation of the majority of Korean-English immigrants who came to the U.S.A at age 12 or later was rated at the same level as Korean monolinguals residing in Seoul, Korea. However, the younger immigrants, whose AOA ranged from 1 to 11, were rated significantly lower than Korean monolinguals. Harada's examination of early Japanese-English bilinguals' VOT of Japanese voiceless stops /p/, /t/ and /k/ showed that the mean VOT values produced by the early Japanese-English bilinguals were greater than those of the monolingual Japanese speakers, regardless of the place of articulation. He suggested that the longer VOT of the L2 English voiceless stops affected the bilinguals' VOT of Japanese voiceless stops.

Despite the general agreement that the earlier one learns an L2, the more likely his/her L1 is influenced by his/her L2, previous studies (Flege and Hillenbrand, 1984; Major, 1992; Peng, 1993; Sancier and Fowler, 1997; Munro, Derwing, & Flege, 1999) have shown that adult learners can also modify their L1 or dialect 1 (D1) as a result of L2 learning or contact with dialect 2 (D2).

For example, Flege and Hillenbrand (1984) found that proficient adult French speakers of English, who had a mean age of 38 years and a mean length of residence (LOR) of 12.2 years in the United States, produced /t/ in the French syllables /tu/ ("tous") and /ty/ ("tu") with a mean VOT value of 54 milliseconds (ms). This mean VOT value was significantly greater than the approximately 20 ms VOT values found in the speech of French monolinguals (Caramazza & Yeni-Komshian, 1974, cited in Flege & Hillenbrand, 1984). Flege and Hillenbrand claimed that these French speakers' L2 English learning affected the production of their L1 French VOT. They hypothesized that these French speakers of English merged the short-lag VOT of French /t/ with the long-lag VOT of English /t/ by identifying the L2 /t/ as an exemplar of the L1 /t/. A single phonetic category was employed for both the French /t/ and the English /t/. As a result, these French speakers of English developed a VOT of French /t/ that was intermediate between that of the French monolinguals and that of the English monolinguals. Two other studies of VOT (Major, 1992; Sancier and Fowler, 1997) also showed that adults might modify their L1 as a result of L2 learning. Major's study revealed that the English VOTs of the adult American immigrants in Brazil deviated from the English monolingual speaker values toward the direction of Brazilian

Portuguese. Sancier and Fowler measured the VOT of voiceless stops /p/ and /t/ in Spanish and English produced by a 27-year old Portuguese learner of English. They found that the speaker's VOT in Portuguese /p/ and /t/ became longer after a 4-month stay in the United States, thus drifting toward the VOT values of the English stops. This indicated that L2 English long-lag VOT affected the learner's L1 Portuguese short-lag VOT. A similar effect of L2 learning on L1 was found in Peng's (1993) study of consonants, which showed that highly proficient adult Amoy speakers of Mandarin tended to produce Amoy /h/ with well-defined striations as observed on wideband spectrograms. The well-defined striations, which indicated secondary uvular vibration caused by velar constriction (Fant, 1960, cited in Peng, 1993), are characteristic of a typical Mandarin /x/. Based on this finding, she claimed the bilinguals' Amoy /h/ productions had the acoustic feature of the "similar" Mandarin /x/ due to the acquisition of Mandarin /x/.

The phonetic interference that happens to adults occurs not only across languages (L1 and L2) but also across dialects (D1 and D2). For example, Munro et al.'s (1999) study of the influence of American English on Canadian English showed that the speech samples of adult Canadians residing in Alabama, U.S.A., were rated to have an American accent by both Canadian listeners in Canada and American listeners in Alabama. Tokens of "wife", "rifle", "like", "driving", "highway" and "goodbye" were selected from the speech samples of the Canadians in Alabama, Canadians in Canada, and Americans in Alabama. Phonetically trained listeners rated these tokens on a 5-point scale, with 1 being "very American" and 5 being "very Canadian." The listeners were instructed to rate the tokens with Canadian raising /ʌɪ/ as "very Canadian," and those with monophthongal /a/ as "very American." Half of the ratings assigned to the word tokens of the Canadians in Alabama were toward the "American" end of the scale indicating the absence of Canadian raising and the presence of monophthongal /a/ in some tokens produced by Canadians in Alabama. The authors proposed that the absence of Canadian raising probably contributed to the speech samples of the Canadians in Alabama being rated as having an American accent. Their study shows that the D1 (Canadian English) of some Canadians in Alabama had some characteristics of the D2 (Alabama English) as a result of contact with the D2.

In sum, it seems that early bilinguals are more susceptible to restructuring their L1 phonetic system than late bilinguals. However, adult learners may also modify their L1 phonetic system in the process of L2 learning.

1.1.2.3. Length of residence

The length of residence (LOR) in an L2 speaking country is usually considered an index of experience with L2. However, previous studies of L2 influence on L1, such as Guion et al. (2000), Guion (2003), Flege et al. (2003), Harada (2003), Yeni-Komshian et al. (2000) and Peng (1993) rarely controlled for LOR to be an independent variable. To the best of my knowledge, only Baker & Trofimovich (2005) controlled LOR in examining L2 influence on L1.

Baker & Trofimovich (2005) found that, compared with Korean monolinguals matched for age, the early Korean-English bilinguals with a longer

LOR (mean=8 years) modified their production of Korean /i/, /u/, /ɛ/. However, early Korean-English bilinguals with a similar AOA but a shorter LOR (mean=1.3 years) did not show this trend. In addition, neither the late Korean-English bilinguals (AOA=15-31 years) with longer LOR (mean=6.9 years) nor those with shorter LOR (mean=0.9 years) differed from the Korean monolinguals in the production of Korean vowels, /i/, /ɛ/, /e/, /u/ and /ɨ/. Baker & Trofimovich's finding suggests that LOR plays a role in L2 influence on L1 only among early bilinguals.

1.1.2.4. Amount of L1 use

One of the factors affecting the nature and strength of the influence of the L1 and the L2 on one another is the amount and circumstances of L1 and L2 use (Grosjean, 1992, 2001). It has been well documented that amount of L1 use affects the extent to which the L1 phonetic system influences the L2 phonetic system (e.g., Flege, Frieda & Nozawa, 1997; Flege, Schirru & MacKay, 2003; Guion et al., 2000; among others). Other factors being equal, the more a bilingual uses his/her L1, the less accurate and more accented his/her L2 production is. However, the relationship between amount of L1 use and extent of L2 influence on L1 is far from clear. Only a few studies of L2 influence on an L1 take amount of L1 use into consideration (McRobbie, 2003; Guion et al., 2000). The findings of these studies are mixed.

McRobbie (2003) carried out an acoustic study of Hungarian immigrants' production of Hungarian [ɒ] and found that the immigrant speakers who had minimal contact with members of the Hungarian community diverged the least from the monolingual Hungarian speakers in their pronunciation of Hungarian [ɒ]. I assume that the less contact one has with his/her L1 community, the less the amount of L1 use there is. Thus, McRobbie's study suggests that the bilinguals using L1 less tended to deviate the least from the monolinguals in the production of Hungarian [ɒ]. Guion et al. (2000) examined the effect of L1 use on pronunciation in Quichua-Spanish bilinguals who learned their L2 Spanish when they started school. Quichua-Spanish bilingual speakers were divided into a low L1 use group, a mid L1 use group and a high L1 use group. It was found the mean ratings of the Quichua sentences for the three groups of Quichua-Spanish bilinguals did not differ much from one another and from the near-monolingual Quichua controls. This indicates that these bilinguals did not have noticeably different degrees of Spanish foreign accent in L1 Quichua. Amount of L1 Quichua use did not affect their pronunciation of L1 sentences. The mixed findings as noted in McRobbie (2003) and Guion et al. (2000) give rise to the need to examine further whether the amount of L1 use is an important factor when examining L2 influence on an L1.

1.1.2.5. Pronunciation proficiency

As reviewed in Yeni-Komshian et al. (2000), if the SLM is correct, segmental changes in an L1 may result as L2 phonetic learning increases. With regard to global pronunciation proficiency, an inverse relationship between L1 and L2 was predicted (Yeni-Komshian et al., 2000). Their study showed that the participants whose AOA ranged from 1 to 9 were rated to have better English pronunciation scores and worse Korean pronunciation scores. The participants with an AOA of 12-23 were judged to be worse in English pronunciation, but better in Korean pronunciation. The only group having equal or slightly above average pronunciation scores in both English and Korean were those with an AOA of 10-11. Indeed, an inverse correlation between L2 and L1 pronunciation was found for most of the Korean bilinguals in this study. However, the focus of Yeni-Komshian et al.'s study was to investigate if the speakers of different age groups behaved differently in their L1 and L2 production. No attempt was made to examine whether good producers of an L2 tended to be those who differed from L1 monolinguals in the production of L1, and vice versa.

1.1.2.6. Summary

To summarize, it is evident that a bilingual's existing L1 phonetic categories may be reorganized in the process of acquiring the corresponding L2 phonetic categories (Flege, 1987; Guion, 2003). It is also established that both early learners (Baker & Trofimovich, 2005; Yeni-Komshian et al., 2000; Harada, 2003) and adult learners (Flege and Hillenbrand, 1984; Peng, 1993; Sancier and Fowler, 1997) may modify their L1 as a result of L2 learning. With regard to LOR, early bilinguals of longer LOR, but not those of shorter LOR, tend to modify their L1 phonetic categories (Baker & Trofimovich, 2005). However, the role that amount of L1 use plays in L2 influence on L1 is not apparent (McRobbie, 2003; Guion et al. 2000). Moreover, while early bilinguals tend to have worse L1 pronunciation than L2 pronunciation, it is not clear whether good producers of L2 within a certain age group tend to be those who differ from L1 monolinguals in the production of L1.

1.1.3. Mandarin Chinese and Canadian English vowels

1.1.3.1. Mandarin vowel system

It is generally accepted that Mandarin Chinese, including both Beijing Mandarin and Taiwanese Mandarin[2], has five vowel phonemes, namely, /i/, /y/, /u/, /ə/ and /a/ (Cheng, 1966; Lin, 1989; Wan & Jaeger, 2003). The high, front, unrounded vowel /i/ has three allophones, which are [ɯ], [ɨ], and [i], respectively. The dental apical vowel [ɯ] occurs after dental sibilants [ts], [ts'] and [s]. The retroflex apical vowel [ɨ] occurs only after retroflex fricatives and affricates [tʂ], [tʂ'] [ʂ] and [ɹ][3]. [i] occurs elsewhere. The allophones of the mid vowel /ə/ are [e], [ɛ], [o], [ɤ], [ɔ], and [ə] (Howie, 1976; Chao, 1968; Wu, 1994; Wan & Jaeger, 2003). [e] occurs before [i]; [ɛ] occurs after a glide or [i] and [y]; [o] occurs before [w]; [ɤ] occurs in an open syllable as a monophthong; [ɔ] occurs after the labial consonants in open syllables; [ə] occurs in closed syllables before nasals. The low vowel /a/ varies allophonically from central to back variants [a] and [ɑ] (Wan & Jaeger, 2003). [a] occurs in an open syllable and before [n] and [j]; [ɑ] occurs before [w] and [ŋ].

Besides monophthongs, Mandarin Chinese has nine diphthongs (/aj/, /ej/, /aʊ/, /ia/, /ie/, /oʊ/, /ua/, /uo/, /ye/) and four triphthongs (/iao/, /iou/, /uai/, /uei/) (Wu & Lin, 1989).

---

[2] The bilinguals in the present study were speakers of Taiwanese Mandarin.

[3] Wan & Jaeger (2003) believed that the dental apical vowel [ɯ] and the retroflex apical vowel [ɨ] were phonetically indistinguishable in Taiwanese Mandarin for the reason that most Taiwanese Mandarin speakers did not produce strongly retroflexed vowel after retroflex consonants.

1.1.3.2. Canadian English vowel system

Canadian English has 13 vowel phonemes, /i/, /ɪ/, /e/, /ɛ/, /æ/, /ʌ/, /u/, /ʊ/, /o/, /ɒ/, /aj/, /aʊ/ and /ɔɪ/, among which /e/ and /o/ are "phonetic" diphthongs (Nearey & Assmann, 1986).

1.1.3.3. Acoustic comparisons of Mandarin and Canadian English vowels

Cross-language acoustic comparisons of English and Mandarin vowels are limited. To my knowledge, only Wang (1997) carried out such a study. Wang (1997) compared Mandarin vowels [i], [ej], [u], [oʊ] and [a] with their Canadian English counterparts in terms of duration and the first and second formant frequencies (F1, F2). The author found that Canadian English [i] and [ej] were significantly longer than their Mandarin counterparts. Mandarin [u], [oʊ] and [a] and their Canadian English counterparts were non-significantly different in duration. With regard to spectral properties, no significant difference was found between Mandarin and Canadian English vowel [i]. Mandarin [ej] was significantly lower and more posterior than its Canadian English counterpart at measurement (a) (30% distance into the vowel), but no significant difference was found at measurement (b) (70% distance into the vowel). Mandarin [oʊ] was significantly lower at both measurement (a) and (b) but only more posterior than its Canadian English counterpart at measurement (b). Mandarin [u] was more posterior than its Canadian English counterpart. The low central Mandarin [a] was significantly lower and more anterior than the low back Canadian English [ɒ]. This study of cross-language similarity is based on acoustic analysis. It provides quantitative information about the similarities and differences between Mandarin and Canadian English vowels. However, acoustic measurements alone may not suffice to determine cross-language similarity (Flege, Bohn & Jang, 1997). For this reason, a perceptual assimilation test is included in the present study.

1.2. The present study

1.2.1. Design

The present study examines the influence of L2 English learning on L1 Mandarin vowel production among a group of Taiwanese Mandarin-English bilinguals[4] to address the unresolved issues arising from the previous studies, such as the effect of L1 use and the correlation between L2 vowel pronunciation proficiency and L1 vowel pronunciation proficiency. The design of this study is described as follows.

First, the AOA of the Mandarin-English bilinguals in the present study is controlled to range from 9 to 13 years (mean=11.3 years). The selection of subjects of this particular AOA range is motivated by the consideration that the bilinguals under study must be those who have established an L1 Mandarin phonetic system before being exposed to L2 English. This is important, because I believe that the effect of L2 learning on L1 can be established only when the bilinguals have mastered L1 phonology. Otherwise, when the bilinguals are found

[4] The nature of this study required that all subjects speak the same dialect of Mandarin. I had attempted to recruit bilingual Beijing Mandarin-English speakers in the pilot study. However, very few subjects met my selection criterion (e.g. AOA=9-13 years; university students; came to Canada from the same city). I therefore decided to recruit only Taiwanese Mandarin speakers coming from Taipei, Taiwan. This resulted in far more Mandarin-English bilingual speakers that met my selection criterion.

to deviate from L1 monolinguals in the production of an L1 phonetic segment, it is hard to say whether the deviation is due to an underdeveloped L1 phonology or due to the effect of the L2 phonology. The bilinguals in the present study are assumed to have established an L1 Mandarin phonetic system before being exposed to English. Previous studies of L1 acquisition have shown that complete mastery of phonology, productive control of most of syntactic structures, and early literacy are achieved by about age eight (Smit, Hand, Freilinger, Bernthal and Bird, 1990; Snow, Burns and Griffin, 1998; cited in Yeni-Komshian, et al., 2000).

Second, the Mandarin-English bilinguals in the present study are divided into a group of high L1 use and a group of low L1 use according to their self-report of L1 use. The aim of this design is to examine the effect of L1 use on L2 influence on L1. In fact, there is no commonly agreed criterion of "low L1 use" and "high L1 use." For example, in Flege, Frieda and Nozawa's study (1997), Italian speakers of low L1 use had an average of 3% Italian use. Italian speakers of high L1 use had an average of 36% Italian use. In Guion et al.'s study (2000), the low L1 use group had an average of 41%[5] Quichua use ; the mid L1 use group had an average of 64% Quichua use; the high L1 use group had an average of 91% Quichua use. As reviewed in Guion et al. (2000), it was possible that some participants in Flege Frieda and Nozawa's (1997) low L1 use (3%) had never fully developed or had lost their Italian production proficiency. For this reason, the present study does not follow Flege Frieda and Nozawa's (1997)

[5] For easy comparison, scores in Guion et al.'s study (2000) were converted to percentages by the author using the formula: score/7*100.

criterion of low and high L1 use. Since the bilinguals in the present study report using Mandarin ranging from 20% to 80% of the time, it may be more reasonable to follow Guion et al.'s criterion (2000). However, none of the bilinguals in the present study report using Mandarin 90% of the time or above. A category consistent with Guion's "high L1 use" is non-existent in the present study. For convenience, the present study assigns Mandarin-English bilinguals to two categories, namely, "low L1 use" and "high L1 use." Those who use Mandarin 40% and less are categorized as "low L1 use." Those who use Mandarin 40% and above are put into the category of "high L1 use."

Third, the present study administers both production tests and perception tests. Most previous studies of L2 influence on an L1 have focused on either acoustic analysis (e.g., Peng, 1993) or perceptual analysis (e.g., Yeni-Komshian et al., 2000). A comprehensive study of L2 influence on an L1 should include both. An acoustic analysis is important because it can quantify the acoustic characteristics of the phonetic segments produced by monolinguals and bilinguals. These quantified acoustic characteristics, such as formant frequencies and duration, can show if and how the bilinguals acoustically differ from the monolinguals. However, the acoustic differences between the bilinguals and the monolinguals do not necessarily have perceptual saliency. Therefore, a perceptual study must also be included to test such saliency. In addition, it is impractical to measure every possible acoustic feature of some given speech samples in acoustic analysis. Some acoustic features (e.g., voice quality) contributing to accent may not be captured in an acoustic analysis (Murray

Munro, 2004, personal communication). However, such acoustic features may be revealed by a perceptual analysis.

Finally, the present study recruits both L1 Mandarin monolingual listeners and L2 English monolingual listeners to do the rating task in the perception tests. The majority of the previous perceptual studies of L2 influence on L1 recruited only monolingual listeners of the L1 to rate the overall pronunciation of the bilinguals' L1 (e.g., Guion et al., 2000; Yeni-Komshian et al., 2000). However, if L2 learners' L1 is "colored" by L2, thus carrying some acoustic characteristic of L2, both L1 and L2 monolingual listeners should be able to detect the "color." That is, L1 monolingual listeners should be able to detect the non-native "color" (unlike L1) whereas L2 monolingual listeners should be able to detect the native "color" (like L2). Selection of L1 vowels and their L2 phonetic counterparts as target vowels makes it possible for monolingual listeners of both L1 and L2 to do the rating task.

1.2.2. Target vowels

The Mandarin vowels examined in this study are /i, y, u, a, aj, aʊ, ej, oʊ/, among which all but /y/ have English counterparts. Selection of the Mandarin vowels having similar English counterparts allows the researcher to examine the predictions in Flege's "category assimilation" and "category dissimilation" whose basis is cross-language phonetic similarity (similar sounds). On the other hand, the selection of /y/, whose counterpart is non-existent in English, is motivated by the fact that effect of L2 phonetic learning on such a dissimilar sound is not predicted in any of the speech learning models (NLM, PAM and SLM) and needs further examination. Since this study is of exploratory nature, no attempt is made to include all Mandarin vowels that do not have obvious English counterparts (e.g., diphthongs /ia/, /ie/, /ua/, /uo/, /ye/ and all triphthongs).

English vowels are also included in the present study for two reasons. First, doing so enables the determination of cross-language vowel similarity between Mandarin and English. Second, it helps determine the correlation between acquisition of L2 vowels and the modification of L1 vowel production. The English vowels selected for the present study are /i, ɪ, e, ɛ, æ, u, ʊ, o, ɒ, ʌ, aj, aʊ/. English /i/, /e/, /u/, /o/, /ɒ/, /aj/ and /aʊ/ each has a Mandarin counterpart, which is /i/, /ej/, /u/, /oʊ/, /a/, /aj/ and /aʊ/, respectively. The inclusion of /ɪ/, /ɛ/, /æ/, /ʌ/ and /ʊ/, which do not have obvious Mandarin counterparts, is motivated by the fact that most of them are difficult for Mandarin learners of English to produce and perceive (Wang, 1997). An acoustic and perceptual analysis of the Mandarin-English bilinguals' production of these vowels may indicate if the bilinguals have successfully formed categories for these L2 vowels, which in turn helps answer questions raised in 1.2.3 (e.g., "Is there an inverse correlation between the Mandarin-English bilinguals' English vowel pronunciation proficiency and their Mandarin vowel pronunciation proficiency?"). English diphthong /ɔɪ/ is not selected because it does not have a Mandarin counterpart.

1.2.3. Research questions

A general goal of the present study is to investigate if the Mandarin-English bilinguals' production of the target Mandarin vowels undergoes modification as a result of L2 English learning. Moreover, the issues arising from the previous studies such as cross-language similarity, effect of L1 use, and correlation between formation of L2 vowel categories and L1 vowel production are addressed by answering the following questions.

Research question 1: Do the Mandarin-English bilinguals differ acoustically from the Mandarin monolinguals in the production of the target Mandarin vowels? If so, do cross-language similarity, amount of L1 use and formation of L2 vowel categories play a role? Production tests are administered to address these issues.

Research question 2: Do the Mandarin-English bilinguals have an accent in their L1 vowel production? If so, what acoustic properties are associated with this accent? Are Mandarin-English bilinguals of high L1 use and those of low L1 use equally judged as accented?

Research question 3: Is there an inverse correlation between the Mandarin-English bilinguals' English vowel pronunciation proficiency and their Mandarin vowel pronunciation proficiency? Perception tests are administered to deal with the issues raised in research question 2 and 3.

CHAPTER 2: MANDARIN AND ENGLISH VOWEL PRODUCTION BY MANDARIN-ENGLISH BILINGUALS

This chapter describes and discusses three production tests. First, Mandarin monolinguals are compared with English monolinguals to determine cross-language vowel similarity. Second, Mandarin-English bilinguals are compared with English monolinguals to establish if Mandarin-English bilinguals have formed L2 English vowel categories. Finally, Mandarin-English bilinguals are compared with Mandarin monolinguals to examine if Mandarin-English bilinguals deviate from Mandarin monolinguals in Mandarin vowel production.

2.1. Methodology

2.1.1. Speakers

Thirteen monolinguals of Taiwanese Mandarin (MonoM), 12 monolinguals of Canadian English (MonoE) and 33 Taiwanese Mandarin-English bilinguals differing in amount of L1 use (BiMH and BiML) participated in this study. They reported no disorders in speaking and hearing.

The characteristics of MonoM, MonoE, BiMH and BiML are summarized in Table 2-1.

Table 2-1: Characteristics of Mandarin monolinguals (MonoM), English monolinguals (MonoE), Mandarin-English bilinguals of high L1 use (BiMH) and Mandarin-English bilinguals of low L1 use (BiML)

		% use of Mandarin	Age (year)	AOA (year)	LOR (year)	Years of English Study	% of English TV	% of English Movie	% of English Radio
MonoE	M	-----	27	-----	-----	-----	-----	-----	-----
	SD	-----	(5)	-----	-----	-----	-----	-----	-----
	Range	-----	19-36	-----	-----	-----	-----	-----	-----
MonoM	M	97	24	24	.1	9	-----	-----	-----
	SD	(3)	(4)	(4)	(.04)	(2.0)	-----	-----	-----
	Range	90-100	17-32	17-32	.08-.16	6-12	-----	-----	-----
BiMH	M	65	22	11.6	9.9	12	66	80	65
	SD	(8)	(2)	(1.2)	(2.2)	(2.5)	(29)	(20)	(28)
	Range	60-80	19-25	10-13	6-14	6-15	10-100	20-100	10-100
BiML	M	30	21	10.9	10.4	12.9	80	91	85
	SD	(9)	(2)	(1.6)	(2.0)	(2.9)	(27)	(14)	(24)
	Range	20-40	18-24	9-13	6-14	8-17	0-100	50-100	30-100

Note. Dashes indicate the values were not available.

The 13 Mandarin monolinguals (5 males and 8 females) were from Taipei, Taiwan. They were either visitors or ESL students in Vancouver. Their LOR and ages ranged from 1 to 2 months (mean=1.5 months) and 17 to 32 years (mean =24 years), respectively. Except for one high school graduate, all of the participants were either studying at a university in Taiwan or had recently graduated from a university in Taiwan. Since English is a required course in Taiwanese schools, they all had studied English in a classroom setting since junior high school. On average, they had studied English for 9 years. However, they rarely used English when they were in Taiwan. The self-rating of their English speaking proficiency in the questionnaire (Appendix 2) showed that the majority rated themselves "not fluent" in speaking English (10 cases of "not fluent", 3 cases of "somewhat fluent" and 0 case of "fluent"). Besides English, most of these speakers reported having some knowledge of Taiwanese[6]. In general, their speaking proficiency in Taiwanese was low. Nine of them rated themselves "not fluent", 4 rated "somewhat fluent" and no one rated "fluent" in speaking Taiwanese. No one reported having knowledge of Hakka, a dialect spoken in Taiwan. The self-reported mean percentage of their Mandarin use at home, in school, at work and with friends was as high as 97%. This indicates that Mandarin was the main language in their daily communication. Their self-report also shows that Mandarin was their first language.

The 12 monolingual speakers of Canadian English (6 males and 6 females) were undergraduates and graduates attending Simon Fraser University. They were all born and raised in Anglophone regions of Canada west of Quebec. In fact, 10 out of the 12 participants were born and raised in the province of British Columbia. Most of them reported having some knowledge of a second language (e.g., French or Spanish) obtained during high school. However, they all reported very low proficiency in the second language they had learned. Their ages ranged from 19 to 36 years (mean= 27 years).

The 33 Mandarin-English bilinguals were undergraduates at Simon Fraser University. All of them had immigrated to Vancouver from Taipei, Taiwan, with their parents when they were children. This group had an AOA ranging from 9 to 13 years (mean=11.3 years). At the time of the study, their LOR and ages ranged from 6 to 14 years (mean=10.2 years) and 18 to 25 years (mean=22 years), respectively. The majority of these speakers spoke Mandarin and English only.

---

[6] Taiwanese is a variant of Amoy Min Nan Chinese. Although most people in Taiwan can speak Taiwanese, the degree of fluency varies widely (Wikipedia, the free encyclopedia, n.d.).

Only 9 out of the 33 speakers reported having some knowledge of Taiwanese in addition to Mandarin and English. However, they rated themselves "not fluent" in speaking Taiwanese and claimed having very few chances to speak it in Vancouver. No one reported speaking Hakka.

Based on the amount of L1 Mandarin use, the 33 Mandarin-English bilinguals were further divided into a group of high Mandarin use (BiMH) (n=16) and a group of low Mandarin use (BiML) (n=17). Participants were asked to answer questions regarding the language most used at home, at school, at work, with friends and in day-to-day affairs (Appendix 1). Three answers were possible. They were "Mandarin", "English" or "Both about the same". Following the practice of Guion et al. (2000), participants were given one point for each "Mandarin" answer, half a point for each "Both about the same" answer, and zero for each "English" answer. The possible maximum total score for a participant was five. Since the question regarding the most used language at work might not apply to every participant, it was possible some participants had answers for only four questions and thus a possible maximum total score of four. To calculate the percentage of L1 Mandarin use, a participant's total score was divided by the number of questions answered, and then multiplied by 100. Those who used Mandarin 40% and less were categorized as "low L1 use" and those who used Mandarin 40% and above were put into the category of "high L1 use". The BiMH's use of Mandarin was significantly more than that of the BiML [$t(31) = 11.71$, $p < .001$]. However, the two groups were comparable in age, AOA, LOR, years of English study, overall percentage of watching English TV programs overall percentage of watching English movies and overall percentage of listening to English radio programs. A one-way ANOVA revealed a main effect of age between the MonoM, MonoE, BiMH, and BiML [$F(1, 31) = 6.9$, $p<.001$]. A Tukey test showed a significant difference only between MonoE and the two Mandarin bilingual groups (for both groups, $p<.01$). This suggests that MonoM, BiMH and BiML were comparable in age and MonoE were older than BiMH and BiML.

2.1.2. Stimuli

English and Mandarin stimuli are summarized in Table 2-2. The initial consonant of all stimuli was controlled to be the same (/p/) whenever possible so that the target vowels have a comparable phonetic context both intra- and cross-linguistically. If a word with initial /p/ was not available in Mandarin, a word with an initial consonant occurring in both Mandarin and English was selected (e.g., /doʊ/ and /ly/). All Mandarin words were in open syllables. English words containing vowels that are Mandarin counterparts (e.g., /i, u, o, ɑ, aj, aʊ, e/) were also in open syllables. In order to match the English words as closely as possible, all Mandarin words were in Tone 4, which is equivalent to a "high-low" in terms of intonation patterns of English (Dow, 1972, cited in Wang, 1995).

2.1.2.1. Mandarin stimuli

The Mandarin vowels examined in this study were /i, y, u, a, aj, aʊ, e, oʊ/. Since no Tone 4 open syllable words containing /oʊ/ and /y/ begin with initial /p/ in Mandarin, a word with initial /d/ was used for /oʊ/ and a word with initial /l/ was used for /y/. The Chinese words containing the Mandarin vowels were "僻[pi] (out-of-the-way), 绿[ly] (green), 铺[pu] (store)，怕[pa] (fear)，派[paj] (send), 炮[paʊ] (canon), 配[pe] (match), 豆[doʊ] (bean)". They were inserted in the sentence frame "*zhe ge zi shi*________." (This character is ________.) in the production test.

2.1.2.2. English stimuli

The English vowels examined in this study were /i, ɪ, e, ɛ, æ, u, ʊ, o, ɒ, ʌ, aj, aʊ/. Words containing these target vowels, "pea, pit, pay, pet, pat, Pooh, put, dough, paw, pub, pie, pow", were inserted in the sentence frame "Now I say ______." to elicit the vowel production data.

Table 2-2: Mandarin and English stimuli

	Target vowels	Words containing target vowels	Sentence frame
Mandarin	/i, y, u, a, aj, aʊ, e, oʊ/	僻[pi], 绿[ly], 铺[pu], 怕[pa], 派[paj], 炮[paʊ],配[pe], 豆[doʊ]	*Zhe ge zi shi*_______. (This character is _____.)
English	/i, ɪ, e, ɛ, æ, u, ʊ, oʊ, ɒ, ʌ, aj, aʊ/[7]	pea, pit, pay, pet, pat, Pooh, put, dough, paw, pub, pie, pow	Now I say _________.

2.1.3. Recording

Participants were recorded in a sound-treated recording booth in the Phonetics Lab at Simon Fraser University using a digital recorder (PMD 670

[7] Acoustic data in the present study (2.2.1) showed that the vowel in English "pay" and Mandarin "[pej]" was a monophthong. In addition, the vowel in English "dough" and Mandarin "[doʊ]"was a diphthong. Henceforth, /e/ and /oʊ/ are used to represent the two vowels, respectively.

Marantz) and a microphone (SHURE KSM109). The recorder was set at a sampling rate of 44KHz and a resolution of 16-bit.

Mandarin monolinguals produced Mandarin stimuli only. English monolinguals produced English stimuli only. Mandarin-English bilinguals produced both English stimuli and Mandarin stimuli. When the Mandarin-English bilinguals were recorded, half of them produced the English stimuli first and Mandarin stimuli second; the other half proceeded in reverse order. Stimuli were presented to participants on cue cards. Each stimulus was presented 3 times in random order. The experimenter used English to give directions when English stimuli were recorded and Mandarin when Mandarin stimuli were recorded.

Before the recording started, participants signed a consent form, completed a background questionnaire (Appendix 1 and 2), and reviewed the stimulus lists. They were instructed to read the stimuli at a normal speaking rate. They were encouraged to correct any errors that they had made while recording.

In total, 1104 Mandarin tokens (8 vowels × 46 subjects× 3 repetitions) and 1620 English tokens (12 vowels × 45 subjects × 3 repetitions) were elicited (The subject pool contained 1 more Mandarin speaker than English speakers).

2.1.4. Acoustic analysis

The duration, F1, F2, F3 and F0 of each target vowel were measured using *Praat* (Version 4.) (Boersma & Weenink, 2005). *Praat* script, "label-vowel.praat" (Welby, 2003), was used to label vowel boundaries. The starting point of a vowel was where the periodicity first appeared in the waveform after the initial consonant. The ending point of the vowel was where F2 disappeared on the spectrogram (Munro, 1992). The pitch track was also displayed to help with the judgement of the starting and ending points of a vowel. The labelled vowels were then submitted to another *Praat* script (Shira, 2003)[8] for analysis. Formants were measured at 30% and 70% of the distance after the beginning of a vowel to address the possibility that some vowels, including both monophthongs and diphthongs, may have upward or downward formant movement (Nearey & Assmann, 1986; Munro, 1992). Vowel duration, F1, F2 and F3 were automatically transferred to a spreadsheet. The measurements obtained by the scripts were manually checked and corrected by referring to the spectrograms. To compensate for the gender differences, the F1, F2 and F3 values in Hertz were converted to the Bark scale using the following formula (Traunmüller, 1990):

$$b=26.8/\left(1+(1960/f)\right)-0.53,$$

in which b is in Bark, and f is in Hz.

The amount of formant movement was calculated using the formant values obtained at a distance of 70% from the beginning of a vowel minus the formant values obtained at a distance of 30% from the beginning of the same vowel. Positive values indicate upward formant movement, while negative values indicate downward formant movement. The upward or downward formant movement was calculated not only for diphthongs but also for monophthongs. The calculation of formant frequency movement is relevant to the present study

[8] The script was slightly modified to satisfy the needs of the present study.

for the following reasons. First, the present study examined the acoustic similarities and differences between Mandarin vowels and their English counterparts. If English monophthongs have significant formant frequency movement (Nearey & Assmann, 1986; Munro, 1992), it is reasonable to include formant frequency movement to achieve a comprehensive comparison between English and Mandarin vowels. Second, the present study also examined the Mandarin-English bilinguals' acquisition of English vowels. In addition to F1, F2 or F3, formant frequency movement should also be a good indicator of a bilingual's success or failure in English vowel acquisition. Finally, the present study examined the influence of English phonetic learning on Mandarin vowel production by Mandarin-English bilinguals. If such an influence exists, the acoustic characteristics of English vowels carried over to the production of Mandarin vowels should include not only F1, F2 or F3 but also formant frequency movement.

F0 was measured because of the concern that reading the target words in the same tone (Tone 4) may not eliminate the possibility that the Mandarin-English bilinguals might be different from the Mandarin monolinguals in tone. Before F0 was manually measured in *Praat,* vowel duration was normalized to adjust for differences in speaking rate (Wang, Jongman & Sereno, 2003). A vowel was stretched or shrunk in *SoundForge* (Version 8) to be the same duration as the median duration of the vowel across all speakers. F0 was measured at approximately 0%, 25%, 50%, 75%, and 100% point of the pitch contour (Wang et al., 2002). To accommodate gender difference, Fo was normalized using Rose's formula (1987):

$$F0_{norm} = (F0_i - meanF0)/s,$$

where $F0_i$ is the F0 of any given point of a pitch contour, *meanF0* is the mean of the frequencies of the five points for a given speaker, *s* is the standard deviation from the mean for a given speaker.

2.2. Results and discussion

This section presents and discusses the results of the production tests. Mandarin vowel production by Mandarin monolinguals is compared with English vowel production by English monolinguals; Mandarin-English bilinguals are compared with English monolinguals in the production of English vowels; and Mandarin-English bilinguals are compared with Mandarin monolinguals in the production of Mandarin vowels. An α value of .05 was the criterion for significance in all statistical analyses in this thesis.

2.2.1. Cross-language acoustic similarities in vowels

This section describes and discusses Mandarin vowel production by Mandarin monolinguals and English vowel production by English monolinguals. The goal is to determine the acoustic similarities between Mandarin vowels and their English counterparts.

2.2.1.1. Results

Duration data

The mean durations for vowels spoken by Mandarin monolinguals and English monolinguals are summarized in Table 2-3. A mixed design ANOVA, with vowel as a within-subjects factor (/a/(/ɒ/)[9], /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/) and language (Mandarin, English) as a between-subjects factor, revealed a main effect of language, [$F(1, 23) = 14.35$, $p < .01$]. English vowels tended to be longer than their Mandarin counterparts. Separate one-way ANOVAs were calculated on each of the vowels, /a/(/ɒ/), /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/. The results revealed significant group differences for /a/ (/ɒ/), [$F(1, 23) = 27.01$, $p < .001$], for /aj/, [$F(1, 23) = 25.38$, $p < .001$], for /aʊ/, [$F(1, 23) = 22.51$, $p < .001$], for /i/, [$F(1, 23) = 11.48$, $p < .01$], and for /oʊ/, [$F(1, 23) = 13.25$, $p < .01$]. In all cases, the English vowel was longer than its Mandarin counterpart. No significant group difference was found for /e/ and /u/.

[9] Enclosed in brackets is the English counterpart of Mandarin /a/.

Table 2-3: Mean durations (ms) and standard deviations of Mandarin vowels /a/, /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/ and their English counterparts, /ɒ/, /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/

Vowel	MonoM	MonoE
/a/ (/ɒ/)	225 (53)**	326 (43)**
/aj/	256 (67)**	378 (53)**
/aʊ/	233 (58)**	330 (42)**
/e/	269 (73)	313 (49)
/i/	232 (46)**	286 (31)**
/oʊ/	256 (58)**	334 (48)**
/u/	241 (63)	278 (47)

Note. Values enclosed in parentheses represent English vowels and standard deviations, respectively. $^{**}p < .01$

Spectral data

The mean F1, F2, and F3[10] values obtained at 30% distance from the beginning of a vowel (labelled F1a, F2a, and F3a) are listed in Appendix 7. An F1 × F2 plot depicting Mandarin vowels produced by Mandarin monolinguals and the English counterparts produced by English monolinguals is provided in Figure 2-1. The means in Bark of F1 and F2 movement values (labelled ΔF1 and ΔF2) are given in Appendix 8. Figure 2-2 illustrates schematically how /aj/, /aʊ/ and /oʊ/ differ in formant movement between Mandarin and English.

[10] F3 values were obtained for rounded vowels only.

Figure 2-1: Mean formant values (Bark) for Mandarin vowels produced by Mandarin monolinguals (n=13) and English vowels produced by English monolinguals (n=12)

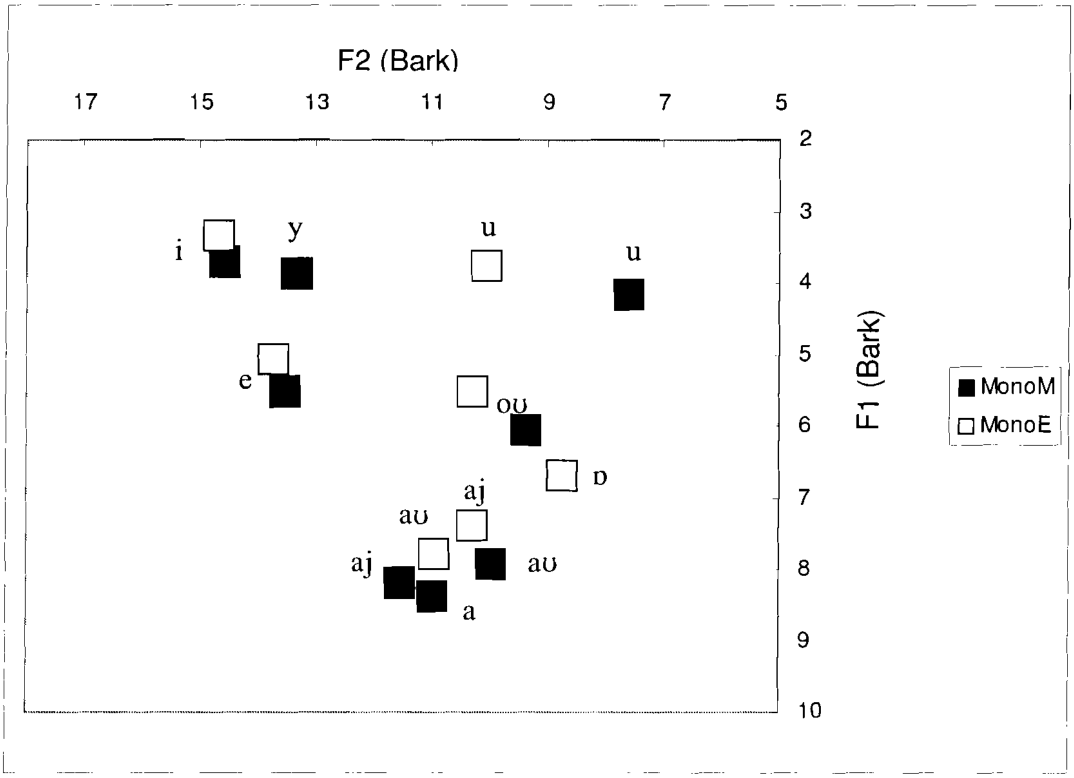
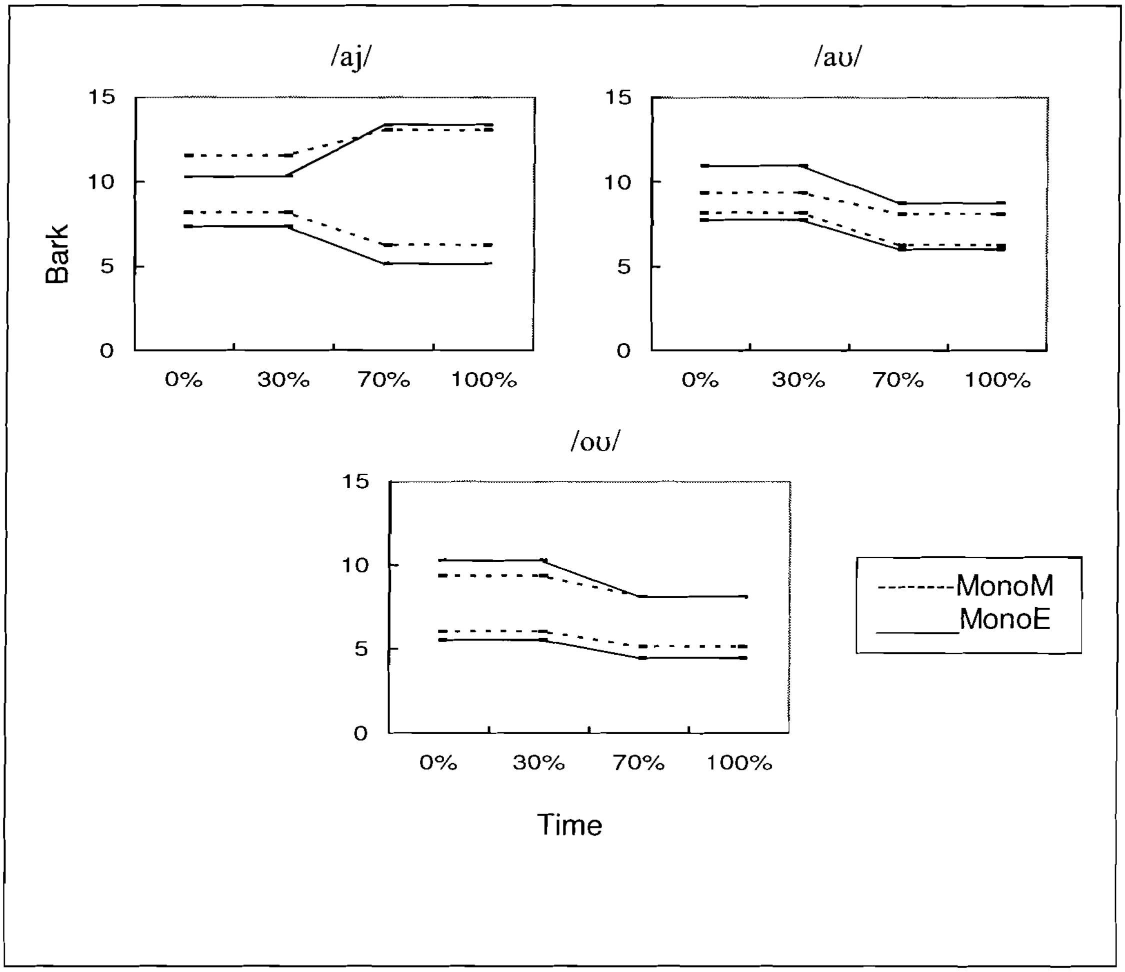


Figure 2-2: Schematic spectrograms representing the first and second formants of /aj/, /aʊ/ and /oʊ/ for Mandarin and English



Note: The formant values at 30% and 70% of the duration are used at 0% and 100% of the duration, respectively.

To compare the differences in spectral properties between Mandarin vowels and their English counterparts, a series of one-way ANOVAs were carried out on the mean values of F1a, F2a, F3a, ΔF1 and ΔF2 for each of the vowels, (/a/(/ɒ/), /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/). The acoustic differences, along with statistical results and phonetic consequences, are listed in Table 2-4.

Table 2-4: Acoustic differences between Mandarin vowels and their English counterparts

Mandarin Vowel	Difference in comparison with English counterpart	Statistical results	Consequences
/a/	Higher F1 Higher F2	$F(1, 23) = 27.27, p < .001$ $F(1, 23) = 53.80, p < .001$	Lower & more anterior
/aj/	Higher F1 Higher F2 smaller upward ΔF2	$F(1, 23) = 4.99, p < .05$ $F(1, 23) = 10.06, p < .01$ $F(1, 23) = 45.60, p < .001$	Lower, more anterior & less open
/aʊ/	Lower F2 smaller downward ΔF2	$F(1, 23) = 7.94, p < .05$ $F(1, 23) = 35.43, p < .001$	More posterior & less open
/e/	No difference	----	----
/i/	Larger downward ΔF1	$F(1, 23) = 8.10, p < .01$	More open
/oʊ/	Higher F1 lower F2 higher F3 less downward ΔF2	$F(1, 23) = 7.05, p < .05$ $F(1, 23) = 11.23, p < .01$ $F(1, 23) = 16.75, p < .001$ $F(1, 23) = 15.47, p < .01$	Lower, more posterior, less rounded & less open
/u/	Higher F1 lower F2	$F(1, 23) = 7.69, p < .05$ $F(1, 23) = 38.49, p < .001$	Lower & more posterior

Note. A dash indicates the information was not available.

The statistical analyses revealed that Mandarin /a/ had significantly higher F1, [$F(1, 23) = 27.27, p < .001$], and significantly higher F2, [$F(1, 23) = 53.80, p < .001$], than English /ɒ/. Therefore, Mandarin /a/ was lower and more anterior than its English counterpart. However, the two vowels did not differ significantly in either F1 movement or F2 movement.

For /aj/, the results showed that Mandarin /aj/ was significantly lower, [$F(1, 23) = 4.99, p < .05$] and more anterior, [$F(1, 23) = 10.06, p < .01$], than its English counterpart. Though the amount of F1 movement of Mandarin /aj/ and that of English /aj/ was non-significantly different, Mandarin /aj/ had significantly smaller amount of upward F2 movement than English /aj/, [$F(1, 23) = 45.60, p < .001$].

Compared with English /aʊ/, Mandarin /aʊ/ was significantly more posterior, [F (1, 23) = 7.94, $p < .05$]. It was also found that Mandarin /aʊ/ had significantly smaller amount of F2 movement than its English counterpart, [F (1, 23) = 35.43, $p < .001$]. No significant difference in height, roundedness and F1 movement were found between Mandarin /aʊ/ and its English counterpart.

For /e/, no significant differences between Mandarin and English were revealed in height, backness, amount of F1 and F2 movement,

There were no significant differences between Mandarin /i/ and English /i/ in height, backness and amount of F2 movement. However, Mandarin /i/ had a significantly larger amount of downward F1 movement than English /i/, [F (1, 23) = 8.10, $p < .01$].

Compared with English /oʊ/, Mandarin /oʊ/ was significantly lower [F (1, 23) = 7.05, $p < .05$], more posterior, [F (1, 23) = 11.23, $p < .01$], and less rounded, [F (1, 23) = 16.75, $p < .001$]. It also had significantly less downward F2 movement, [F (1, 23) = 15.47, $p < .01$].

Mandarin /u/ was significantly lower, [F (1, 23) = 7.69, $p < .05$], and more posterior than English /u/, [F (1, 23) = 38.49, $p < .001$]. However, Mandarin /u/ and its English counterpart did not differ significantly from each other in roundedness, amount of F1 and F2 movement.

Summary

All English vowels except /e/ and /u/ were significantly longer than their Mandarin counterparts. Moreover, all Mandarin vowels except /e/ were significantly different from their English counterparts in at least one acoustic dimension. Specifically, Mandarin /a/ and /aj/ were lower and more anterior than English /ɒ/ and /aj/[11], respectively. Mandarin /aj/ also had less upward F2 movement, which indicates it was less open as compared with English /aj/. Mandarin /aʊ/ was more posterior than English /aʊ/. It also had less downward F2 movement. Mandarin /i/ had more downward F1 movement than English /i/. Mandarin /oʊ/ was lower, more posterior, less rounded and less downward in F2 movement than English /oʊ/. Finally, Mandarin /u/ was lower and more posterior than English /u/.

2.2.1.2. Discussion

The results of the cross-language vowel duration measurement in the present study were inconsistent with Wang's study (1997) except for vowel /u/, in which both studies revealed a non-significant duration difference between Mandarin and English. While the present study revealed that a majority of the English vowels were longer than their Mandarin counterparts, Wang found that English had significantly longer duration only in /i/ and /e/. This inconsistency was probably due to the different phonetic contexts of the stimuli used in the two studies. In Wang's study, words containing the target English vowels were in a closed syllable (/bVt/) while words containing the target Mandarin vowels were in an open syllable (/bV/). In the present study, however, words containing the target English vowels that have Mandarin counterparts and words containing the

[11] The height and backness of diphthongs such as /aj/, /aʊ/, and /oʊ/ mentioned in this thesis were determined by the formant frequencies obtained at the first portion of the vowels (30%).

target Mandarin vowels were in an open syllable (/pV/. The vowel shortening effect in a closed syllable may account for the non-significant duration difference between Mandarin and English in most of the vowels in Wang's study.

Although Wang (1997) found that Mandarin /e/ was lower and more posterior than its English counterpart, the same spectral differences were not confirmed by the data in this study. Like Wang (1997), this study found that Mandarin /i/ was spectrally identical to English /i/ in height and backness. However, this study had an additional finding in that Mandarin /i/ had significantly downward F1 movement, which, to the best of my knowledge, has never been reported in the literature. The finding that Mandarin diphthongs /aj/, /aʊ/, and /oʊ/ had significantly less F2 movement was consistent with Wu's finding (1992) that Mandarin diphthongs such as /aj/ are less open than those of English. Mandarin /a/ was considerably lower and more anterior than English /ɒ/, which conforms to Wang's (1997) study. Also conforming to Wang's (1997) study are the results of Mandarin /oʊ/ and /u/. /oʊ/ was found to be lower and more posterior than its English counterpart and /u/ differed from English /u/ in backness and height.

The acoustic data of the present study suggest that Mandarin /e/ was temporally and spectrally identical to its English counterpart. The remaining Mandarin vowels all differed from their English counterparts in at least one acoustic dimension. However, it is difficult to determine cross-language similarities and differences solely on the basis of acoustic measurement, because of the various uncontrolled differences, such as stress, speaking rate, formality and vocal tract size that may complicate determining cross-language phonetic similarities (Flege, Bohn and Jang, 1997). Therefore, cross-language vowel similarities can not be adequately determined without a perception test. Nonetheless, the acoustic analysis in this section sets the basis for subsequent perception studies. Furthermore, the acoustic comparison of Mandarin and English vowels establishes the Mandarin and English norms for analysis of non-native vowel production by Mandarin-English bilinguals. Moreover, Mandarin vowels and English vowels have to be acoustically different to test L2 influence on L1.

2.2.2. Bilingual English vowel production

This section describes and discusses the extent to which Mandarin-English bilinguals differ from English monolinguals in English vowel production.

2.2.2.1. Results

Duration data

Table 2-5 displays the mean durations of English vowel productions by BiMH, BiML and MonoE. A repeated measures ANOVA, with vowel as within-subjects factor (/i/, /ɪ/, /e/, /ɛ/, /æ/, /u/, /ʊ/, /o/, /ɒ/, /ʌ/, /aj/, /aʊ/) and speaker group as between-subjects factor (BiMH, BiML, MonoE), revealed an effect of vowel by speaker group interaction, [$F(22, 42) = 1.58$, $p < .05$]. However, there was no main effect of speaker group. A series of one-way ANOVAs on the vowel data with speaker group (BiMH, BiML, MonoE) as between-subjects factor revealed no significant group difference in any vowel except /æ/, [$F(2, 42) = 3.37$, $p < .05$].

A post hoc Tukey test revealed a significant difference only between BiMH and MonoE [$p < .05$]. No significant difference was found between BiML and MonoE.

Table 2-5: Mean durations (ms) and standard deviations of English vowel production by Mandarin-English bilinguals of high L1 use (BiMH) (n=16), Mandarin-English bilinguals of low L1 use (BiML) (n=17) and English monolinguals (MonoE) (n=12)

Vowel	BiMH	BiML	MonoE
/ɒ/	287 (47)	292 (71)	326 (43)
/aj/	335 (56)	347 (82)	379 (53)
/aʊ/	290 (43)	307 (64)	330 (42)
/e/	315 (50)	328 (83)	313 (49)
/oʊ/	313 (43)	330 (72)	334 (48)
/ʌ/	141 (30)	151 (21)	138 (27)
/i/	280 (40)	278 (64)	286 (31)
/ɪ/	110 (24)	123 (26)	122 (29)
/u/	284 (38)	264 (76)	278 (47)
/ʊ/	117 (23)	128 (24)	122 (28)
/æ/	156 (34)*	174 (25)	185 (32)
/ɛ/	130 (27)	136 (30)	137 (29)

Note. Values enclosed in parentheses represent standard deviations. *$p < .05$.

Spectral data

The mean F1a, F2a, F3a (Bark) of English vowel productions by BiMH, BiML and MonoE and a summary of the mean ΔF1 and ΔF2 (Bark) of English vowel productions by the same speaker groups are listed in Appendix 9 and 10.

Figure 2-3 is a plot of vowel space of English vowel production by MonoE and BiMH. Figure 2-4 shows the vowel space of English vowel production by MonoE and BiML. Figure 2-5 illustrates schematically how the Mandarin-English bilinguals differ from MonoE in the formant movements of English /aj/, /aʊ/, /oʊ/, /ɪ/ and /ɛ/.

Figure 2-3: Mean formant frequencies (Bark) for English vowels produced by English monolinguals (MonoE) (n=12) and Mandarin-English bilinguals of high L1 use (BiMH) (n=16)

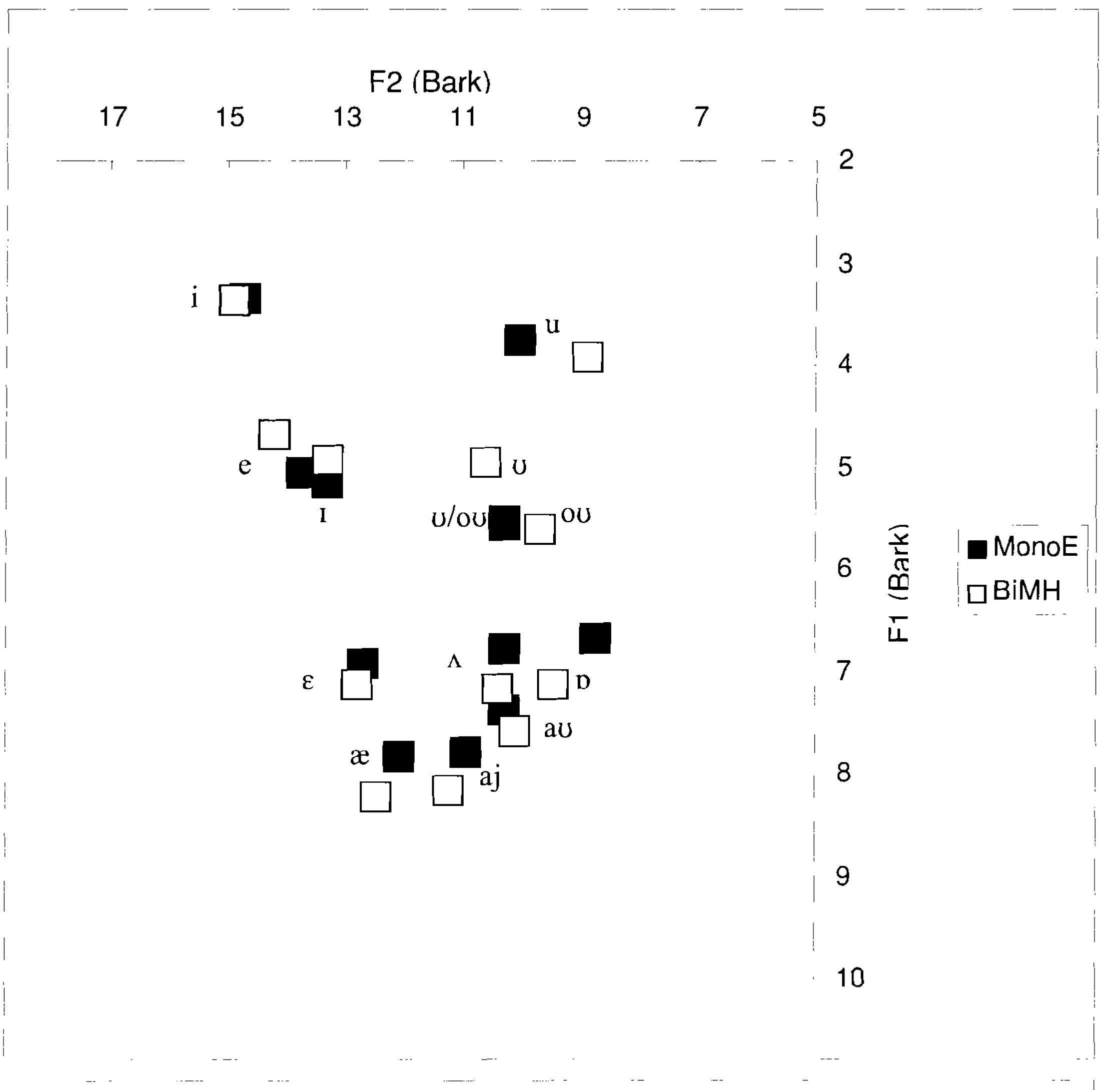


Figure 2-4: Mean formant frequencies (Bark) for English vowels produced by English monolinguals (MonoE) (n=12) and Mandarin-English bilinguals of low L1 use (BiML) (n=17)

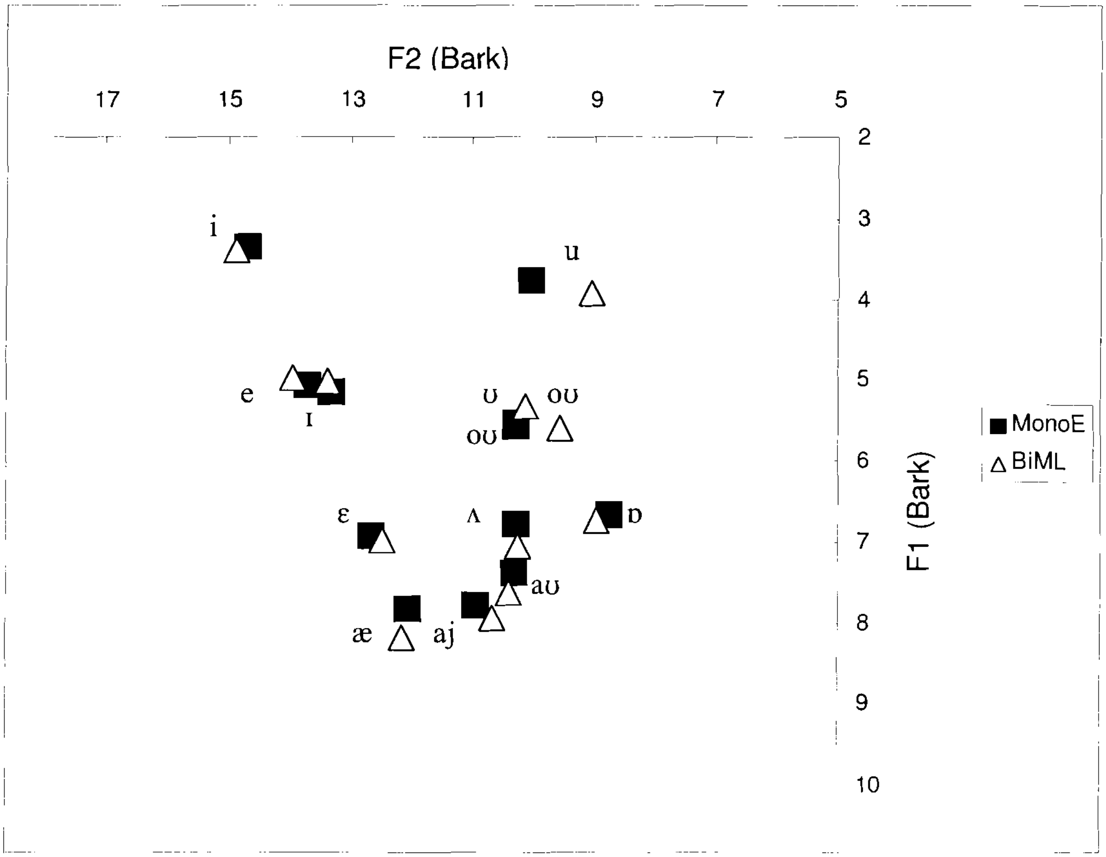
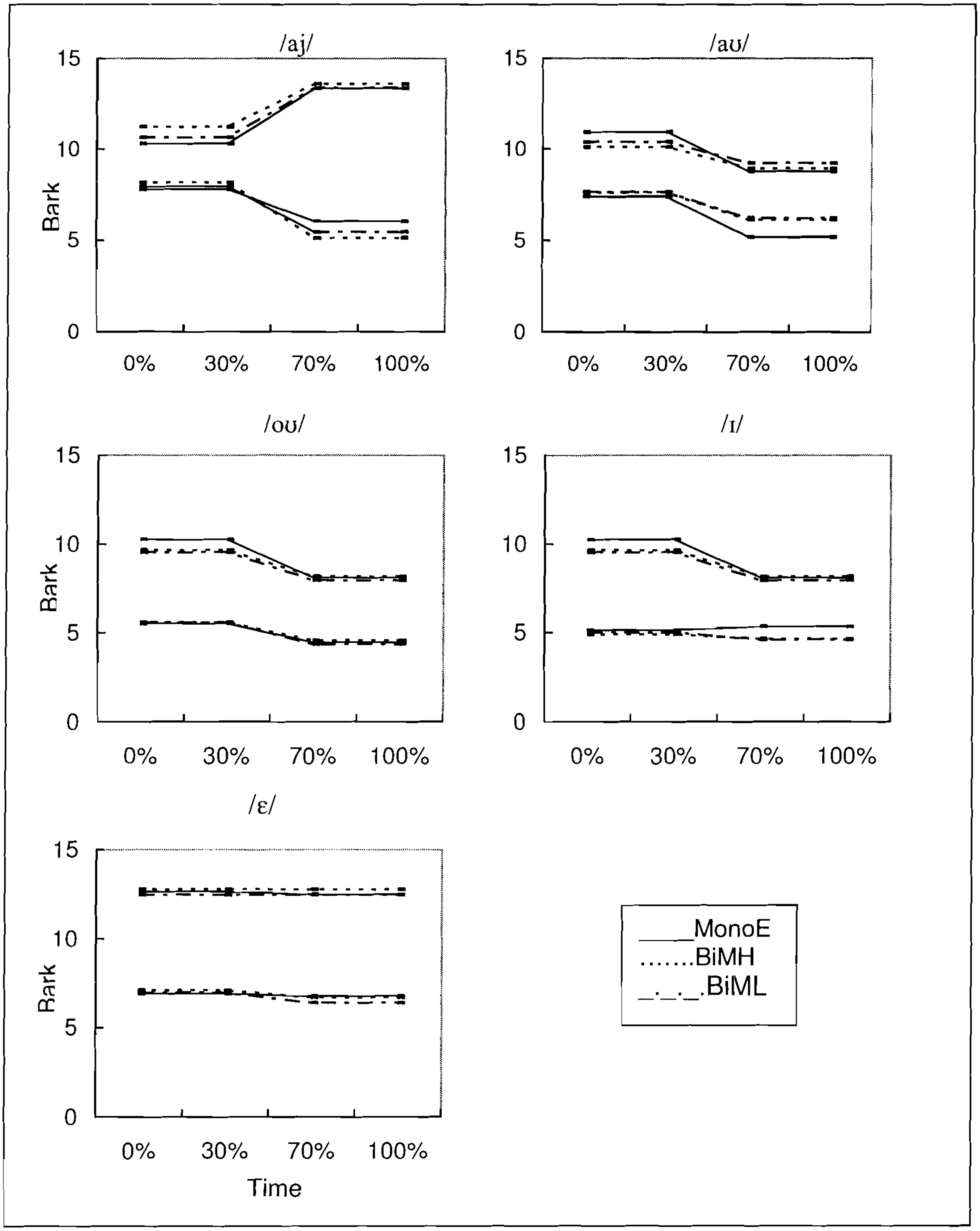


Figure 2-5: Schematic spectrograms representing the first and second formants of /aj/, /aʊ/, /oʊ/, /ɪ/ and /ɛ/ for English monolinguals (MonoE) (n=12), Mandarin-English bilinguals of high L1 use (BiMH) (n=16) and Mandarin-English bilinguals of low L1 use (BiML) (n=17)



Note: The formant values at 30% and 70% of the duration are used at 0% and 100% of the duration, respectively.

Separate one-way ANOVAs were calculated on the mean F1a, F2a, F3a, ΔF1 and ΔF2 for each of the 12 English vowels. The results showed that BiMH and BiML did not differ significantly from MonoE in any of the spectral properties of vowels /ɒ/, /e/, /ʌ/, /i/, /ʊ/ and /æ/. For vowels /aj/, /aʊ/, /oʊ/, /ɪ/, /u/, /ɛ/, however, the three groups differed significantly from each other in at least one dimension of the spectral properties. The acoustic differences, along with statistical results and phonetic consequences, are given in Table 2-6.

Table 2-6: Acoustic differences between Mandarin-English bilinguals (BiMH & BiML) and English monolinguals (MonoE) in English vowel production

English Vowel	Difference in comparison with MonoE	Statistical results	Consequences
/aj/	Less upward ΔF2 (BiMH)	$p < .05$	Less open
/aʊ/	smaller upward ΔF2 (BiMH & BiML)	$p < .01$	Less open
/oʊ/	less downward ΔF2	$p < .05$	Less open
/ɪ/	downward ΔF1 (BiMH & BiML) less upward ΔF2 (BiMH & BiML)	$p < .01$ (BiMH); $p < .001$ (BiML) $p < .05$ (BiMH); $p < .01$ (BiML)	Less open
/u/	lower F2 (BiMH) less downward ΔF2 (BiMH)	$p < .05$ $p < .01$	More posterior Less open
/ɛ/	More downward ΔF1 (BiML)	$p < .05$	More open

Note. A dash indicates the information was not available. "BiMH" and "BiML" enclosed in the brackets indicate which group the information applies to.

For the diphthongs /aj/, /aʊ/, /oʊ/, Mandarin-English bilinguals tended to have less F2 movement when compared to MonoE. For /aj/ [$F(2, 42) = 3.33$, $p < .05$], a post hoc Tukey test revealed a significant difference only between BiMH and MonoE [$p < .05$]. No significant difference was found between BiML and MonoE [$p = .443$].

For /aʊ/ [$F(2, 42) = 6.96$, $p < .01$], a post hoc Tukey test revealed a significant difference between BiMH and MonoE [$p < .01$], and between BiML and MonoE [$p < .01$].

For /oʊ/ [$F(2, 42) = 4.88$, $p < .05$], a post hoc Tukey test revealed a significant difference between BiMH and MonoE [$p < .05$] and between BiML and MonoE [$p < .05$].

For monophthongs /ɪ/, /u/, /ɛ/, the three groups differed significantly in at least one of the spectral properties of F2a, ΔF1, and ΔF2. For /ɪ/, a significant group difference was found in both ΔF1 [$F(2, 42) = 10.57$, $p < .001$] and ΔF2 [$F(2, 42) = 7.65$, $p < .01$]. While MonoE had an upward F1 movement (positive value), BiMH and BiML had a downward F1 movement. A post hoc Tukey test revealed a significant difference between BiMH and MonoE [$p < .01$], and between BiML and MonoE [$p < .001$]. Though the three groups all had the same direction of F2 movement (downward movement as indicated by the negative values), they differed significantly in the amount of F2 movement. Significantly less F2 movement was observed in BiMH [$p < .05$] and BiML [$p < .01$].

For /u/, a significant group difference in F2a [$F(2, 42) = 3.50$, $p < .05$], and ΔF2 [$F(2, 42) = 5.42$, $p < .01$], was revealed. A post hoc Tukey test showed that BiMH's /u/ was significantly more posterior than MonoE's /u/ [$p < .05$]. No such difference was found between BiML and MonoE [$p = .094$]. BiMH showed significantly less downward F2 movement as compared with MonoE and BiML [$p < .01$].

For /ɛ/, a significant group difference in F1 movement was observed, [$F(2, 42) = 3.70$, $p < .05$]. A post hoc Tukey test showed it was BiML [$p < .05$], rather than BiMH, that had significantly more downward F1 movement in comparison with MonoE.

Summary

With regard to duration, the two groups of Mandarin-English bilinguals did not differ from the English monolinguals except for the vowel /æ/, where BiMH had a significantly shorter duration than MonoE. Spectrally, the two groups of Mandarin-English bilinguals did not differ from the English monolinguals in the English vowels /ɒ/, /e/, /ʌ/, /i/, /ʊ/ and /æ/. The English vowels in which the Mandarin-English bilinguals differed from the English monolinguals were /aj/, /aʊ/, /oʊ/, /ɪ/, /u/, and /ɛ/. Compared with MonoE, less F2 movement in the English vowels /aj/ and /u/, and a more posterior position in English /u/, were observed for BiMH only. For English /aʊ/, /oʊ/ and /ɪ/, less F2 movement was observed for both BiMH and BiML. While F1 movement in the opposite direction (downward) was found for both BiMH and BiML in English vowel /ɪ/, more downward F1 movement was observed for BiML only in the English vowel /ɛ/.

2.2.2.2. Discussion

The Mandarin-English bilinguals showed a native-like duration pattern in the majority of the English vowels examined in this study. Duration did not seem to pose a big problem in the process of their acquisition of the English vowels. Traditionally, F1 and F2 are used to measure if an L2 English learner has successfully acquired an L2 English target vowel. If the same criterion were used to examine the successful acquisition of L2 English vowel acquisition here, it could be claimed that, in most cases, the Mandarin-English bilinguals in this study had successfully acquired the L2 English vowels. However, as Nearey & Assmann (1986) and Munro (1992) pointed out, the traditional approach might fail to reveal differences in formant movement. In fact, most of the significant group differences in this study were observed in either F1 or F2 movement. If differences in formant movement are taken into consideration, it can be claimed that the Mandarin-English bilinguals in this study had successfully acquired English vowels /ɒ/, /e/, /ʌ/, /i/, /ʊ/ and /æ/, but were not equally successful in the acquisition of the English vowels /aj/, /aʊ/, /o/, /ɪ/, /u/, and /ɛ/.

Consistent with previous studies (e.g., Flege, Fireda and Nozawa, 1997; Flege, Schirru and MacKay, 2003; Guion et al., 2000), an effect of L1 use on L2 was found in the production of /aj/ and /u/, in which BiMH rather than BiML differed from MonoE in either F2 movement or F2. However, such an effect was not found in the production of /aʊ/, /oʊ/, and /ɪ/, in which both BiMH and BiML were unsuccessful, suggesting that such vowels pose some difficulty for both groups of learners. Nevertheless, compared with BiML, BiMH had problems with far more vowels. Overall, the findings in the present study suggest that the more a bilingual uses his/her L1, the less accurate his/her L2 production is (e.g., Flege, Frieda and Nozawa, 1997; Flege, Schirru and MacKay, 2003).

2.2.3. Bilingual Mandarin vowel production

In this section, Mandarin-English bilinguals are compared with Mandarin monolinguals in the production of Mandarin vowels. As noted in chapter 1, this is the focus of the present study.

2.2.3.1. Results

Duration Data

Mean durations of Mandarin vowel productions by Mandarin-English bilinguals and Mandarin monolinguals are given in Table 2-7. A repeated measures ANOVA, with vowel as within-subjects factor (/a/, /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/, /y/), and speaker group as between-subjects factor (MonoM, BiMH, BiML), revealed a non-significant main effect of speaker group and a significant vowel by speaker group interaction, [$F(14, 43) = 2.66$, $p = .001$]. Separate one-way ANOVAs on each of the eight vowels with speaker group as a between-subjects factor (MonoM, BiMH, BiML) revealed no significant group difference in any of the vowels except /oʊ/. For /oʊ/, a main effect of group was observed, [$F(2, 43) = 3.55$, $p < .05$]. However, a post hoc Tukey test showed only a marginal difference between BiML and MonoM [$p = .051$]. No significant difference was found between BiMH and MonoM. In general, Mandarin-English bilinguals did not differ from Mandarin monolinguals in the mean durations of Mandarin vowels.

Table 2-7: Mean durations (ms) and standard deviations of Mandarin vowel production by Mandarin-English bilinguals of high L1 use (BiMH) (n=16), Mandarin-English bilinguals of low L1 use (BiML) (n=17) and Mandarin monolinguals (MonoM) (n=13)

Vowel	BiMH	BiML	MonoM
/a/	197 (51)	219 (57)	225 (53)
/aj/	238 (59)	270 (61)	256 (67)
/aʊ/	214 (47)	251 (66)	233 (58)
/e/	224 (55)	257 (67)	269 (73)
/i/	213 (48)	222 (59)	232 (46)
/oʊ/	215 (44)	213 (44)	256 (58)
/u/	204 (49)	211 (64)	241 (63)
/y/	257 (60)	277 (66)	280 (62)

Spectral Data

The mean values of F1a and F2a of all Mandarin vowels and of F3a of Mandarin /aʊ/, /oʊ/, /u/, /y/ as produced by both Mandarin monolinguals and Mandarin-English bilinguals are given in Appendix 11. The mean ΔF1 and ΔF2 of the Mandarin vowel productions by Mandarin monolinguals and Mandarin-English bilinguals are listed in Appendix 12. Figure 2-6 is a plot indicating the vowel space of Mandarin vowel productions by MonoM, BiMH and BiML. Figure

2-7 illustrates schematically how the Mandarin-English bilinguals differ from MonoM in the formant movements of /aj/ and /i/.

Figure 2-6: Mean formant frequencies (Bark) for Mandarin vowels produced by Mandarin monolinguals (MonoM) (n=13), Mandarin-English bilinguals of high L1 use (BiMH) (n=16) and Mandarin-English bilinguals of low L1 use (BiML) (n=17)

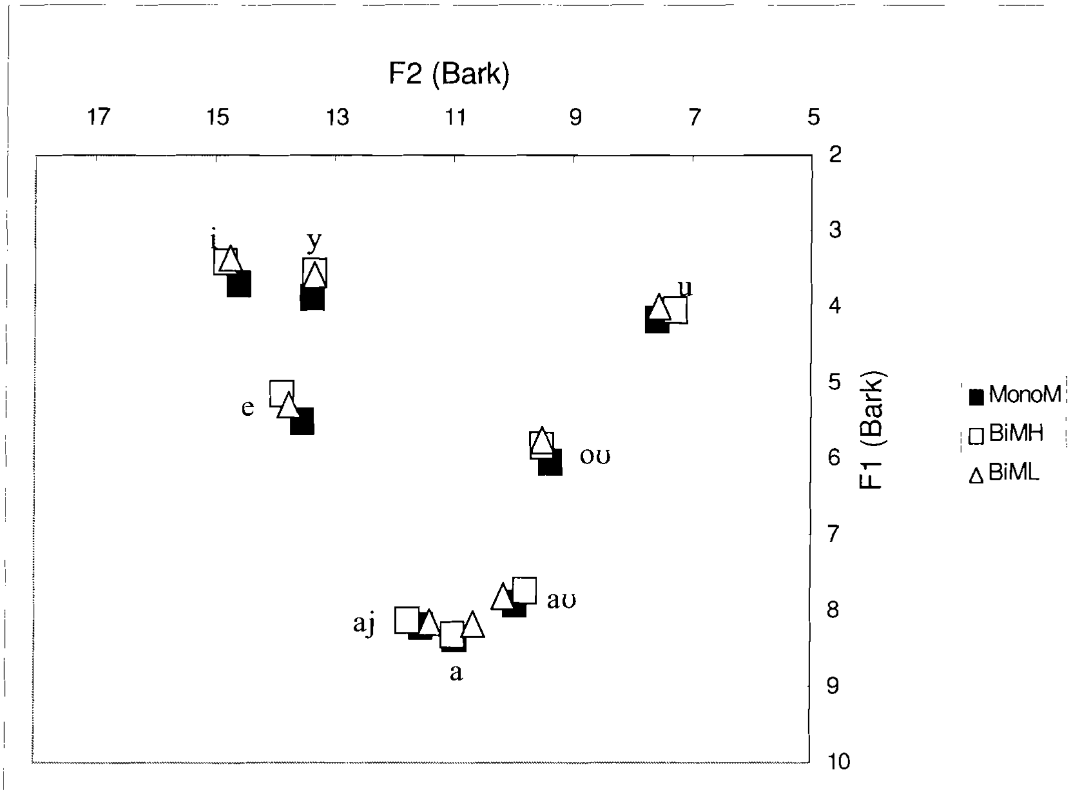
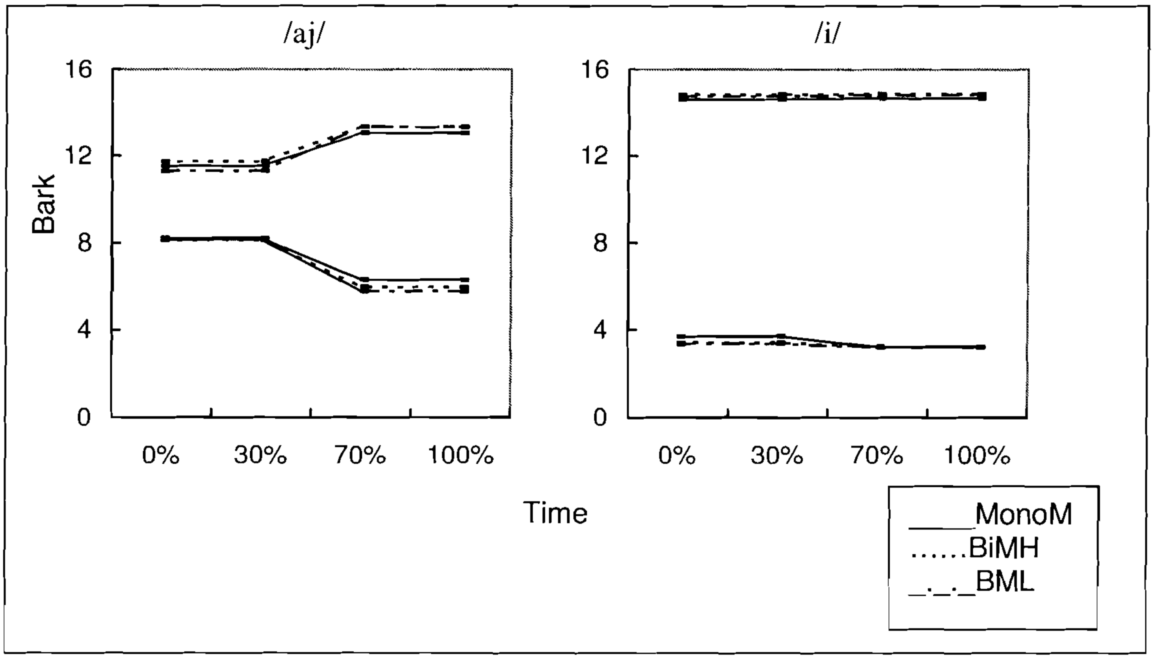


Figure 2-7: Schematic spectrograms representing the first and second formants of /aj/ and /i/ for Mandarin monolinguals (MonoM) (n=13), Mandarin-English bilinguals of high L1 use (BiMH) (n=16) and Mandarin-English bilinguals of low L1 use (BiML) (n=17)



Note: The formant values at 30% and 70% of the duration are used at 0% and 100% of the duration, respectively.

Separate one-way ANOVAs were calculated on the mean values of F1a, F2a, F3a, ΔF1 and ΔF2 for each of the Mandarin vowels with group (MonoM, BiMH, BiML) as a between-subjects factor. The analysis revealed no significant group difference in the mean values of F1a, F2a, and F3a in any of the Mandarin vowels, indicating that BiMH, BiML and MonoM did not differ significantly from each other in vowel height, backness and rounding. However, a significant group difference was found in the mean values of ΔF1 in vowel /i/, [$F(2, 43) = 4.93$, $p < .05$], and the mean values of ΔF2 in vowel /aj/, [$F(2, 43) = 4.40$, $p < .05$]. For ΔF1 in /i/, a post hoc Tukey test showed that, compared with MonoM, both BiMH [$p < .05$] and BiML [$p < .05$] had significantly less downward movement. For ΔF2 in /aj/, a post hoc Tukey test revealed that BiML had significantly more upward movement than MonoM [$p < .05$]. BiMH did not differ significantly from MonoM in this respect. No significant group difference in ΔF1 and ΔF2 was found for other Mandarin vowels.

Tone data

As noted in 2.1.4, the measurement of F0 was motivated by the concern that the Mandarin-English bilinguals might deviate from Mandarin monolinguals in tone when the words containing the target vowels were produced.

Separate one-way repeated measures ANOVAs were conducted comparing MonoM, BiMH, and BiML's Tone 4 productions for each of the eight Mandarin vowels, with position of measurement (0%, 25%, 50%, 75%, 100%) as within-subjects factor, and group (MonoM, BiMH and BiML) as between-subjects factor. No significant effect was observed for group or group by position interaction in the Tone 4 productions of any of the Mandarin vowels, indicating that all three groups produced the tone homogenously.

Summary

The acoustic analysis did not reveal any significant differences between the Mandarin-English bilinguals and the Mandarin monolinguals in vowel duration, formant frequencies (F1, F2 or F3), and Tone 4 contours in any of the target Mandarin vowels. The only significant group difference was found in formant movement for the vowels /i/ and /aj/. For vowel /i/, both BiMH and BiML had significantly less downward F1 movement than MonoM, approximating the pattern of English /i/. For /aj/, BiML, but not BiMH, had significantly more upward

F2 movement than MonoM, indicating that BiML's /aj/ was more open than MonoM's /aj/ and thus approximates English /aj/.

2.2.3.2. Discussion

BiMH and BiML's less downward F1 movement in Mandarin /i/ and BiML's more upward F2 movement in Mandarin /aj/ can be interpreted as the result of L2 vowel learning. As mentioned in 2.2.2, neither BiMH nor BiML differed from MonoE in F1 movement or other acoustic dimensions of English /i/, indicating successful acquisition of this English vowel. In the case of English /aj/, however, BiML did not differ from MonoE in F2 movement and other acoustic dimensions while BiMH did, indicating BiML's successful acquisition of English /aj/. It is hypothesized that the Mandarin-English bilinguals who successfully acquired English /i/ (BiMH and BiML) and /aj/ (BiML) carried over the stable F1 movement of English /i/ and the more upward F2 movement of English /aj/ to the production of Mandarin /i/ and /aj/, respectively. In sum, consistent with previous studies (Flege, 1987, 1995, 2003; Guion, 2003), the findings in the present study indicate that successful acquisition of L2 vowels plays a role in L1 vowel modification.

It must be noted, however, that the effect of L2 vowel learning on L1 vowel production did not show up in all L1 vowels whose L2 counterparts have been successfully acquired. For example, the acoustic analysis in 2.2.2 revealed that Mandarin-English bilinguals did not differ from English monolinguals in duration and spectral properties of English /ɒ/ and /e/, indicating their successful acquisition of these two vowels. However, no corresponding modification of

Mandarin vowels /a/ and /e/ was observed of the Mandarin-English bilinguals. It is hypothesized that cross-language vowel similarity is a contributing factor (e.g., Williams, 1980; Flege, 1995, 2003; Baker & Trofimovich, 2005). As can be seen in Figure 2-1 of 2.2.1, Mandarin /a/ and English /ɒ/ were the least similar acoustically among all target vowels. Further evidence of their dissimilarity is available in Table 3-2 of the following chapter, in which Mandarin /a/ and English /ɒ/ received the lowest and second lowest rating, respectively. Since Mandarin /a/ and English /ɒ/ are different from each other to a larger extent than other vowels, Mandarin-English bilinguals may categorize them as belonging to two different categories after being exposed to them for an extended period of time. As a result, Mandarin /a/ was not affected by English /ɒ/. With regard to Mandarin and English /e/, the acoustic analysis in 2.2.1 revealed that they did not differ in any of the dimensions measured in the present study and were therefore "identical." Given the identical nature of these two vowels, no effect of English /e/ on Mandarin /e/ should be expected. In sum, the effect of L2 phonetic learning on an L1 can not be demonstrated if an L1 phonetic segment is too dissimilar (e.g., Mandarin /a/ and English /ɒ/) or identical to its L2 counterpart (e.g., Mandarin /e/ and English /e/).

The effect of the amount of L1 use was revealed in Mandarin-English bilinguals' production of Mandarin /aj/, where BiML had significantly more upward F2 movement than BiMH when compared with MonoM.

Despite the successful category formation of English /i/ and /aj/ by Mandarin-English bilinguals, the SLM's prediction of "category dissimilation" was not observed in the present study. Instead of deflecting away from the English monolingual's English /i/ and /aj/ to keep a contrast between Mandarin vowels and their English counterparts, Mandarin-English bilinguals' Mandarin /i/ and /aj/ became similar to English /i/ and /aj/ (e.g., less downward F1 movement in /i/ and more upward F2 movement in /aj/). It seems that "category assimilation" occurs even though a bilingual has successfully established an L2 sound.

To summarize, Mandarin-English bilinguals, especially those with low L1 use, have modified the categories of some L1 Mandarin vowels (e.g., Mandarin /i/ and /aj/) as a result of L2 vowel learning. Yet, it is not known if the acoustic differences between Mandarin-English bilinguals and Mandarin monolinguals as revealed in the production test are salient in perception. In addition, although the measurement of acoustic dimensions such as duration, formant frequencies, formant movement and tone does not reveal a difference between Mandarin-English bilinguals and Mandarin monolinguals in the production of some vowels (e.g., /i/, /y/, /aʊ/, /e/, /oʊ/), it does not necessarily mean that Mandarin-English bilinguals do not have an accent in their L1 vowel production. Other acoustic dimensions that are not measured in the production test (e.g., voice quality) may contribute to Mandarin-English bilinguals' accented L1 vowel production. Moreover, it is not clear if the Mandarin-English bilinguals whose L1 vowel production deviates from L1 monolingual norms tend to have good L2 vowel proficiency, or whether the Mandarin-English bilinguals who have poor L2 vowel proficiency do not deviate from L1 monolingual norms. In order to address these issues, perception tests were also administered. These are discussed in the following chapter.

CHAPTER 3: GOODNESS RATING OF MANDARIN-ENGLISH BILINGUALS' VOWEL PRODUCTION

This chapter describes and discusses two goodness rating experiments. The first experiment addresses three main issues, namely, cross-language vowel perceptual similarity, Mandarin-English bilinguals' accent in L1 vowel production and the acoustic properties contributing to this accent. The second experiment evaluates the correlation between Mandarin-English bilinguals' English vowel pronunciation proficiency and their Mandarin vowel pronunciation proficiency.

3.1. Experiment 1

3.1.1. Methodology

3.1.1.1. Listeners

Twenty Mandarin listeners and 20 English listeners were initially recruited for the perception test. Four Mandarin listeners did not qualify due to their first language being Taiwanese. Four English listeners were excluded because they had learned Mandarin or were proficient speakers of other languages (e.g., French, Spanish). In the end, 16 Mandarin listeners (7 males and 9 females) and 16 English listeners (6 males and 10 females), all of whom reported normal hearing, participated in the perception test.

All Mandarin listeners were from Taipei, Taiwan. At the time of the study, they were either adult ESL students studying at the Harbour Centre campus of

Simon Fraser University, recent immigrants or short-term visitors. They had all received a university education in Taiwan. Their ages ranged from 20-39 years (mean=29). Their LOR ranged from half a month to 3 months (mean=2 months). In general, their English speaking proficiency was low. Twelve participants reported "not fluent" in English speaking; the other four reported "somewhat fluent." They all reported having some knowledge of Taiwanese. However, their spoken Taiwanese was generally "not fluent" (12 cases of "not fluent" and 4 cases of "somewhat fluent"). No one reported having knowledge of Hakka. The ESL students studying at Simon Fraser University completed the perception test in a quiet room at Harbour Centre campus. The recent immigrants and visitors were tested in public libraries across the Lower Mainland, British Columbia.

The English listeners were recruited from the undergraduate population at the Burnaby campus of Simon Fraser University. Their ages ranged from 18-26 years (mean=21 years). They were all born and raised in Anglophone regions of Canada west of Quebec. None of them had background in Mandarin. They completed the test in the Phonetics Lab of Simon Fraser University.

3.1.1.2. Stimuli

The vowel tokens collected in the production tests were used for this experiment. For English, only "pea, Pooh, pie, paw, pow, pay, dough", each of which had a Mandarin counterpart, were selected. In total, there were 104 Mandarin tokens by Mandarin monolinguals (8 vowels ×13 subjects × 1 repetition), 264 Mandarin tokens by Mandarin-English bilinguals (8 vowels ×33 subjects × 1 repetition) and 84 English tokens by English monolinguals (7 vowels × 12 subjects × 1 repetition).

There is no commonly agreed-upon criterion for selecting stimuli from among multiple repetitions in a perception test. In the present study, the second repetitions[12] of the target words in the production test were selected. Based on intuition and common sense, the first repetition of a stimulus may not be as stable as latter repetitions. On the other hand, the last repetition may have more practice effect (Murray Munro, personal communication, 2007).

To eliminate the effect of the initial consonants on vowel perception, the initial consonants of the target words were held constant through digital editing. Two Mandarin monolingual listeners (1 male and 1 female) and 2 English monolingual listeners (1 male and 1 female) were randomly selected from the subject pool. These selected speakers' initial consonants of the target words were edited out and were used to replace the initial consonants of all speakers on the basis of language (Mandarin, English) and gender (male, female). The editing was conducted using *Audacity* (Version 1.2.6). The average RMS intensity (loudness) of these edited words was then normalized to -16dB in *Sound Forge* (Version 8.0). A native Mandarin speaker from Taiwan evaluated the naturalness of these edited words and judged all tokens as natural.

[12] As noted in chapter 1, there are three repetitions for each stimulus in the production test. If a second repetition is not usable due to poor recording quality, the first or third repetition is selected instead.

3.1.1.3. Procedure

Stimuli were divided into two blocks, in which Mandarin stimuli and English stimuli were mixed, and were presented to listeners for goodness rating via E-Prime 1.0 (Psychology Software Tools) on a laptop computer (Compaq Presario V2000). For the Mandarin listeners, the first block consisted of 241 stimuli. There were 49 English stimuli by 7 English monolinguals (7 vowels × 7 subjects ×1 repetition), 56 Mandarin stimuli by 7 Mandarin monolinguals (8 vowels ×7 subjects × 1 repetition) and 136 Mandarin stimuli by 17 Mandarin-English bilinguals (8 vowels × 17 subjects × 1 repetition). The second block included 211 stimuli, among which there were 35 English stimuli by 5 English monolinguals (7 vowels × 5 subjects ×1 repetition), 48 Mandarin stimuli by 6 Mandarin monolinguals (8 vowels × 6 subjects ×1 repetition) and 128 Mandarin stimuli by 16 Mandarin-English bilinguals (8 vowels × 16 subjects ×1 repetition). Mandarin listeners were instructed to rate the goodness of the word they heard on a 7-point scale, with "1" being the worst and "7" the best exemplar of the Mandarin target word (see Appendix 3 for detailed oral instruction). For English listeners, the Mandarin stimuli containing vowel /y/ were excluded, because Mandarin /y/ does not have an English counterpart. As a result, the first block had a total of 217 stimuli, comprising 24 stimuli (1 vowel × 24 Mandarin speakers) less than the total number of stimuli for Mandarin listeners. The second block had a total of 189 stimuli, being 22 stimuli (1 vowel × 22 Mandarin speakers) less than the total number of stimuli for Mandarin listeners. The English listeners were instructed to complete the same task as the Mandarin listeners did, with "1" being the worst and “7” the best exemplar of the English target word (see Appendix 4 for detailed oral instruction).

Before the experiment started, listeners completed a consent form and a questionnaire. All listeners used headphones (Technics RP-HT400). The experiment began with a practice session to familiarize listeners with the test procedure. The inter-trial interval was set to be 4 seconds. Stimulus presentation in the actual test was counter-balanced. Half of the listeners listened to Block 1 first, and then listened to Block 2. The other half of the listeners proceeded in reverse order. Mandarin listeners completed Block 1 in 24 minutes and Block 2 in 21 minutes. English listeners took 20 minutes and 19 minutes respectively to complete blocks 1 and 2. All participants had a 5-minute break between the two blocks. When the listeners finished the test, they completed a debriefing questionnaire (Appendix 5 and 6).

3.1.2. Results

The perceptual similarities between Mandarin vowels and their English counterparts, and differences between Mandarin monolinguals and Mandarin-English bilinguals in rating scores are presented in the following sections.

3.1.2.1. Cross-language vowel perceptual similarities

As noted in 3.1.1.3, English vowel production by English monolinguals was included in the perception test for Mandarin listeners. Similarly, Mandarin vowel production by Mandarin monolinguals was included in the perception test for English listeners. Mandarin listeners in the present study had to rate the goodness of these English vowels with reference to the prototypical Mandarin vowels existing in their minds. English listeners had to do the same task by referring to the prototypical English vowels. The inclusion of the English and Mandarin stimuli made it possible to examine cross language (between English and Mandarin) vowel perceptual similarity.

Inter-rater reliability

Inter-rater reliability estimates by vowel were computed for Mandarin listeners and English listeners by using Cronbach's α (Cronbach, 1951). The results are provided in Table 3-1. The cut-off point of Cronbach's α is .70, above which inter-rater reliability is acceptable (Nunnaly, 1978). An acceptable inter-rater reliability was observed for /aj/, /e/, /oʊ/ and /u/ in both groups of listeners. However, the Cronbach's values for English /ɒ/ and Mandarin /a/, English vowel /aʊ/ and Mandarin vowel /i/ failed to meet the criterion for acceptable reliability. The ratings assigned to the vowels having unacceptable inter-rater reliability must be interpreted with caution.

Table 3-1: Mean inter-rater reliability scores for Mandarin listeners (n=16) judging MonoE and English listeners (n=16) judging MonoM

Vowel	Cronbach's α for Mandarin Listeners' judgement of MonoE	Cronbach's α for English Listeners' judgement of MonoM
/ɒ/(/a/)	.59	.48
/aj/	.71	.88
/aʊ/	.62	.90
/e/	.90	.76
/i/	.90	.48
/oʊ/	.77	.74
/u/	.86	.71

Goodness rating scores

The rating scores assigned to the English monolinguals' productions of the English vowels by the Mandarin listeners and those assigned to the Mandarin monolinguals' productions of the Mandarin vowels by the English listeners are listed in Table 3-2.

Table 3-2: Mean rating scores assigned to English monolinguals by Mandarin listeners (n=16) and rating scores assigned to Mandarin monolinguals by English listeners (n=16) on a 7-point scale where 1= "bad" and 7= "good"

Vowel	English vowels/ Mandarin listeners	Mandarin vowels/English listeners
/i/	3.93 (.99)	4.39 (.87)
/e/	3.48 (.98)	4.69 (.61)
/aj/	3.05 (.54)	4.49 (.86)
/u/	2.94 (.78)	4.32 (.46)
/oʊ/	2.66 (.61)	3.65 (.25)
/ɒ/(/a/)	2.42 (.45)	2.07 (.53)
/aʊ/	2.33 (.41)	3.65 (1.02)

Note. Values enclosed in parentheses represent standard deviations.

The Mandarin listeners' scores descended in the order of /i/, /e/, /aj/, /u/, /oʊ/, /ɒ/, /aʊ/, showing different extents of similarity to the Mandarin vowel counterparts. The three vowels to which English listeners assigned higher scores were also /i/, /e/, /aj/. In addition, both Mandarin listeners and English listeners assigned the lowest scores to /ɒ/(/a/) and /aʊ/. It has to be noted, though, that the rating scores for Mandarin /a/ and English /ɒ/ must be interpreted with caution due to the unacceptable inter-rater reliability. Although the Cronbach's α for the ratings of Mandarin /i/ failed to meet the criterion of an acceptable inter-rater reliability, the inter-rater reliability for the ratings of English /i/ was acceptable. Based on the ratings of English /i/, it may be argued that Mandarin /i/, /e/, /aj/ and their English counterparts are highly similar. Although the Cronbach's value for the ratings of English /aʊ/ was unacceptable, the inter-rater reliability for the ratings of Mandarin /aʊ/ was acceptable. Based on the ratings of Mandarin /aʊ/, it may be argued that Mandarin /aʊ/ and its English counterpart are among the vowels that are the least similar between English and Mandarin.

Relating the production data in 2.2.1 and the perceptual data in this section, Mandarin /e/ and its English counterpart may be considered "identical", because the production data did not reveal any difference in any acoustic dimension between Mandarin /e/ and its English counterpart and it was rated the best by English listeners and the second best by Mandarin listeners. Although the acoustic analysis showed Mandarin /i/ had a significant downward F1 movement, the perceptual analysis showed it was highly similar to its English counterpart. Thus, Mandarin /i/ may also be considered "identical" to its English counterpart, at least perceptually. The remaining Mandarin vowels had at least two acoustic dimensions that differed from their English counterparts. Thus, it is not surprising that they were rated worse than /i/ and /e/. These remaining vowels may be considered "similar" to their English counterparts. However, they are similar to their English counterparts to different degrees, with /aj/, /u/, and

/oʊ/ being more similar to their English counterparts than /a/ and /aʊ/ were. In other words, Mandarin /a/ and /aʊ/ were the least similar to their English counterparts.

3.1.2.2. Rating scores assigned to Mandarin-English bilinguals

Presented in the following subsections are inter-rater reliability, group differences, individual differences and the relation between individual differences and acoustic data. As can be recalled in 3.1.1.3, the English stimuli produced by the English monolinguals were included in the perception test for both Mandarin listeners and English listeners. The inclusion of the English stimuli served as a reference point when English listeners rated the goodness of Mandarin vowel production by Mandarin monolinguals and Mandarin-English bilinguals. It was not the focus of the perception test. The main interest of the perception test was to compare Mandarin-English bilinguals with Mandarin monolinguals. Thus, English monolinguals' perception data were excluded from analysis in this section. Only the perception data for the two groups of Mandarin-English bilinguals (BiMH and BiML) and those of the Mandarin monolinguals (MonoM) were submitted for the analysis of inter-rater reliability and for statistical analyses.

Inter-rater reliability

Inter-rater reliability estimates by vowel were computed for Mandarin listeners and English listeners by using Cronbach's α. The results are provided in Table 3-3. The values for Mandarin listeners are all above.70, a cut-off point for acceptable reliability (Nunnaly, 1978). The values for English listeners are above .70 in all vowels except /a/ and /i/. Since English listeners did not quite agree with each other on the rating of /a/ and /i/, the rating scores assigned by English listeners to these two vowels were excluded from further analysis.

Table 3-3: Mean inter-rater reliability scores for Mandarin listeners (n=16) and English listeners (n=16)

Mandarin Vowel	Cronbach's α for Mandarin Listeners	Cronbach's α for English Listeners
/a/	.87	.56
/aj/	.87	.89
/aʊ/	.86	.88
/e/	.85	.81
/i/	.75	.47
/oʊ/	.80	.78
/u/	.79	.74
/y/	.76	----

Note. Dash indicates the value was not relevant as /y/ is nonexistent in English.

Group differences

Rating scores assigned by Mandarin listeners

The mean ratings for each vowel assigned by Mandarin listeners to each speaker group (pooled across listeners) is given in Figure 3-1. The general tendency is that MonoM received the highest ratings and MonoE the lowest ratings, with the two groups of Mandarin-English bilinguals receiving intermediate ratings. It is also observed that, in most cases, BiML's ratings were lower than those of BiMH.

A Shapiro-Wilks test for normality indicated that the rating scores assigned to all three speaker groups (MonoM, BiMH, BiML) were normally distributed in the vowels /a/, /aʊ/, /e/, /i/, and /oʊ/. However, significant deviations from normality were observed in the rating scores assigned to MonoM in /aj/ (Shapiro-

Wilks = .782, $p < .01$), BiML in /u/ (Shapiro-Wilks = .879, $p < .05$) and BiMH in /y/ (Shapiro-Wilks = .877, $p < .05$). For the normally distributed data, parametric tests (ANOVA or *t*-test) were conducted. For those that deviated from normality, a nonparametric test (Mann-Whitney U) was conducted.

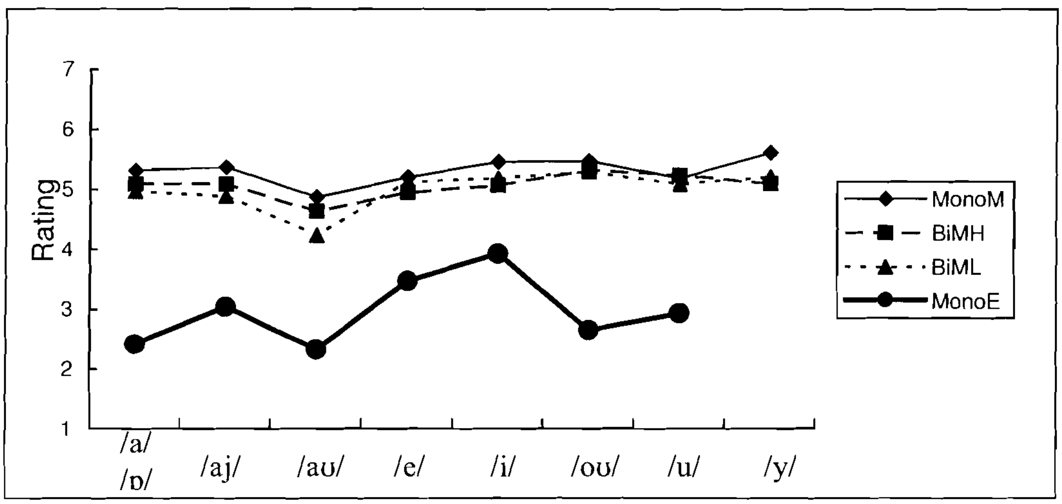
A repeated measures ANOVA was conducted on the rating scores assigned to each of the three speaker groups for /a/, /aʊ/, /e/, /i/, /oʊ/ with speaker group as a between-subjects factor (MonoM, BiMH, BiML). The analysis revealed neither a significant main effect of speaker group [$F(2, 43) = 1.10$, $p = .34$] nor significant vowel by speaker group interaction [$F(2, 43) = .77$, $p = .630$], indicating a non-significant difference between MonoM, BiMH, and BiML in the rating scores assigned to /a/, /aʊ/, /e/, /i/, and /oʊ/.

For the rating scores assigned to /aj/, Mann-Whitney independent samples test (U) revealed a non-significant difference between MonoM and BiMH and between MonoM and BiML. For the rating scores assigned to /u/, two separate analyses were conducted. Mann-Whitney U test was conducted on the rating scores of MonoM and BiML due to the significant deviation from normality in the rating scores of BiML. Since the rating scores of BiMH were normally distributed, an independent samples *t*-test (two-tailed) was conducted on the rating scores of MonoM and BiMH. Neither the Mann-Whitney U test nor the *t*-test revealed a significant between group effect, indicating no significant group difference between the Mandarin-English bilinguals and the Mandarin monolinguals in the rating scores for /u/. For the same reason, two parallel tests were conducted on the rating scores assigned to vowel /y/. Both the Mann-Whitney U test [$Z = -2.42$,

$p < .05$] and the t-test [$t\,(28) = 2.46$, $p < .05$] showed a significant between group effect.

In summary, despite the trend that the rating scores assigned to the Mandarin-English bilinguals were intermediate between the Mandarin monolinguals and the English monolinguals, a significant between-group difference of rating scores was observed only in /y/, a vowel nonexistent in English.

Figure 3-1: Mean Ratings assigned by Mandarin listeners (n=16)



Rating scores assigned by English listeners

The mean ratings for each vowel assigned by English listeners to each speaker group (pooled across listeners) are given in Figure 3-2. MonoE tended to receive the highest ratings while MonoM tended to get the lowest ratings. BiMH did not differ much from MonoM in the mean ratings of all vowels except /aʊ/, in which BiMH was lower than MonoM. BiML had intermediate ratings in all vowels except /i/. In this vowel, BiML, BiMH and MonoM received similar rating scores.

An examination of the rating scores assigned by English listeners indicated a significant deviation from normality in the rating scores of BiMH in vowel /aʊ/ (Shapiro-Wilks = .806, $p < .01$). All rating scores of the three Mandarin speaker groups and the remaining vowels were normally distributed[13]. Accordingly, an independent samples Mann-Whitney U test was conducted on the rating scores of MonoM and BiMH in /aʊ/. An independent samples *t*-test was conducted on the rating scores of MonoM and BiML in the same vowel. A one-way repeated measures ANOVA was conducted on the rating scores of the three Mandarin speaker groups for the remaining vowels[14].

The independent samples Mann-Whitney test indicated that BiMH and MonoM differed significantly in the rating scores of /aʊ/, [Z = - 2.33, $p < .05$]. The independent samples *t*-test revealed a marginal difference between BiML and MonoM, [$t(28) = -1.92$, $p = .065$]. The one-way repeated measures ANOVA, with vowel (/aj/, /e/, /oʊ/, /i/) as within-subjects factor and speaker group as between-subjects factor (MonoM, BiMH, BiML), showed a significant main effect of speaker group, [$F(2, 43) = 6.47$, $p < .01$], but a non-significant vowel by group interaction, [$F(2, 43) =.41$, $p = .873$]. Four one-way ANOVAs (one for each of the four vowels) with speaker group as a between-subjects factor (MonoM, BiMH,

---

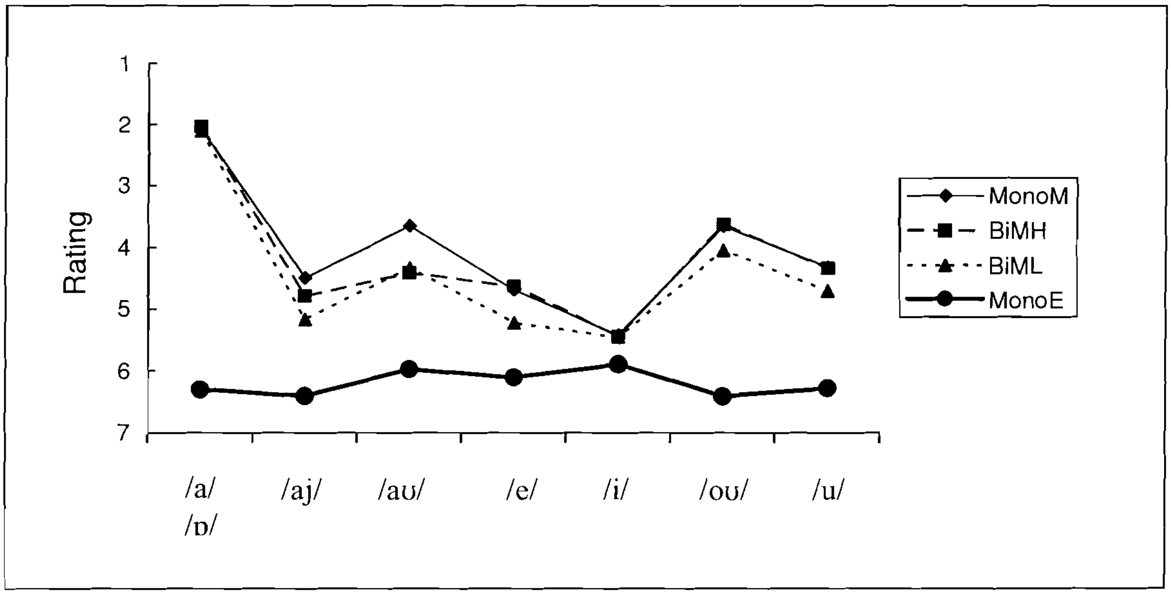
[13] Deviations from normality were also found in the rating scores of BiMH and BiML in vowel /a/. But they were not reported here, because the rating scores of /a/ were excluded from further analysis due to a lower than acceptable inter-rater reliability.

[14] The rating scores of /a/ and /i/ were not included in the analysis.

BiML) revealed a significant main effect of speaker group only in the rating scores of /e/, [$F(2, 43) = 4.61$, $p < .05$]. A Post Hoc Tukey HSD analysis showed a significant difference between the rating scores of BiML and those of BiMH [$p < .05$]. However, no significant difference was found between MonoM and BiMH. A marginal difference [$p = .057$] was found between MonoM and BiML.

To sum up, from the perspective of the English listeners, the Mandarin-English bilinguals did not differ significantly in ratings from the Mandarin monolinguals in the target vowels except /aʊ/, in which BiMH differed significantly from MonoM.

Figure 3-2: Ratings assigned by English listeners (n=16)



Note. The vertical axis has been reversed to permit easy comparison with Figure 3-1 (Munro et al., 1999).

Correlation

If Mandarin-English bilinguals' productions of Mandarin vowels carried some characteristics of English vowels and were therefore judged by Mandarin listeners as bad instances of Mandarin vowels, English listeners might be able to detect the English characteristics in the Mandarin-English bilinguals' productions of Mandarin vowels and judge them as good instances of English vowels. If this is the case, the ratings assigned to the Mandarin-English bilinguals by Mandarin listeners should be negatively correlated with the ratings assigned to the same group of speakers by English listeners. To test this hypothesis, correlations were computed between the ratings assigned to the Mandarin-English bilinguals by Mandarin listeners and those assigned by English listeners to the vowels /aj/, /aʊ/, /e/, /oʊ/, and /u/. The ratings assigned to Mandarin /y/ were not included in the correlation study due to the non-existence of its counterpart in English. The ratings assigned to /a/ and /i/ were also excluded due to the unacceptable inter-rater reliability for English listeners.

An examination of the ratings assigned to the above five vowels by Mandarin listeners and English listeners revealed significant departures from normality in the ratings assigned to /aʊ/ by English listeners (Shapiro-Wilks = .900, $p < .01$), and in the ratings assigned to /e/ (Shapiro-Wilks = .922, $p < .05$) and /u/ (Shapiro-Wilks = .904, $p < .01$) by Mandarin listeners. Accordingly, Spearman rank correlation, which is a nonparametric correlation test, was computed on the ratings of /aʊ/, /e/, and /u/. A Pearson correlation was computed on the ratings of /aj/ and /oʊ/.

A significant negative correlation was found between the ratings assigned to /aʊ/ by Mandarin listeners and the ratings assigned to the same vowel by

English listeners (r_s = -.378, $p < .05$). No significant correlation was found in the ratings assigned to the remaining vowels.

Summary

When compared with MonoM, both BiMH and BiML obtained significantly lower rating scores for Mandarin /y/. With regard to the rating scores assigned by English listeners to Mandarin /aʊ/, BiMH was significantly lower than MonoM. BiML was marginally lower than MonoM. In addition, a significant negative correlation was revealed between the ratings assigned to /aʊ/ by Mandarin listeners and the ratings assigned to the same vowel by English listeners.

Individual differences

Ratings assigned by Mandarin listeners

The analysis in this section aims to examine the individual Mandarin-English bilinguals whose vowel production was judged as accented by Mandarin listeners. In Flege, Munro & MacKay's study (1995) of Italian learners of English, those who obtained a mean rating falling two standard deviations below the mean rating assigned to native English speakers were considered to have accented English pronunciation. The same accentedness criterion was adopted in the present analysis.

Ratings obtained for the Mandarin monolinguals are presented in Appendix 13. Ratings assigned to all Mandarin monolinguals except two (speaker 105 in /a/, /aj/, /aʊ/ and /e/; speaker 107 in /i/, and /y/) are within two standard deviations of the Mandarin monolinguals' group mean. It is not uncommon that a small number of native speakers may be judged as having foreign accents in studies of accentedness (e.g., Flege, Munro & MacKay, 1995; Bongaerts, Mennen, & van der Slik, 2000). This may be due to human error or to other unknown reasons (Murray Munro, personal communication, 2007). A further look at speaker 105 and 107's answers to the questionnaire did not show anything unusual. Since it is a normal phenomenon to have a few speakers being judged as accented, it was decided not to remove the two speakers' data in the present analysis. In the following analysis of the individual Mandarin-English bilinguals' rating scores, the mean ratings and the standard deviations obtained for all 13 Mandarin monolinguals for each of the target vowels were used to calculate if a Mandarin bilingual's rating score was two standard deviations below the Mandarin monolinguals' mean.

Ratings obtained for the Mandarin-English bilinguals are listed in Appendix 14. As many as 13 Mandarin-English bilinguals (6 BiMH, 7 BiML) had rating scores falling two standard deviations below the Mandarin monolinguals' mean in /y/. This is consistent with the finding in group differences that both BiMH and BiML had significantly lower ratings than MonoM.

Although no significant group differences in ratings were found between the Mandarin-English bilinguals and the Mandarin monolinguals in /a/, /aʊ/, /aj/, /e/, /i/, /oʊ/ and /u/, some individual Mandarin-English bilinguals did have ratings falling two standard deviations below the Mandarin monolinguals' mean. For /oʊ/, 5 Mandarin-English bilinguals (2 BiMH, 3 BiML) were judged accented; for each of the vowels /i/ and /a/, 4 Mandarin-English bilinguals (2 BiMH, 2 BiML each) were judged accented; for /e/, 2 Mandarin-English bilinguals (1 BiMH, 1 BiML) were judged accented; for each of the vowels /aj/, /aʊ/ and /u/, 1 Mandarin-English bilingual was judged accented. As the numbers in the brackets indicate, BiML did not outnumber BiMH in being judged as accented by Mandarin listeners.

Ratings assigned by English listeners

For the analysis of individual differences in ratings assigned by English listeners, Flege, Munro & MacKay's (1995) nativeness criterion was adopted. Specifically, Mandarin speakers who received a mean rating for a target Mandarin vowel (e.g., /aj/) that fell within two standard deviations of the English monolinguals' mean rating for a corresponding English vowel (e.g., /aj/) were considered to have produced the Mandarin vowel like an authentic English vowel.

The ratings assigned by English listeners to Mandarin monolinguals and to Mandarin-English bilinguals are listed in Appendix 15 and 16, respectively. An examination of the ratings assigned to the 12 English monolinguals revealed that all speakers had ratings that fell two standard deviations within the English monolinguals' mean except for one speaker who received a rating of 6.06 out of 7 in /oʊ/. This was probably due to a small standard deviation (SD=.15) causing this speaker to fall two standard deviations below the English monolinguals' mean (mean=6.41).

One Mandarin monolingual and 4 Mandarin-English bilinguals (2 BiMH, 2 BiML) obtained a rating for Mandarin /aj/ that fell within two standard deviations of the English monolinguals' mean rating for English /aj/. For Mandarin /aʊ/, 1 Mandarin monolingual and 3 Mandarin-English bilinguals (1 BiMH, 2 BiML) met the nativeness criterion as discussed above. For Mandarin /oʊ/ and /u/, no Mandarin speakers met the nativeness criterion.

For Mandarin /e/, however, as many as 3 Mandarin monolinguals and 10 Mandarin-English bilinguals (2 BiMH, 8 BiML) received a rating that fell within two standard deviations of the English monolinguals' mean rating for English /e/. Since Mandarin monolinguals (23%) and Mandarin-English bilinguals (30%) were comparable in the percentage of speakers whose production of Mandarin /e/ was perceived as authentic English /e/, the evidence is not adequate to conclude that more Mandarin-English bilinguals than Mandarin monolinguals produced Mandarin /e/ as authentic English /e/. In general, one group of speakers (e.g., BiML) did not outnumber the other group of speakers (e.g., BiMH) in being judged as native (e.g., /aj/ and /aʊ/).

Correlation

The data for the Mandarin-English bilinguals who were judged as accented by Mandarin listeners and those who were judged as native by English listeners are summarized in Table 3-4.

Table 3-4: # of Mandarin-English bilinguals being judged as accented by Mandarin listeners (n=16) and those being judged as native by English listeners (n=16)

Mandarin Vowel	Number of speakers judged as accented by Mandarin Listeners	Number of speakers judged as native by English Listeners	Correlation
/a/	4 (2 H, 2 L)	----	no
/aj/	1	4 (2 H, 2 L)	yes
/aʊ/	1	3 (1 H, 2 L)	yes
/e/	2 (2 H)	10 (2 H, 8 L)	no
/i/	4 (2 H, 2 L)	----	no
/oʊ/	5 (2 H, 3 L)	none	no
/u/	1	none	no
/y/	13 (6 H, 7 L)	----	n/a

Note. A dash indicates the value was not available. H and L stand for "speakers of high L1 use" and "speakers of low L1 use", respectively.

An examination of the individual differences provided evidence that the ratings assigned to /aj/ and /aʊ/ by Mandarin listeners and those assigned to the same vowels by English listeners were negatively correlated. For example, the rating assigned to the speaker who was judged as accented by Mandarin listeners was 2.38. English listeners gave the same speaker a rating as high as 5.38. Similarly, the ratings assigned to the 3 speakers who were judged as native by English listeners were 6.06, 6.00, 5.94, and 6.31, respectively. Mandarin listeners gave them a rating of 4.13, 4.38, 4.88 and 4.50, respectively. The same pattern emerges in the rating scores assigned to Mandarin-English bilinguals in the production of /aʊ/. For example, while Mandarin listeners judged a Mandarin-English bilingual as accented (rating=2.13) in the production of /aʊ/, English listeners gave the same speaker a rating as high as 5.06. Similarly, while English listeners gave 3 speakers (all of whom were judged having native-like production of /aʊ/) a rating of 5.38, 5.63, and 5.50, respectively, Mandarin listeners assigned the same speakers a low rating of 4.31, 3.33 and 3.81.

With regard to the rating scores assigned to Mandarin-English bilinguals in the production of the remaining vowels (/e/, /oʊ/, /u/), no negative correlation was found. When speakers were judged as accented and were assigned low scores by Mandarin listeners, they were assigned low scores by English listeners as well. Conversely, when speakers were judged as native and were assigned high scores by English listeners, they were also assigned high scores by Mandarin listeners.

Summary

Consistent with the group differences, as many as 13 Mandarin-English bilinguals were judged as accented by Mandarin listeners in the production of Mandarin /y/. Despite the non-significant group difference in rating scores between Mandarin-English bilinguals and MonoM in /a/, /aʊ/, /aj/, /e/, /i/, /oʊ/ and /u/, some individual Mandarin-English bilinguals were judged as accented by Mandarin listeners. Similarly, English listeners judged some individual Mandarin-English bilinguals as native, even though they were actually judging Mandarin vowels. In general, there was no evidence indicating that BiML outnumbered BiMH in being judged as accented by Mandarin listeners and as native by English listeners. In the analysis of individual differences in the ratings of /aj/ and /aʊ/, a negative correlation was found between the ratings assigned by Mandarin listeners and the ratings assigned by English listeners. No negative correlation was observed between the ratings assigned by Mandarin listeners and those assigned by English listeners in /e/, /oʊ/ and /u/.

Individual differences and acoustic data

As can be seen above, some Mandarin-English bilinguals were judged either as accented by Mandarin listeners or as native by English listeners, or both. An obvious question to ask is what acoustic properties are inherent in these speakers' vowel productions that led to them being judged as accented by Mandarin listeners and as native by English listeners. The following section attempts to explore this issue by relating the individual differences in ratings to acoustic data such as duration, formant frequencies, formant movement and fundamental frequencies. The possible acoustic properties contributing to individual Mandarin-English bilinguals' accent are summarized in Table 3-5.

Table 3-5: Possible acoustic properties attributing to Mandarin-English bilinguals' accentedness

Mandarin Vowel	Number of speakers judged as accented by Mandarin Listeners	Possible acoustic properties	Number of speakers judged as native by English Listeners	Possible acoustic properties
/y/	13 (6 H, 7 L)	Lower F1	----	----
/oʊ/	5 (2 H, 3 L)	tone deviation, exaggerated duration	none	----
/a/	4 (2 H, 2 L)	tone deviation	----	----
/i/	4 (2 H, 2 L)	tone deviation	----	----
/e/	2 (2, H)	tone deviation, short duration	10 (2 H, 8 L)	----
/aj/	1	Larger upward ΔF2	4 (2 H, 2 L)	larger upward ΔF2
/aʊ/	1	tone deviation	3 (1 H, 2 L)	larger downward ΔF2
/u/	1	Larger downward ΔF1	none	----

Note. Dash indicates the value was not available. H and L stand for "speakers of high L1 use" and "speakers of low L1 use", respectively.

Ratings assigned by Mandarin listeners and acoustic data

/y/

This is the only Mandarin vowel in which a significant group difference in rating scores was observed. A comparison was made between the acoustic data of the 13 Mandarin-English bilinguals who were judged as accented (poor speakers) and those of the 13 Mandarin-English bilinguals who were judged as native (good speakers). The 13 good speakers were those who received good ratings in a descending order among the 33 Mandarin-English bilinguals. The 13 MonoM were also included in the statistical analysis. Nine one-way ANOVAs (one for each acoustic dimension: duration, F1a, F2a, F3a, ΔF1, ΔF2, F0 peak, F0 valley, ΔF0[15]), with speaker group as a between-subjects factor (poor speakers, good speakers, MonoM) as a between-subjects factor, were conducted. The analysis revealed a significant group difference in F1a only, [$F(2, 36) = 3.895$, $p < .05$]. A Post Hoc Tukey HSD analysis revealed that the poor speakers, but not the good speakers, had significantly lower F1a than MonoM ($M = 3.87$ Bark), $p < .05$, indicating that, compared with MonoM, the poor speakers produced a significantly higher Mandarin /y/.

/oʊ/

A comparison similar to the analysis of /y/ was made between the acoustic data of the 5 Mandarin-English bilinguals who were judged as accented (poor

---

15 F0 peak was the highest F0 value on the F0 contour, and F0 valley was the lowest F0 value. ΔF0 (F0 range) was the difference between F0 peak and F0 valley.

speakers) and those of the 5 Mandarin-English bilinguals who were judged as native (good speakers). Nine separate one-way ANOVAs (one for each of the nine acoustic dimensions as mentioned in /y/) were conducted. No significant group difference was revealed in any of the nine acoustic dimensions. A closer look at each speaker's acoustic data revealed that the acoustic cues that might have led to these speakers being judged as accented varied from speaker to speaker. For example, speaker 213 produced Tone 4 instead of Tone 1. The difference between her F0 peak (228 Hz) and F0 valley (213 Hz) was as small as 9 Hz. Speaker 226 was had an exaggerated duration of 457 ms, which was much longer than both Mandarin monolinguals' mean (mean=256, *S.D*=58) and Mandarin-English bilinguals' mean (mean=220, *S.D*=60). An examination of speaker 216, 224 and 229's acoustic data did not show anything unusual. The acoustic parameters contributing to these speakers' accent were not identified.

/a/

For the vowel /a/, there were 4 Mandarin-English bilinguals who were judged as accented. Since the number of speakers was too small, no attempt was made to do a statistical analysis. Instead, each speaker's production of /a/ was examined separately.

Speakers 232 and 218 produced Tone 4 /pa/ like Tone 1 /pa/. An examination of the pitch track in *Praat* showed that their pitch contours were flat. For speaker 232, the difference between the F0 peak (207 Hz) and the F0 valley (199 Hz) was only 8 Hz. For speaker 218, the difference between the F0 peak

(116 Hz) and the F0 valley (110 Hz) was as small as 4Hz. An examination of speaker 227 and 233's acoustic data did not show a significant difference between these 2 speakers and the remaining 31 speakers. It is not known why they were assigned such a low rating score.

/i/

Like in /pa/, Speaker 232's production of Tone 4 /pi/ sounded like Tone 1 /pi/. The difference between her F0 peak (220 Hz) and F0 valley (215 Hz) was as small as 5 Hz. An examination of speaker 224, 227 and 229's acoustic data did not identify the acoustic properties that caused them to be perceived as accented.

/e/

Speaker 231 produced Tone 4 /pe/ as Tone 1 /pe/. There was a difference of 22 Hz between her F0 peak (222 Hz) and her F0 valley (201 Hz). As well, the duration of her /e/ was as short as 158 ms, the second shortest among the 33 Mandarin-English bilinguals. Speaker 227 was found to have produced a /pe/ as short as 162 ms, which was the third shortest among the Mandarin-English bilinguals.

/aj/, /aʊ/ and /u/

/aj/, /aʊ/ and /u/ each had 1 speaker being judged as accented. Speaker 229's accentedness was probably due to his large upward F2 movement in the production of /aj/. His F2 movement in /aj/ was 3.2 bark, the second largest movement among all 33 Mandarin-English bilinguals. Like her production of /ej/, speaker 231's production of Tone 4 /paʊ/ was actually Tone 1 /paʊ/. The difference between her F0 peak (196 Hz) and F0 valley (188 Hz) was 8 Hz only. An examination of speaker 210's acoustic data in the production of /u/ showed that his F2 movement of /u/ was -1.21 bark, the second largest among all Mandarin-English bilinguals.

Ratings assigned by English listeners and acoustic data

/aj/

A closer look at the acoustic data of the 4 Mandarin-English bilinguals whose productions of Mandarin /aj/ were judged as native productions of English /aj/ showed that the 4 speakers had substantially larger upward F2 movement (4.95 bark, 3.64 bark, 3.61 bark, and 2.74 bark, respectively) than the Mandarin-English bilinguals (mean=1.98, *S.D*=.86) and the Mandarin monolinguals (mean =1.51, *S.D*=.46).

/aʊ/

An examination of the acoustic data of the 3 Mandarin-English bilinguals whose productions of Mandarin /aʊ/ were heard as native productions of English /aʊ/ revealed that speakers 229, 206 and 210 had F2 movement of -5.79 bark, -2.84 bark and -1.55 bark, respectively. Speakers 229 and 210 ranked the first and second in F2 movement among all 33 Mandarin-English bilinguals. Speaker 210 ranked seventh. In general, these 3 speakers' F2 movement was much larger than the mean F2 movement of all Mandarin-English bilinguals (mean=-1.19, *S.D*=1.05) and of Mandarin monolinguals (mean=-.80, *S.D*=.40).

/e/

Since there was no evidence in the data that more Mandarin-English bilinguals than Mandarin monolinguals produced Mandarin /e/ like an authentic English /e/, no attempt was made to relate the individual Mandarin-English bilinguals' ratings to their acoustic data.

Summary

Compared with MonoM, the Mandarin-English bilinguals who were judged as accented by Mandarin listeners produced a significantly higher /y/. For vowels /a/, /aj/, /aʊ/, /e/, /i/, /oʊ/ and /u/, the acoustic dimensions that might have contributed to some speakers being judged as accented by Mandarin listeners varied from speaker to speaker. These acoustic dimensions include tone, duration, and F2 movement. However, the acoustic dimensions that were associated with the other Mandarin-English bilinguals being judged as accented by Mandarin listeners were not identifiable. For /aj/ and /aʊ/, the Mandarin-English bilinguals who were judged as native by English listeners had a much larger F2 movement than those who were judged as accented.

3.1.3. Discussion

This section responds to the issues raised in research question 2, which was: Do the Mandarin-English bilinguals have an accent in their L1 vowel production? If so, what acoustic properties are associated with this accent? Are the Mandarin-English bilinguals of high L1 use and those of low L1 use equally judged as accented?

Discussion, along with responses to question 2, is provided below.

3.1.3.1. Mandarin-English bilinguals' accent and L1 vowel modification

In this study, many Mandarin-English bilinguals have an accent in their L1 vowel production (e.g., /y/). However, not all cases of accent in all vowels indicate L1 vowel modification. L1 vowel modification is revealed in the Mandarin-English bilinguals' accent in /y/, /aʊ/ and /aj/, but not in /a/, /e/, /i/, /oʊ/ and /u/.

A clear pattern of accentedness was observed for many Mandarin-English bilinguals in the production of /y/. This result is surprising, given that Mandarin /y/ does not have an English counterpart. According to Flege's principle of *equivalence classification* (1987, 1991, 1992, 1995), only an L1 sound that is similar to its L2 counterpart is predicted to undergo reorganization. A possible interpretation is that the crowded vowel space in the vicinity of /y/ of the bilingual speakers triggered the raising of /y/ to allow for sufficient contrast. As can be recalled in 2.2.2, the Mandarin-English bilinguals did not differ from English monolinguals either in F1, F2 or duration. They differed from English monolinguals only in the formant movements of English /ɪ/. Although their production of English /ɪ/ was not as native as English monolinguals, it might be safe to say that they probably have established a category for English /ɪ/. The addition of this new vowel category makes the high, front portion of the vowel space more crowded. To maintain the contrast between L1 and L2 phonetic categories existing in a common phonological space, bilinguals have to adjust their L1 and L2 phonetic categories (Flege, 1995, 2003). It is probably due to the crowded nature of the vowel space that the Mandarin-English bilinguals raised their /y/ to keep it perceptually distinct from its surrounding vowels[16]. The Mandarin-English bilinguals' raising of L1 /y/ suggests that an L1 sound that does not have an L2 counterpart and is therefore not "similar" to an L2 sound may also be adjusted to maintain perceptual contrast in the shared L1 and L2 vowel space.

The second Mandarin vowel that showed L1 vowel modification was /aʊ/. While tone deviation[17] possibly contributed to one Mandarin-English bilingual's accentedness as judged by Mandarin listeners, a larger downward F2 movement was probably responsible for 3 Mandarin-English bilinguals' nativeness as judged by English listeners. The Mandarin-English bilinguals' larger downward F2 movement in Mandarin vowel production suggests the carry-over effect of English /aʊ/ due to the effect of Flege's *equivalence classification.* In other words, the Mandarin-English bilinguals' L1 /aʊ/ might have been "coloured" by English /aʊ/ and is therefore considered native-like.

The third Mandarin vowel in which L1 vowel modification occurs is /aj/. Some accented Mandarin-English bilinguals' larger upward F2 movement in Mandarin /aj/ suggests the influence of English /aj/. Assuming that Flege's

---

[16] According to Stevens (1999), although lip rounding lowers F2 and F3 of high vowels (this effect was not observed within the data of the present study), it does not lower F1. Therefore, lip-rounding is probably not a factor contributing to the raised /y/.

[17] It is not known if tone deviation affects vowel quality.

principle of *equivalence classification* was at work, these Mandarin-English bilinguals assimilated L2 English /aj/ with their L1 Mandarin /aj/, a result of which was that some characteristics of L2 /aj/, such as larger upward F2 movement, was carried over to the corresponding L1 vowel that, consequently, was modified.

With regard to /a/, no indication of L1 vowel modification is available due to the tone deviation observed in two of the four accented Mandarin-English bilinguals[18]. It is not known if tone deviation affects vowel quality.

As many as 10 Mandarin-English bilinguals were judged native by English listeners in the production of Mandarin /e/. At the same time, 2 Mandarin-English bilinguals were judged accented by Mandarin listeners. It is tempting to conclude that these Mandarin-English bilinguals may have modified Mandarin /e/. However, there is a reason to doubt it. The Mandarin monolinguals and the Mandarin-English bilinguals are comparable in the percentage being judged as native by English listeners (23% vs. 30%). It is possible that English listeners might have attended to some common acoustic feature in the production of Mandarin /e/ by the Mandarin monolinguals and the Mandarin-English bilinguals. Further evidence to support this account is the lack of a negative correlation between the ratings assigned to /e/ by Mandarin listeners and those assigned to the same vowel by English listeners. In many cases, the Mandarin-English bilinguals whom Mandarin listeners judged as poor speakers were also judged as poor speakers by English listeners. Similarly, Mandarin-English bilinguals whom Mandarin

[18] Acoustic properties contributing to two other speakers' accentedness were not identified.

listeners judged as good speakers of Mandarin /e/ were judged as good speakers of English /e/ by English listeners as well. The identical acoustic properties of Mandarin and English /e/ probably contributed to many Mandarin speakers being judged native by English listeners. The 2 Mandarin-English bilinguals who were judged accented by Mandarin listeners had tone deviation and short duration, respectively, which, as discussed earlier, may not necessarily affect vowel quality. In sum, there is no obvious indication that the Mandarin-English bilinguals had reorganized Mandarin /e/ in the shared L1 and L2 phonetic space.

There is also no evidence indicating a reorganization of Mandarin /i/, /oʊ/ and /u/. As discussed above, tone and duration deviation observed in some Mandarin-English bilinguals may not be related to vowel modification. The fact that no Mandarin-English bilinguals were judged as native by English listeners and that no negative correlation existed between the ratings assigned to /oʊ/ by Mandarin listeners and those assigned to the same vowel by English listeners further supported the idea that the Mandarin-English bilinguals probably have not modified their Mandarin /oʊ/ category. For Mandarin /u/, a conclusion as to whether it has been modified is not possible due to the small number of accented Mandarin-English speakers (1 accented judged by Mandarin listeners and 0 native judged by English listeners).

3.1.3.2. Acoustic properties contributing to Mandarin-English bilinguals' accent

Lower F2 probably contributed to some Mandarin speakers' accented production of /y/ as judged by Mandarin listeners. As discussed in 3.1.3.1, the raising of /y/ may be due to the enhancement of perceptual contrast in the crowed vowel space. The acoustic properties that might have been responsible for some Mandarin-English bilinguals being judged as accented by Mandarin listeners and as native by English listeners in the production of the remaining vowels varied from speaker to speaker and from vowel to vowel.

On several occasions, Tone 4 target Mandarin words were produced as Tone 1[19]. A possible interpretation of this tone deviation is that some Mandarin-English bilinguals' categories for Tone 1 and Tone 4 may have been modified due to English acquisition, a result of which is tone error in production. It has been reported that English listeners do not attend to pitch direction as much as they do to pitch height (Gandour, 1983, 1984) and they perceive high tones (Tone 1 and Tone 4) as stressed (White, 1981). As a result, they may have one category for both Tone 1 and Tone 4, and confuse Tone 1 and Tone 4 (Wang, Spence, Jongman & Sereno, 1999; Wang, Jongman & Sereno, 2003; Miracle,1989). The Mandarin-English bilinguals who mispronounced Tone 4 as Tone 1 probably behaved like the native English speakers who confused Tone 1 and Tone 4 in perception and made errors in production. Of course, the hypothesis that the Mandarin-English bilinguals who mispronounced Tone 4 as Tone 1 had modified their tone categories could not be verified unless a separate perception test examining Tone 1 and Tone 4 were administered. Future research should take this into consideration.

[19] This is a flaw of the present study. An identification test should have been designed to exclude such tokens from the rating experiments.

Larger downward or upward F2 movement is another acoustic feature that contributed to some Mandarin-English bilinguals being judged as accented by Mandarin listeners and as native by English listeners (e.g., Mandarin /aʊ/ and /aj/). As can be seen in 2.2.1, English /aʊ/ and /aj/ had a larger downward and upward F2 movement, respectively, than their Mandarin counterparts. For these Mandarin-English bilinguals, the characteristics of English /aʊ/ and /aj/ were probably transferred to their production of Mandarin as a result of learning English.

It must be noted, though, that the acoustic dimensions leading to a few Mandarin-English bilinguals being judged as accented are not identifiable. It is possible that other acoustic dimensions that were not measured in the present study (e.g., voice quality) may have contributed to their accentedness.

3.1.3.3. Mandarin-English bilinguals' accent and amount of L1 use

Neither the group differences nor the individual differences showed an effect of the amount of L1 use on the Mandarin-English bilinguals' accent. For example, both BiMH and BiML were significantly lower than MonoM in the rating of Mandarin /y/. A similar scenario was revealed in the analysis of individual differences, where, in most cases, there was a balanced number of BiMH and BiML whose ratings fell two standard deviations below the Mandarin monolinguals' mean and two standard deviations within the English monolinguals' mean.

This finding is inconsistent with McRobbie (2003) who found an effect of the amount of L1 use. Hungarian immigrants who used less L1 diverged the least from the monolingual Hungarian speakers in the production of Hungarian [ɒ]. However, all of her speakers except two were adults when they immigrated to Canada. The speakers in the present study were all children aged 9-13 when they moved to Canada. The inconsistency between the finding of the present study and that of her study is probably due to the different AOA of the two groups of speakers.

The finding in the present study is consistent with Guion et al.'s (2000) study of Quichua-Spanish bilinguals and Yeni-Komshian et al.'s (2000) study of Korean-English bilinguals in that effect of L1 use was not observed. However, there is a substantial difference between the findings of the present study and those of Guion et al. (2000) and Yeni-Komshian et al.'s (2000) study. While neither the bilinguals of high L1 use nor those of low L1 use in Guion et al. (2000) and Yeni-Komshian et al.'s (2000) study were judged accented by native speakers, some Mandarin-English bilinguals in the present study, be it high L1 users or low L1 users, were judged as accented by native Mandarin listeners (e.g., in the production of Mandarin /y/). The inconsistency between the finding of the present study and those of the previous studies (Guion et al., 2000; Yeni-Komshian et al., 2000) is probably due to the different tasks being carried out by listeners. In Guion et al. (2000) and Yeni-Komshian et al.'s (2000) study, the listeners' task was to rate sentences for degree of foreign accent. In the present study, listeners were asked to rate the goodness of vowels. In rating vowels, listeners' attention focuses on vowels only. If there is some deviation from the norm in the production of a vowel, native listeners may easily pick it up. However, when rating sentences, listeners have far more information to attend to. With regard to phonetic information, they have to attend to both segmental and suprasegmental information. As well, they may also need to attend to syntactic and semantic information. It is possible that a sentence may still be rated as a good sentence of a native language even if there is some deviation in a certain segment (e.g., vowel). The deviation of a segment may be ignored or not have priority in determining the overall pronunciation of a sentence. Nevertheless, this interpretation cannot be confirmed unless two separate perception tests, one on vowels and the other on sentences containing the same vowels, are administered. Future research should take care of this.

In sum, some Mandarin-English bilinguals did sound accented to Mandarin listeners and sound native to English listeners, whether their amount of L1 use was high or low. A possible interpretation is that an L2 will exert influence on an L1 if a bilingual regularly uses and is exposed to L2. As can be recalled in 2.1.1, the BiMH's amount of L1 use ranged from 60% to 80% and the BiML's amount of L1 use ranged from 20% to 40%. None of these bilinguals used Mandarin exclusively in all situations. Even the speakers who had the highest amount of L1 use still spoke English 20% of time. Besides, living in an English speaking country, they may have had more exposure to English than Mandarin. As can be seen in 2.1.1, when compared with Mandarin input, both BiMH and BiML watched more English TV programs, saw more English movies, and listened to more English radio programs. The two groups did not differ significantly from each other in the amount of English input received from TV, movies, and radio. This could be expected in an English speaking country like Canada where English media is more predominate than Mandarin media.

3.1.3.4. Summary

The analysis of group differences and individual differences in Experiment 1 indicates that some Mandarin-English bilinguals have an accent in their production of Mandarin vowels. Despite this accent, only some L1 vowels have been modified as a result of L2 learning. For example, while there is obvious evidence of L1 vowel modification in /y/, /aʊ/ and /aj/, there is no indication that the remaining L1 vowels (/a/, /e/, /i/, /oʊ/ and /u/) have been modified. In addition, the acoustic properties possibly contributing to some Mandarin-English bilinguals' accent include lower F1 (/y/), larger upward (/aj/) or downward F2 (/aʊ/) movement, tone deviation (/a/, /aʊ/, /e/, /i/, /oʊ/) and extremely short (/e/) or long duration (/oʊ/). Furthermore, the effect of the amount of L1 use is not revealed. The Mandarin-English bilinguals of high L1 use and those of low L1 use are equally judged as accented.

It must be noted that the goodness rating of vowels in this experiment may not have a cause-effect relation due to confounding factors such as initial consonants, tone, etc.

3.2. Experiment 2

This experiment aims to answer research question 3, "Is there an inverse correlation between the Mandarin-English bilinguals' English vowel pronunciation proficiency and their Mandarin vowel pronunciation proficiency?" If the L1 and L2 phonetic systems coexist in a common phonological space in a bilingual's mind, a stronger L2 system should correlate with a weaker L1 system, and vice versa. It has been reported that there is an inverse relationship between a bilingual's L1 overall pronunciation proficiency and his/her L2 overall pronunciation proficiency (Yeni-Komshian et al., 2000). In the same fashion, the higher English vowel rating scores obtained by Mandarin-English bilinguals should correlate with their lower Mandarin vowel rating scores, and vice versa. To test this hypothesis, the Mandarin-English bilinguals' English vowel productions were presented to phonetically trained native English listeners for goodness rating. Following this, a correlation study was conducted to examine the relationship between the Mandarin-English bilinguals' English vowel proficiency and their Mandarin vowel proficiency.

3.2.1. Methodology

3.2.1.1. Listeners

Since the goodness rating test required listeners to focus their attention on the vowel in each target word, 2 phonetically trained monolingual English listeners in Linguistics Department of Simon Fraser University were recruited to do the goodness rating test. Both reported normal hearing. The 2 listeners were instructed to rate the goodness of the vowel in each CV word on a 5-point scale, with “1” indicating a bad and “5” a good exemplar of the target vowel. A 5-point scale was deemed appropriate for evaluating vowel production in this experiment. English monolinguals were on the “good” end and Mandarin-English bilinguals would approximate or deviate from the “good” end.[20]

3.2.1.2. Stimuli

The second repetitions of the 12 English CV words (see 2.1.2.1) produced by 33 Mandarin-English bilinguals and 4 English monolinguals in the production test were selected as stimuli for the goodness rating test. The 4 English speakers were randomly selected from among the 12 English monolinguals. Their inclusion was for reference only. The average RMS intensity (loudness) of these English words were normalized to -16dB in *Sound Forge* (Version 8.0). In total, 396 tokens (12 vowels × 33 Mandarin-English bilingual speakers) were analyzed.

3.2.1.3. Procedure

Stimuli were divided into 12 blocks, 1 for each of the 12 vowels. The stimuli (CV words) were each presented once to listeners for goodness rating via *E-Prime* (Version 2000) on a laptop computer (Compaq Presario V2000). Before the experiment started, the listeners completed a consent form and a questionnaire (Appendix 1). The listeners used headphones (Technics RP-HT400) in the experiment. The experiment began with a practice session to familiarize the listeners with the test procedure. The inter-trial interval was set to

[20] A 7-point scale was used in Experiment 1, in which there were four groups of speakers, namely, Mandarin monolinguals, English monolinguals, Mandarin-English bilinguals of high L1 use and Mandarin-English bilinguals of low L1 use. To achieve a finer evaluation in Experiment 1, the 7-point scale was deemed appropriate.

be indefinite to create a self-paced task. The listeners were instructed to focus their attention on the goodness of the vowel. In the actual test, the presentation of the 12 blocks was counter-balanced. There was a 5-minute break when the listeners finished the first half of the test (6 blocks). When they finished the test, they were asked to complete a debriefing questionnaire (Appendix 6).

3.2.2. Results

An examination of the rating data in this experiment showed that the majority of the Mandarin-English bilinguals received good rating scores in their productions of English vowels /i, e, æ, u, oʊ, ɒ, ʌ, aj, aʊ/. Although /ʊ/ was previously found to be difficult to acquire for Mandarin speakers (Wang, 1997), an examination of the rating data indicated little variation. A vast majority of the Mandarin-English bilinguals (31 out of 33) received a score ranging from 4 to 5; 2 people were in the score range of 3 to 3.5; no one received a score that was less than 2.5. Due to the lack of variation in the rating scores of /i, e, æ, u, oʊ, ɒ, ʌ, aj, aʊ, ʊ/, the rating scores of these English vowels were not included in this correlation study. An examination of two other English vowels /ɪ/ and /ɛ/, which were previously found difficult for Mandarin speakers to acquire (Wang, 1997), showed a full range of scores ranging from low to high. The rating scores of these two English vowels were selected for this correlation study. It was decided that the Mandarin-English bilinguals' English phonetic proficiency level could be reasonably indexed by the rating scores of the English front vowel /ɛ/ and /ɪ/. If a Mandarin speaker is good at producing these difficult vowels, he/she probably does not have much difficulty producing other English vowels correctly. Mandarin /y/ was selected, because it was the only Mandarin vowel in which a significant group difference was found and a large number of individuals (n=13) were judged as accented by Mandarin listeners.

3.2.2.1. Inter-rater reliability

Inter-rater reliability estimates by vowel /ɛ/ and /ɪ/ were computed for the English evaluators by using Cronbach's α. For /ɛ/, Cronbach's α was .853. For /ɪ/, it was .826. This indicates that the two English evaluators had a high agreement in their rating of these two English vowels. To recall (3.1.2.2), for Mandarin /y/, Mandarin listeners also reached an acceptable agreement (Cronbach's α=.76).

3.2.2.2. Correlation coefficients

The ratings assigned by the Mandarin listeners to Mandarin-English bilinguals' production of Mandarin /ly/ and those assigned by the English evaluators to Mandarin-English bilinguals'[21] productions of English /pɛt/ and /pɪt/ are listed in Table 3-6.

An examination of the data showed significant departures from normality in the rating scores of Mandarin /y/ (Shapiro-Wilks = .901, $p < .01$), English /ɛ/ (Shapiro-Wilks = .772, $p < .001$) and English /ɪ/ (Shapiro-Wilks = .746, $p < .001$). Because of this, a nonparametric correlation test, namely, Spearman rank correlation, was conducted. The Spearman correlation analysis revealed no

[21] Mandarin-English bilinguals were not further divided into BiMH and BiML due to the balanced number of "high L1 use" and "low L1 use" speakers being judged as accented in Mandarin /y/.

significant correlation between the rating scores of any of the English vowels and the rating scores of the Mandarin vowel. To be specific, the Mandarin-English bilinguals' rating scores of English /ɛ/ were weakly correlated with the Mandarin-English bilinguals' rating scores of Mandarin /y/ [r_s = .009, p = .480]. Similarly, a weak correlation was revealed between the Mandarin-English bilinguals' rating scores of English /ɪ/ and their rating scores of Mandarin /y/ [r_s = -.045, p = .401].

Summing up, the correlation study did not reveal an inverse correlation between the ratings of English vowel /ɛ/ and /ɪ/ and the ratings of Mandarin /y/. As can be seen in Table 3-6, the Mandarin-English bilinguals who were judged by Mandarin listeners as accented in the production of Mandarin /y/ included both good and bad speakers of English vowel /ɛ/ and /ɪ/.

Table 3-6: Goodness ratings assigned by Mandarin listeners (n=16) to Mandarin-English bilinguals' production of Mandarin /ly/ and ratings assigned by English evaluators (n=2) to Mandarin-English bilinguals' production of English /pɛt/ and /pɪt/

Subject	Rating of Mandarin /ly/	Rating Of English /pɛt/	Rating of English /pɪt/
1	3.31*	3.5	5
2	3.75*	4.5	4.5
3	4.06*	2	4.5
4	4.44*	4	4.5
5	4.73*	5	4.5
6	4.8*	5	4.5
7	4.87*	2.5	2.5
8	4.88*	5	5
9	4.88*	5	3
10	4.88*	5	5
11	4.88*	5	4.5
12	4.88*	2	1
13	4.94*	3.5	3
14	5.06	3	4.5
15	5.06	3.5	5
16	5.25	5	2
17	5.38	5	5
18	5.38	4	5
19	5.38	5	1.5
20	5.44	4.5	4
21	5.5	1	4
22	5.53	4.5	4.5
23	5.56	5	4.5
24	5.56	2.5	4
25	5.56	5	5
26	5.56	5	4.5
27	5.56	5	5
28	5.63	1.5	5
29	5.81	1.5	4.5
30	5.87	5	5
31	5.88	4.5	3.5
32	6	5	5
33	6.06	5	5

Note. Numbers with asterisk indicate ratings that are two standard deviations below the Mandarin monolinguals' group mean. Ratings of /pɛt/ and /pɪt/ were on a five-point scale whereas ratings of /ly/ were on a seven-point scale.

3.2.3. Discussion

The present study provided no evidence of an inverse correlation between the Mandarin-English bilinguals' English vowel pronunciation proficiency and their Mandarin vowel pronunciation proficiency. Good English vowel pronunciation did not predict bad Mandarin vowel pronunciation, and vice versa.

The findings of the present study are not comparable to the findings of Yeni-Komshian's study (2000). In that study, the researchers examined the correlation between L1 pronunciation proficiency and L2 pronunciation proficiency among three groups of Korean-English bilinguals differing in AOA. The focus of the study was differences among groups. The group of speakers with AOA ranging from 10-11 in Yeni-Komshian et al.'s study was comparable to the group of speakers examined in the present study. However, the focus of Yeni-Komshian et al.'s study was to compare the bilinguals' English pronunciation scores with their Korean pronunciation scores to see if one is better than the other. The present study does not aim to examine if the Mandarin-English bilinguals, as a group, have better pronunciation in Mandarin or better pronunciation in English. Instead, it aims to determine whether the Mandarin-English bilinguals who were poor in Mandarin vowel pronunciation would be those who were good in English vowel pronunciation.

The lack of an inverse correlation between the Mandarin-English bilinguals' English vowel pronunciation proficiency and their Mandarin vowel pronunciation proficiency among the Mandarin-English bilinguals in the present study indicates that English vowel pronunciation proficiency may not be a crucial factor contributing to poor Mandarin vowel pronunciation. Factors, such as quantity and quality of L1 and L2 input (Flege, 1987; Major, 1992, 1993), and socio-linguistic factors such as intention to identify with the host culture (Major, 1993) may be more important.

CHAPTER 4: GENERAL DISCUSSION AND CONCLUSION

4.1. General discussion

4.1.1. L1 vowel modification and influence of L2 learning

The present study provided some limited supporting evidence for the claim that L2 learning influences L1 vowel production (e.g., Flege, 1995, 2003). For example, the less downward F1 movement in the production of Mandarin /i/ by BiMH and BiML was probably due to the less downward F1 movement of English /i/ being carried over to the production of Mandarin /i/. Similarly, the more upward F2 movement in the production of Mandarin /aj/ by BiML might be the result of the larger upward F2 movement of English /aj/ being carried over to the production of Mandarin /aj/. A similar carry-over effect was found in the production of Mandarin /aʊ/, in which some Mandarin-English bilinguals had significantly larger downward F2 movement. The influence of L2 learning is further revealed in the Mandarin-English bilinguals' production of Mandarin /y/, which was pushed upward possibly to achieve a perceptual contrast when a new L2 vowel (/ɪ/) was established and the high, front portion of the vowel space got crowded.

Taken together, these findings indicate that some Mandarin-English bilinguals have modified their L1 vowels as a result of L2 learning. Similar findings of L2 influence on L1 are found in Baker & Trofimovich (2005), Guion et al. (2000), Guion (2003), Flege et al. (2003), Harada (2003), Yeni-Komshian et al. (2000), Peng (1993), Major(1992), Flege & Efting (1988) and Flege (1987). The findings in the present study, along with the findings in the previous studies, provide support for the two assumptions made in the SLM. First, the two phonetic systems coexisting in a common phonological space interact in a bidirectional way (Flege, 1995). That is, the L1 phonetic system exerts an influence on the L2 phonetic system, and vice versa. Second, the L1 phonetic categories established in childhood do not remain static over the life span (Flege, 1995); instead, they may undergo modification when similar L1 and L2 sounds interact in the process of L2 learning.

It must be noted that not all the assumptions made in the SLM were supported by the findings of the present study. Flege (2003) proposed the mechanisms of "category assimilation" and "category dissimilation" to account for the interaction of L1 and L2 phonetic categories. The findings in the present study did not show the occurrence of "category dissimilation," even though the category of an English vowel similar to its Mandarin counterpart was established (e.g., /aj/). For example, BiML did not differ from MonoE in the production of English /aj/ in any of the acoustic dimensions measured in the present study. It is probably safe to say they have established a category for English /aj/. According to "category dissimilation", BiML were expected to produce Mandarin /aj/ less similar to English /aj/ to maintain a phonetic contrast. However, BiML was found to have significantly larger upward F2 movement than MonoM in the production of /aj/, indicating their Mandarin /aj/ was similar to English /aj/ in F2 movement. In fact, "category assimilation" occurred.

An interesting finding in the present study is that an obvious pattern of accentedness was observed in the Mandarin-English bilinguals' production of Mandarin /y/, a vowel nonexistent in English. To my knowledge, none of the speech perception and production models, including the NLM, the PAM and the SLM, predict what will happen to an L1 sound that does not have an L2 counterpart when the L1 phonetic system and the L2 phonetic system interact in a bilingual's mind. The NLM and the PAM do not make predictions about the status of L1 sounds when the L1 and the L2 phonetic systems interact. The SLM only makes predictions about the status of an L1 sound that is similar to an L2 sound (e.g., Mandarin /aj/ and English /aj/). The present study shows that some Mandarin-English bilinguals modified their Mandarin /y/, possibly due to the addition of the English vowel /ɪ/ in the common L1 and L2 vowel space. This finding suggests that the study of the influence of the L2 phonetic system on the L1 phonetic system must consider the whole phonetic space (Yue Wang, personal communication, 2005) assuming that L1 and L2 share a common phonetic space (Flege, 1995, 2003) and neighboring sounds influence each other. In other words, a study of this nature must take into consideration all the phonetic segments in the shared L1-L2 phonological space. It may not be sufficient just to look at the interaction between an L1 phonetic segment and its L2 counterpart (e.g., English /i/ versus Mandarin /i/). As shown above, an L1 phonetic segment (e.g., Mandarin /y/) that does not have an English counterpart may undergo reorganization due to a neighboring phonetic segment (e.g., English /ɪ/).

4.1.2. L1 vowel modification and cross language similarity

The present study examined eight Mandarin vowels, seven of which had English counterparts. However, not all seven Mandarin vowels showed the effect of L2 influence. For example, only two vowels, /aj/ and /aʊ/, revealed an obvious pattern of L2 influence on the L1. This may be related to cross language similarity. The examination of cross-linguistic vowel perceptual similarity in the present study showed that Mandarin vowels' similarity to their English counterparts descended in the order of /i/, /e/, /aj/, /u/, /oʊ/, /aʊ/, /a/. Mandarin /aj/ and /aʊ/ were neither the most nor the least similar to their English counterparts. If a Mandarin vowel is identical to its English counterpart (e.g., /i/, /e/), Mandarin-English bilinguals would probably use one category for both vowels and non-native production of the "identical" sound is not expected. This provides support for Williams' claim (1980) that an effect of L2 learning on L1 can not be demonstrated if an L1 phonetic segment and its L2 counterpart are not reliably different in the first place. However, if an L1 phonetic segment and its L2 counterpart are too different, an effect of the L2 learning upon L1 vowel is probably not expected either. In sum, the effect of L2 influence on L1 is mostly likely to be observed in L1 sounds that are similar to, but at the same time reliably different from, their L2 counterparts.

4.1.3. L1 vowel modification and amount of L1 use

Consistent with previous studies (Flege, Frieda & Nozawa, 1997; Flege, Schirru & MacKay, 2003; Piske & MacKay, 1999; Guion et al., 2000; Yeni-Komshian et al., 2000), the present study reveals that the amount of L1 use played a role in the L1 influence on L2. For example, while the Mandarin-English bilinguals of high L1 use differed from MonoE in F2 movement of English /aj/ and F2 of English /u/, those of low L1 use did not.

However, the amount of L1 use, in general, did not play a role in L2 influence on L1. For instance, both Mandarin-English bilinguals of high L1 use and those of low L1 use were found to have modified some Mandarin vowels. Despite the distinction between low and high L1 use, all of the Mandarin-English bilinguals in the present study used and were exposed to English on a regular basis. It seems that a bilingual's L1 is affected in one way or another as long as he or she uses and is exposed to L2 on a regular basis. Previous studies of VOT (Flege and Hillenbrand, 1984; Flege, 1987; Flege & Efting, 1987; Sancier and Fowler, 1997), fricatives (Peng, 1993) and accent (Munro et al., 1999) provide supporting evidence for this claim.

It has to be admitted that determining the amount of L1 use may be problematic. Following the practice of previous studies (Flege, Frieda & Nozawa, 1997; Piske & MaKcay, 1999; Guion et al., 2000; Yeni-Komshian et al., 2000), the present study elicited information about L1 use by asking subjects to report the amount of the current use of their L1. While the amount of L1 use is well indexed by such a report, it fails to provide information regarding the amount of

L1 use in the first few years of L2 exposure. If information about the amount of L1 use in the first few years of L2 acquisition is available, the amount of L1 use may be better indexed. It is true that eliciting such information is difficult because subjects may not be able to remember the amount of L1 use in the long past. However, it is not impossible to manage. Future research should take this issue into consideration.

4.1.4. L1 vowel modification and L2 vowel acquisition

The production test showed that both the Mandarin-English bilinguals of high L1 use and those of low L1 use had acquired English /i/, which was indicated by the finding that the Mandarin-English bilinguals did not differ significantly from the English monolinguals in any of the acoustic dimensions measured in the present study. As compared with the Mandarin monolinguals, these Mandarin-English bilinguals produced Mandarin /i/ with significantly less downward movement, indicating the carry-over effect of the less downward F1 movement of English /i/. Another finding the production test revealed was that the Mandarin-English bilinguals of low L1 use had successfully acquired English /aj/. Their F2 movement in the production of English /aj/ did not differ significantly from the English monolinguals. The carry-over effect of a larger upward F2 movement was found in these speakers' production of Mandarin /aj/. Compared with the Mandarin monolinguals, the Mandarin-English bilinguals of low L1 use produced Mandarin /aj/ with significantly larger upward F2 movement. These findings suggest that successful acquisition of an L2 vowel does indeed influence the production of a corresponding L1 vowel.

It must be noted, however, that the effect of L2 vowel acquisition interacts with cross-language similarity. If an L1 vowel is acoustically identical to its L2 counterpart, no change of vowel will occur. On the other hand, if an L1 vowel is acoustically too dissimilar from its L2 counterpart, no change of vowel will occur either.

4.1.5. Relationship between L1 and L2 pronunciation proficiency

The correlation study did not reveal an inverse correlation between the Mandarin-English bilinguals' English vowel pronunciation proficiency and their Mandarin vowel pronunciation proficiency. This indicates that good English vowel pronunciation does not necessarily index bad Mandarin pronunciation, and vice versa. L2 vowel pronunciation proficiency may not be a crucial factor contributing to poor Mandarin vowel pronunciation. Other factors, such as quantity and quality of L1 and L2 input (Flege, 1987; Major, 1992, 1993), intention to identify with the host culture (Major, 1993), etc., may be more important.

4.1.6. Relating production and perception

The acoustic analysis in 2.2.3 did not show a significant difference between the Mandarin-English bilinguals and the Mandarin monolinguals in duration, formant frequencies, formant movement and tone of Mandarin /a/, /aʊ/, /e/, /oʊ/, /u/ and /y/. A significant group difference in formant movement was revealed in the production of /i/ and /aj/ only. In the production of Mandarin /i/, both BiMH and BiML had significantly less downward F1 movement than MonoM. In the production of /aj/, BiML had significantly more upward F2 movement than MonoM.

The analysis of the rating scores assigned to Mandarin /y/ by Mandarin listeners showed a significant group difference between the Mandarin-English bilinguals and the Mandarin monolinguals. An examination of the rating scores assigned to Mandarin /aʊ/ by English listeners revealed a significant difference between BiMH and MonoM, and a marginal difference between BiML and MonoM. No significant difference between the Mandarin-English bilinguals and the Mandarin monolinguals was revealed in the rating scores of the remaining vowels.

If the results of the production test and those of the perception test are compared, some inconsistencies arise. First, the perception test revealed a significant group difference in the rating scores of /y/ whereas the production test did not show a group difference in any of the acoustic dimensions that were measured for the same vowel. Second, the production test revealed a significant group difference in F1 movement in /i/ and F2 movement in /aj/, even though the perception test did not show a significant group difference in the rating scores of the same vowel.

The first inconsistency is not surprising, because the non-significant difference in the production test did not rule out the possibility that a certain number of individual speakers acoustically deviated from the norm of a native vowel. Findings in the analysis of individual differences provided evidence for this claim. When the 13 poor speakers and the same number of good speakers as judged by Mandarin listeners were compared with MonoM in the acoustic dimensions of /y/, it was the 13 poor speakers, but not the 13 good speakers, who were found to have a significantly higher /y/ than the MonoM. Similarly, though the production test did not reveal a significant difference between the Mandarin-English bilinguals and MonoM in the F2 movement of /aʊ/, the analysis of the individual differences in rating scores showed that the three Mandarin-English bilinguals who were judged as native by English listeners did have a substantially larger downward F2 movement. Therefore, it is important to analyze inter-speaker variability in both perception and production and then relate perception results to production results.

A possible interpretation for the inconsistency in the case of Mandarin /i/ is that the significant difference of F1 movement between the Mandarin-English bilinguals and MonoM as revealed in the production test may not be salient in the perception test. The difference of the F1 movement was probably unnoticeable to the listeners who did the goodness-rating task. After all, Mandarin /i/ and English /i/ were perceptually "identical" despite the acoustic difference in F1 movement. Though some Mandarin-English bilinguals were judged as accented by Mandarin listeners, F1 movement was not found to be an acoustic cue contributing to their accentedness.

The results of the production test and those of the perception test in Mandarin /aj/ are, in fact, not completely inconsistent, because, as can be seen in the analysis of individual differences, some Mandarin-English bilinguals were judged either as accented by Mandarin listeners or as native by English listeners, and these speakers were all found to have substantially larger upward F2 movement. However, it is probably not reasonable to expect that all Mandarin-English bilinguals who had a larger upward F2 movement would be judged as accented by Mandarin listeners and as native by English listeners, because, in addition to F2 movement, many other acoustic cues are available when listeners do a goodness-rating task. The goodness rating may not be based solely on one acoustic cue. In addition, the weighting of the acoustic cues in goodness rating may vary from case to case.

Taken together, the results in the present study show that it is of great importance and necessity to include both production tests and perception tests in a study of this nature.

4.1.7. Other issues

The present study examined the similarities between Mandarin and English vowels in both a production test and a rating test. While the acoustic analysis revealed an acoustic difference between a Mandarin vowel and its English counterpart (e.g., F1 movement of /i/), the perception test showed the vowel was "identical" cross-linguistically. The results in the present study provide further evidence for the claim that it may be difficult to determine cross language similarities solely on the basis of acoustic analysis (Flege, Bohn and Jang, 1997).

The present study included both Mandarin and English monolingual listeners in the perception test. An examination of the correlations between the ratings of Mandarin listeners and those of English listeners revealed that both listener groups could detect the English "colour" in some Mandarin-English bilinguals' productions of some Mandarin vowels (e.g., /aj/ and /aʊ/). The present study provides proof for the effectiveness of including both L1 and L2 listeners in the detection of bilinguals' accentedness in L1 production (Munro et al., 1999).

4.2. Limitations

The present study has several limitations. First, it only examined bilinguals of one language (Taiwanese Mandarin) in one location (Vancouver, Canada). To draw a firm conclusion regarding L2 effects on L1 vowels, bilinguals of another language (e.g., Italian) or bilinguals of the same language (Mandarin) in a different location (e.g., Melbourne, Australia) should be investigated. Second, the stimuli in the perception tests only included the second repetition of the vowel productions. If all three repetitions of a vowel were included, Mandarin-English bilinguals' performance in the three different repetitions could be observed. It is possible that the Mandarin-English bilinguals may show different patterns in different repetitions of a vowel. Third, following the practice of previous studies (Flege, Frieda & Nozawa, 1997; Piske & MacKay, 1999; Guion et al., 2000; Yeni-Komshian et al., 2000), the present study determined the amount of L1 use by subjects' self report of current L1 use. However, information regarding the subjects' amount of L1 use in the first few years of L2 exposure may be more important. Finally, when cross language similarities and differences were examined, the initial consonants of the Mandarin stimuli were Mandarin consonants and those of the English stimuli were English consonants. The initial consonants might have influenced listeners' judgement of cross-language similarities and differences. In addition, tone and other acoustic features that are not necessarily vowel features might have played a role in the judgement of cross-language vowel similarities and differences. These limitations should be taken into consideration in future research.

4.3. Conclusion

The present study confirms previous research that the L1 system established in childhood does not remain static. Instead, it may undergo modification when the L1 phonetic system and the L2 phonetic system interact in a common phonological space. The modification of L1 segments is closely related to cross-language similarity and acquisition of L2 segments. The L1 segments that are similar, but not identical, tend to undergo modification. Successful acquisition of an L2 segment may also trigger the modification of a corresponding L1 segment. However, it must be noted that an L1 segment that does not have an L2 counterpart may also be reorganized as a result of L2 learning. Therefore, phonetic segments (e.g., vowels) must be examined in the whole system in which all L1 segments reside (e.g., all vowels in a vowel space). The amount of L1 use does not seem related to the modification of L1 segments. It seems that a bilingual's L1 will undergo modification as long as he or she uses and is exposed to an L2 on a regular basis. It is also found that L2 pronunciation proficiency and L1 pronunciation proficiency are not inversely correlated.

Whether bilinguals have good or bad L2 pronunciation, it is possible for them to have good L1 pronunciation.

Since cross-language similarity and acquisition of L2 segments are closely related to L1 modification, they deserve to be further explored. Future research should focus on how cross-language similarity and acquisition of L2 segments correlate with L1 modification. Moreover, given the finding that an L1 segment that is neither "identical" nor "similar" to an L2 segment may undergo modification, subsequent research and theories should take into account both "similar" and neighbouring dissimilar sounds. Furthermore, though L2 proficiency is not found to be a factor predicting L1 proficiency, other factors, such as language input and motivation to identify with the host culture, may correlate with L1 proficiency and should be considered in future research.

In examining the effect of L2 phonetic learning on L1 vowels, the present study contributes to the less well-studied field of L2 influence on L1. In particular, it suggests the necessity to include dissimilar L1 segments in speech learning theories.

APPENDICES

Appendix 1: English monolinguals and Mandarin-English bilinguals' background information

1. Name: _______________(last), _____________________(First)
2. Sex: Male Female
3. Age: ___________________
4. Age of arrival in Canada: _____________ (If your first language is English, please skip to Question 6)
5. Age at first exposure to English: _________
6. Birthplace: ________________________________ (city/town, country)
7. Place(s) where you spent most of your childhood?

   _________________________ (city/town, country)
8. Other places you lived more than six months:

   Place 1 __________ Age __________ Duration __________

   Place 2 __________ Age __________ Duration __________

   Place 3 __________ Age __________ Duration __________
9. First language: Mandarin, English
10. Please list language(s) you speak and rate your fluency level in speaking, listening, reading and writing according to the criteria provided.

    1. fluent 2. somewhat fluent 3. not fluent

Language 1 __________ speaking___ listening___ reading___ writing___

Language 2 __________ speaking___ listening___ reading___ writing___

Language 3 __________ speaking___ listening___ reading___ writing___

Language 4 __________ speaking___ listening___ reading___ writing___

6. Which language do you use the most

at home? __________

at school? __________

at work? __________

with friends? __________

7. What is the overall percentage of your use of English and Mandarin (in spoken
form) in your day-to-day affairs? (If your first language is English, please skip to Question 14)

English: ________% **Mandarin**________%

8. What is the overall percentage of English TV programs and Mandarin TV programs you watch? (If you don't watch TV, skip to Question 9)

English programs ________% **Mandarin programs** ________%

9. What is the overall percentage of movies in English and movies in Mandarin you watch? (If you don't watch movies, skip to Question 10)

Movies in English ________% **Movies in Mandarin** ________%

10. What is the overall percentage of English radio programs and Mandarin radio programs you listen to? (If you don't listen to radio, skip to Question 11)

English programs ________% **Mandarin programs** ________%

11. How often do you go back to Taiwan? And how long is each visit?

12. Is your hearing normal? **Yes; No** (circle one)

13. Contact information:

Tel:

E-mail:

16. Would you like to be contacted for a subsequent experiment? **Yes; No** (circle one)

Appendix 2: Mandarin monolinguals' background information

問卷調查

1. 姓名: ____________________ 2. 性別：男，女 （請選擇）

3. 教育程度：__________ 4. 專業:__________________

4. 年齡：____________歲 5. 出生地: ____________縣/市

6. 你的童年是在何地度過的：____________縣/市

7. 生活超過六個月的地方（請列出具體地點, 年齡, 時間）：

地點 1：__________縣/市 ； 年齡：_______ ； 時間: ________月
地點 2：__________縣/市 ； 年齡：_______ ； 時間: ________月
地點 3：__________縣/市 ； 年齡：_______ ； 時間: ________月

8. 你何時來加拿大？_________年________月

9. 你的母語：國語，台灣話，客家話 （請選擇）

10. 你講台灣話，客家話或其它方言嗎？熟練程度如何？請用以下標准判斷：1=流利，2=一般流利，3=不流利；並將相應的數字填在空格上。

台灣話：__________ 聽力_____ 口語_____ 閱讀_____ 寫作_____
客家話：__________ 聽力_____ 口語_____ 閱讀_____ 寫作_____
其它：__________ 聽力_____ 口語_____ 閱讀_____ 寫作_____

11. 會何種外語? 熟練程度如何? 請用以下標准判斷：1=流利，2=一般流利，3=不流利；並將相應的數字填在空格上。

外語 1 _________ 聽力______ 口語______ 閱讀______ 寫作______
外語 2 _________ 聽力______ 口語______ 閱讀______ 寫作______

12. 幾歲開始學習英語？________歲 學習英語共多少年？________年

13. 在以下情況下，你使用哪種語言？

在家裡: _____% (國語); ______% (台語); _____% (其它, 請說明)
在學校: _____% (國語); ______% (台語); _____% (其它, 請說明)
工作: _____% (國語); ______% (台語); _____% (其它, 請說明)

和朋友在一起: _____% (國語); ______% (台語); _____% (其它, 請說明)

14. 聽力和視覺是否正常：是， 否 （請選擇）

15. 聯系方式：
電話： 電子信箱：

Appendix 3: Oral Instruction to Mandarin listeners in the perception test

- 你将听到三组人说汉语：a.居住在台湾的国语为母语的台湾人,不会英语　b.很早就移民到加拿大的台湾人,英语较好　c.加拿大本地人

-该实验包括三个部分，首先是练习部分，接下来是正式实验的第一部分，第一部分结束后，你有 5 分钟的休息，休息结束后，你要做正式实验的第二部分。

-在实验中，你將聽到一些聲音文件，每個聲音文件代表以下八個詞中的一個：

僻，　綠，　鋪，　怕，　豆，　配，　派，　炮

-請注意看“+”號，並注意聽播放的聲音文件。如：

Trial 4 of 8

+

-聲音文件播放完畢后，代表該聲音文件的詞將出現在屏幕上，

同時出現的還有數字 1-7。如：

綠

差　　　　　　　　　　　　　　好

1　2　3　4　5　6　7

-你的任務是點擊數字鍵 1-7 來判斷你聽到的詞的好壞，1 表示“差”，7 表示“好”。 所谓的 “差” 指的是你听到的词听起来不像母语中的目标词，即它同你脑中的母语目标词有太大的偏差。所谓的 “好” 指的是你听到的词听起来很像母语中的目标词，即它同你脑中的母语目标词十分吻合。
当然，所谓的好坏有程度上的差别，所以我们鼓励你做判断时使用 1-7 当中的所有数字。

-你必須等答案頁出現后再點擊你的選擇。

-你必須在 4 秒鐘內做出選擇。

-請對聽到的每個聲音文件都做出判斷。若錯過某個聲音文件，不用擔心，

請繼續做下一個。

Appendix 4: Oral Instruction to English listeners in the perception test

- You are about to hear English speech samples from three groups of speakers: Mandarin speakers who knew little English; Mandarin speakers whose English was quite good; and native English speakers

-In total, you will do three tests. The first test is a practice test. The second one is the first part of the real test. After you are done with the first part of the real test, you will have a 5-minute break. After the break, you will do the 2nd part of the real test.

-In the tests, you will hear some audio sound files, each of which stands for one of the following seven words: pea, Pooh, paw, dough, pay, pie, pow.

-Focus your attention on the cross and listen to the sounds. For example,

```
Trial 4 of 8

      +
```

-On each screen following the audio clip, you will be presented with the target word that the sound represents. Also presented are numbers 1-7. For example,

pay

bad good

1 2 3 4 5 6 7

- Your task is to judge the goodness of the word you hear on a seven point scale. 1 stands for "bad" pronunciation, 7 for "good" pronunciation. By "bad", we mean the word you hear is a bad exemplar of the native target word; in other words, it does not sound like the native word at all. By "good", we mean it is a good exemplar of the native target word; in other words, it sounds very much like the native word.

As you can see, there are seven numbers on the badness to goodness continuum; you are encouraged to use all of them.

-You must wait until the prompt screen appears before giving your responses!!!

-You will have 4 second to respond before the next clip begins.

- Please respond to every prompt, but do not worry if you miss one. Just go on to the next clip.

Appendix 5: Debriefing questionnaire for Mandarin listeners in the perception test

實驗后問卷調查

問題 A: 在你聽錄音判斷詞的好壞時，你將注意力放在 （请在选项前填写 X）

_____ 1. 詞的長短

_____ 2. 詞的聲調（四聲）

_____ 3. 詞的響亮度

_____ 4. 說話人的音色

_____ 5. 1-4 全部

_____ 6. 其它: ____(請說明)

如果你在問題 A 中選擇的答案超過兩項, 請排列它們在你判定詞的好壞時的重要性。 (請將表示重要性的号码填写在 X 符号前)

問題 B: 關於該實驗，你有什麼建議和意見嗎?

Appendix 6: Debriefing questionnaire for English listeners in the perception test

Debriefing Questionnaire

Question A: What did you attend to when you were listening to the words and did the goodness rating? (please check)

_____ 1. duration

_____ 2. tone

_____ 3. loudness

_____ 4. voice quality

_____ 5. all of the above

_____6. other: ____(please specify)

If you have checked more than two answers in Question A, please rank their importance in your goodness rating? (put the rank numbers beside your check mark)

Question B: What suggestions or comments do you have about the experiment?

Appendix 7: Mean F1a, F2a, F3a (Bark) and standard deviations of Mandarin vowels /a/, /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/ and their English counterparts, /ɒ/, /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/

Vowel	F1a		F2a		F3a	
	MonoM	MonE	MonoM	MonE	MonoM	MonE
/a/ (/ɒ/)	8.37** (.89)	6.68** (.70)	10.97** (.83)	8.74** (.67)	—	—
/aj/	8.20* (.87)	7.38* (.97)	11.53** (.85)	10.30** (1.07)	—	—
/aʊ/	7.91 (.82)	7.79 (.83)	9.96** (.82)	10.94** (.91)	14.97 (.44)	14.61 (.60)
/e/	5.52 (.82)	5.06 (.71)	13.53 (.49)	13.72 (.62)	—	—
/i/	3.70 (.61)	3.34 (.51)	14.60 (.79)	14.69 (.73)	—	—
/oʊ/	6.04* (.50)	5.53* (.46)	9.37** (.64)	10.27** (.72)	15.36** (.48)	14.57** (.48)
/u/	4.17* (.37)	3.76* (.38)	7.60** (.73)	10.05** (1.20)	14.74 (.61)	14.49 (.63)

Note. Values enclosed in parentheses represent standard deviations. Dashes indicate the value was not available for unrounded vowels. *$p < .05$. **$p < .01$.

Appendix 8: Mean ΔF1 and ΔF2 values (Bark) and standard deviations of Mandarin vowels /a/, /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/ and their English counterparts, /ɒ/, /aj/, /aʊ/, /e/, /i/, /oʊ/, /u/

Vowel	ΔF1		ΔF2	
	MonoM	MonE	MonoM	MonE
/a/ (/ɒ/)	.02 (.17)	.02 (.33)	-.05 (.13)	.03 (.25)
/aj/	-1.91 (1.06)	-2.20 (.94)	1.51** (.46)	3.05** (.67)
/aʊ/	-1.02 (.68)	-1.75 (1.06)	-.80** (.40)	-2.17** (.72)
/e/	-1.33 (.49)	-1.06 (.55)	.60 (.25)	.54 (.18)
/i/	-.50** (.31)	-.18** (.25)	.06 (.15)	.01 (.14)
/oʊ/	-.91 (.40)	-1.06 (.33)	-1.25** (.48)	-2.16** (.67)
/u/	-.40 (.40)	-.36 (.33)	-.74 (.37)	-1.11 (.62)

Note. Values enclosed in parentheses represent standard deviations. **$p < .01$.

Appendix 9: Mean F1a, F2a, F3a (Bark) and standard deviations of English vowel production by Mandarin-English bilinguals of high L1 use(BiMH), Mandarin-English bilinguals of low L1 use(BiML) and English monolinguals(MonoE)

Vowel	F1a			F2a			F3a		
	BiMH	BiML	MonoE	BiMH	BiML	MonoE	BiMH	BiML	MonoE
/ɒ/	7.14 (1.10)	6.73 (1.15)	6.68 (.70)	9.45 (1.14)	8.96 (1.08)	8.74 (.67)	—	—	—
/aj/	8.16 (1.11)	7.95 (1.16)	7.38 (.97)	11.24 (1.11)	10.66 (.95)	10.30 (1.07)	—	—	—
/aʊ/	7.59 (1.18)	7.63 (1.03)	7.79 (.83)	10.11 (1.32)	10.39 (1.02)	10.94 (.91)	15.15 (.58)	14.99 (.67)	14.61 (.60)
/e/	4.69 (.46)	4.98 (.73)	5.06 (.71)	14.21 (.82)	13.96 (.73)	13.71 (.62)	—	—	—
/oʊ/	5.61 (.43)	5.60 (.81)	5.53 (.46)	9.67 (.95)	9.55 (.85)	10.27 (.72)	15.19 (.63)	14.99 (.99)	14.57 (.48)
/ʌ/	7.17 (.92)	7.06 (.88)	6.79 (.93)	10.38 (.89)	10.25 (.87)	10.29 (1.04)	—	—	—
/i/	3.37 (.55)	3.38 (.59)	3.34 (.51)	14.87 (1.02)	14.88 (.96)	14.69 (.73)	—	—	—
/ɪ/	4.92 (.53)	5.02 (.70)	5.15 (.64)	13.30 (.96)	13.38 (.90)	13.31 (.54)	—	—	—
/u/	3.92 (.54)	3.93 (.51)	3.76 (.38)	8.88* (1.40)	9.06 (1.06)	10.05 (1.20)	14.96 (.66)	14.85 (.74)	14.49 (.63)
/ʊ/	4.95 (.60)	5.32 (.75)	5.57 (.90)	10.62 (1.00)	10.12 (.97)	10.27 (1.15)	15.22 (.68)	15.16 (.75)	14.95 (53)
/æ/	8.23 (1.10)	8.18 (.96)	7.84 (.91)	12.47 (.79)	12.16 (.77)	12.05 (.60)	—	—	—
/ɛ/	7.13 (.84)	6.99 (.79)	6.93 (.92)	12.79 (.78)	12.48 (.79)	12.67 (.55)	—	—	—

Note. Values enclosed in parentheses represent standard deviations. Dashes indicate the value was not available for unrounded vowels. *p <.05.

Appendix 10: Mean ΔF1 and ΔF2 values (Bark) and standard deviations of English vowel production by Mandarin-English bilinguals of high L1 use(BiMH), Mandarin-English bilinguals of low L1 use(BiML) and English monolinguals(MonoE)

Vowel	ΔF1			ΔF2		
	BiMH	BiML	MonoE	BiMH	BiML	MonoE
/ɒ/	-.49 (.61)	-.41 (.68)	.02 (.33)	-.29 (.53)	-.28 (.91)	.04 (.25)
/aj/	-3.04 (1.11)	-2.48 (1.00)	-2.20 (.94)	2.37* (.71)	2.73 (.71)	3.05 (.67)
/aʊ/	-1.47 (1.20)	-1.42 (.79)	-1.75 (1.06)	-1.18** (.81)	-1.18** (.81)	-2.17 (.72)
/e/	-.83 (.29)	-.99 (.54)	-1.06 (.55)	.52 (.32)	.60 (.32)	.54 (.18)
/oʊ/	-1.03 (.52)	-1.24 (.57)	-1.06 (.33)	-1.49* (.63)	-1.58* (.53)	-2.16 (.67)
/ʌ/	-.48 (.41)	-.46 (.44)	-.52 (.39)	-.06 (.21)	.03 (.25)	-.13 (.19)
/i/	-.31 (.40)	-.17 (.25)	-.18 (.25)	.06 (.14)	.09 (.18)	.01 (.14)
/ɪ/	-.26** (.27)	-.40** (.47)	.22 (.30)	-.09* (.11)	-.04** (.12)	-.24 (.19)
/u/	-.42 (.30)	-.46 (.30)	-.36 (.33)	-.48* (.47)	-.77 (.43)	-1.11 (.62)
/ʊ/	-.11 (.27)	-.28 (.44)	.06 (.40)	1.18 (.65)	1.50 (.68)	1.42 (.57)
/æ/	-.14 (.17)	-.04 (.34)	-.08 (.39)	-.10 (.20)	-.18 (30)	-.26 (.36)
/ɛ/	-.41 (.43)	-.57* (.50)	-.12 (.34)	-.02 (.22)	-.01 (.14)	-.18 (.26)

Note. Values enclosed in parentheses represent standard deviations. $^{**}p < .01$. $^{*}p < .05$.

Appendix 11: Mean F1a, F2a, F3a (Bark) and standard deviations of Mandarin vowel production by Mandarin-English bilinguals of high L1 use(BiMH), Mandarin-English bilinguals of low L1 use(BiML) and Mandarin monolinguals(MonoM)

Vowel	F1a			F2a			F3a		
	BiMH	BiML	Mono M	BiMH	BiML	Mono M	BiMH	BiML	Mono M
/a/	8.32 (1.07)	8.18 (.90)	8.37 (.89)	10.99 (1.03)	10.65 (.89)	10.97 (.83)	—	—	—
/aj/	8.12 (.97)	8.14 (.85)	8.20 (.87)	11.75 (.88)	11.39 (1.04)	11.53 (.85)	—	—	—
/aʊ/	7.73 (.95)	7.80 (.83)	7.91 (.82)	9.77 (.79)	10.16 (.95)	9.96 (.82)	15.23 (.72)	15.05 (.56)	14.97 (.44)
/e/	5.14 (.58)	5.30 (.72)	5.52 (.82)	13.86 (.73)	13.74 (.52)	13.53 (.49)	—	—	—
/i/	3.41 (.63)	3.37 (.55)	3.70 (.61)	14.82 (1.03)	14.75 (.95)	14.60 (.79)	—	—	—
/oʊ/	5.84 (.54)	5.76 (.64)	6.04 (.50)	9.51 (.80)	9.51	9.37 (.64)	15.36 (.48)	15.32 (.63)	15.36 (.48)
/u/	4.05 (.51)	4.01 (.46)	4.17 (.37)	7.29 (.78)	7.56 (.82)	7.60 (.73)	14.92 (.95)	14.92 (.88)	14.74 (.61)
/y/	3.54 (.65)	3.58 (.50)	3.87 (.54)	13.32 (.55)	13.34 (.69)	13.35 (.51)	14.88 (.73)	14.92 (.73)	14.82 (.41)

Note. Values enclosed in parentheses represent standard deviations. Dashes indicate the value was not available for unrounded vowels. **p* <.05.

Appendix 12: Mean ΔF1 and ΔF2 values (Bark) and standard deviations of Mandarin vowel production by BiML (Mandarin-English bilinguals of low L1 use), BiML (Mandarin-English bilinguals of high L1 use) and MonoM (Mandarin monolinguals)

Vowel	ΔF1			ΔF2		
	BiMH	BiML	MonoM	BiMH	BiML	MonoM
/a/	-.08 (.33)	-.01 (.21)	.02 (.17)	-.04 (.12)	-.03 (.13)	-.05 (.13)
/aj/	-2.16 (1.01)	-2.38 (0.79)	-1.91 (1.06)	1.59 (.34)	2.04* (.72)	1.51 (.46)
/aʊ/	-1.09 (.55)	-1.15 (.55)	-1.02 (.68)	-.82 (.44)	-1.06 (.65)	-.80 (.40)
/e/	-1.07 (.42)	-1.21 (.37)	-1.33 (.49)	.48 (.38)	.67 (.28)	.60 (.25)
/i/	-.18* (.22)	-.18* (.38)	-.50 (.31)	-.04 (.10)	-.05 (.11)	-.06 (.15)
/oʊ/	-78 (.37)	-96 (.45)	-91 (.40)	-1.23 (.52)	-1.35 (.41)	-1.25 (.48)
/u/	-33 (.37)	-34 (.27)	-.40 (.40)	-.41 (.39)	-60 (52)	-.74 (.37)
/y/	-.26 (.35)	-.20 (.26)	-.37 (.27)	.05 (.15)	.11 (.14)	.16 (.19)

Note. Values enclosed in parentheses represent standard deviations. $^*p < .05$.

Appendix 13: Goodness ratings assigned to Mandarin monolinguals by Mandarin listeners

	/a/	/aj/	/aʊ/	/e/	/i/	/oʊ/	/u/	/y/
speaker	Rating							
101	5.31	5.63	5.63	5.69	6.19	5.81	5.56	6.06
102	4.44	5.25	5.25	5.63	5.63	5.56	6	6
103	6.31	5.75	5.75	5.25	5.87	6.38	5	5.75
104	5.19	6	6	5.25	5	5.13	5.06	5.67
105	3.75*	2.75*	2.75*	3.19*	5.31	4.69	4.5	5.33
106	5.88	5	5	5.5	5.75	4.81	5.19	5.38
107	4.75	4.63	4.63	3.81	4.25	5.06	4.38*	4.88*
108	5.69	5.06	5.06	6.38	5.75	6.19	5.38	5.6
109	5.88	4	4	4.69	5	5.19	5.38	5.81
110	5.44	5.13	5.13	4.88	5.75	5.50	4.69	5.63
111	5.31	4.25	4.25	5.81	5.06	5.69	5.5	5.81
113	5.44	5.25	5.25	6	5.56	5.69	5.38	5.38
114	5.81	4.56	4.56	5.69	5.94	5.56	5.44	5.81
Mean (S.D)	5.32 (.68)	4.87 (.86)	4.87 (.86)	5.21 (.89)	5.47 (.52)	5.48 (.50)	5.19 (.46)	5.62 (.32)

Note. Numbers in shade and with asterisk are ratings that are two standard deviations below the Mandarin monolinguals' group mean. "S. D" stands for "standard deviation".

Appendix 14: Goodness ratings assigned to Mandarin-English bilinguals by Mandarin listeners on a 7-point scale (1= bad; 7= good)

	/a/	/aj/	/aʊ/	/e/	/i/	/oʊ/	/u/	/y/
speaker	Rating							
201	6.5	6.44	4.69	5.75	6.06	6.44	5.81	5.25
202	4.44	4.33	3.5	4.31	4.63	5.20	5	5.5
203	4.47	5.19	4.75	4.13	5	5.25	4.75	4.44*
204	5.19	5.94	5.56	4.75	5.13	4.88	5.81	4.73*
205	5.88	5.69	6.06	5.8	5.56	5.57	5.88	6
206	4.93	4.13	4.31	4.56	5.56	6.06	4.94	4.87*
207	5.29	5.31	4.5	5.8	6.13	6.44	5.88	5.87
209	6.13	5.25	3.19	5.94	5.44	5.69	5.56	4.8*
210	4.63	3.88	3.33	4.88	5.19	4.94	4.25*	4.88*
211	6.06	5.06	5.63	5.88	6	5.69	5.75	5.56
212	6.31	6.07	5.81	6.19	5.94	6.50	5.81	5.63
213	5.67	5.13	5.25	4.06	5.56	4.38*	5.81	5.06
214	4.81	5.25	4.38	5.31	6	5.81	5.25	5.38
215	5.94	5.56	4.44	5.75	5.56	5.75	5.69	5.88
216	4.69	4.94	3.47	4.25	5.06	4.38*	5	4.88*
217	5.38	5.19	4.31	4.13	5.27	4.57	4.33	4.88*
218	3.8*	5.50	4.25	5.38	5.06	5.44	4.77	5.81
219	5.75	4.38	3.44	5.06	4.88	5.75	4.64	4.06*
220	5.13	4.88	4.63	5.69	4.75	5.38	4.88	5.56
221	4.56	4.88	3.44	4.75	4.5	5.13	4.79	5.38
222	5.25	5.13	5.31	5.56	4.94	6.00	5.75	6.06
223	4.63	5.81	5.31	5.5	4.56	4.81	5	3.31*
224	5	5.63	4.38	5.19	3.5*	4.13*	4.8	5.06
225	5.94	4.88	6.13	5.63	5.38	5.81	5.06	5.56
226	4.75	4.31	3.88	5.56	5.94	4.13*	4.33	4.94*
227	3.69*	3.25	3.69	2.86*	3.87*	4.81	5.33	3.75*
228	6.31	6.06	4.31	5.63	4.94	5.94	4.5	5.53
229	5.25	2.38*	3.81	4.5	3.94*	3.94*	4.29	4.88*
230	4.31	4.88	5.69	5.38	5.25	5.73	5.75	5.56
231	5.25	4.63	2.13*	3.06*	4.94	5.31	5	5.56
232	2.94*	4.31	3.6	5	3.94*	5.50	5.5	4.88*
233	2.88*	4.50	4.69	4	4.69	4.75	4.8	5.38
234	4.13	5.88	4.5	6.19	6.19	5.31	5.75	5.44
MonoM Mean (S.D)	5.32 (.68)	4.87 (.86)	4.87 (.86)	5.21 (.89)	5.47 (.52)	5.48 (.50)	5.19 (.46)	5.62 (.32)

Note. Numbers in shade and with asterisk indicate ratings that are two standard deviations below the Mandarin monolinguals' group mean. "S. D" stands for "standard deviation".

Appendix 15: Goodness ratings assigned to Mandarin monolinguals by English listeners on a 7-point scale (1= bad; 7= good)

	/aj/	/aʊ/	/e/	/oʊ/	/u/
speaker					
101	4.94	3.44	4.56	3.63	4.25
102	3.56	2.93	3.94	3.60	4.00
103	4.38	3.19	4.06	3.75	3.38
104	5.19	4.00	5.56*	3.94	4.81
105	3.75	3.56	4.50	3.31	4.63
106	3.06	3.81	5.44*	3.88	4.81
107	3.81	3.19	4.94	3.75	4.44
108	4.67	4.69	4.81	4.00	4.88
109	6.13*	2.13	4.06	3.69	3.75
110	5.56	5.50*	4.25	3.67	4.25
111	4.06	5.13	4.75	3.63	3.88
113	4.38	3.81	4.31	3.56	4.50
114	4.94	2.13	5.81*	3.06	4.63
MonoEMean (S.D)	6.41 (.27)	5.97 (.40)	6.1 (.38)	6.41 (.15)	6.28 (.22)

Note. Numbers in shade and with asterisk are ratings that are within two standard deviations of the English monolinguals' group mean. "S. D" stands for "standard deviation".

Appendix 16: Goodness ratings assigned to Mandarin-English bilinguals by English listeners on a 7-point scale (1= bad; 7= good)

	/aj/	/aʊ/	/e/	/oʊ/	/u/
speaker	Rating				
201	5.69	4.63	4.75	3.40	3.88
202	4.75	4.88	5.56*	4.89	5.56
203	5.25	3.88	6.25*	5.10	4.88
204	3.94	4.20	4.81	3.20	4.50
205	5.25	4.00	5.19	4.10	4.69
206	6.06*	5.38*	4.75	4.40	5.13
207	3.75	3.75	5.00	3.00	4.75
209	2.88	2.19	4.13	3.55	4.56
210	4.00	5.63*	5.56*	4.10	4.25
211	5.06	3.81	4.63	3.30	4.31
212	4.69	3.81	3.44	2.85	4.25
213	5.75	4.75	5.00	3.80	3.63
214	4.06	4.94	5.44*	3.70	3.88
215	5.50	4.57	4.88	3.80	3.69
216	5.31	3.94	4.94	4.80	5.25
217	5.00	4.81	4.69	2.85	4.75
218	5.06	4.31	4.19	5.05	4.31
219	6.00*	5.00	5.81*	3.55	4.81
220	5.44	4.31	5.50*	3.55	4.25
221	5.13	4.31	5.63*	5.30	5.75
222	5.06	4.38	5.00	3.65	4.50
223	4.13	4.19	4.69	3.20	3.25
224	4.00	4.40	4.19	3.35	4.50
225	5.38	4.44	5.13	3.85	4.13
226	5.81	4.56	5.56*	3.65	3.88
227	5.00	4.81	5.00	4.80	4.88
228	4.69	2.94	5.50*	3.37	4.88
229	5.38	5.50*	6.19*	4.55	4.88
230	5.94*	4.56	4.75	3.95	4.69
231	4.31	5.06	3.47	2.30	4.25
232	5.67	5.06	4.69	4.35	5.25
233	6.31*	5.00	4.13	4.70	5.56
234	4.38	2.13	4.47	2.65	3.81
MonoE Mean (S.D)	6.41 (.27)	5.97 (.40)	6.1 (.38)	6.41 (.15)	6.28 (.22)

Note. Numbers in shade and with asterisk indicate ratings that are within two standard deviations of the English monolinguals' group mean. "S. D" stands for "standard deviation".

REFERENCES

Baker W., & Trofimovich P. (2005). Interaction of native and second language vowel system(s) in early and late bilinguals. *Language and Speech*, *48*, 1–27.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman, & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore: York Press.

Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.3) [Computer program]. Retrieved May 26, 2005, from http://www.praat.org/

Bongaerts, T., Mennen, S., & van der Slik, F. (2000). Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica*, 54: 2, 298-308.

Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press

Cheng, R. 1966. Mandarin phonological structure. *Journal of Linguistics*, *2*, 135-262.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly, 15*, 443-455.

Flege, J. & Hillenbrand, J. (1984). Limits on phonetic accuracy in adult foreign language speech production. *Journal of the Acoustical Society of America*, *76*, 708-721.

Flege, J. E. (1987). The production of "new" and "similar" phonemes in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47-65.

Flege, J. & Eefting, W. (1987). The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics*, *15*, 67-83.

Flege, J. & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America*, *83*, 729-740.

Flege, J. E. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Heubner, & C. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theory.* (pp. 249-284). Philadelphia: John Benjamins.

Flege, J. E. (1992). Speech learning in a second language. In C. Ferguson, L. Menn, & C. Stoel-Gammon (eds.), *Phonological development: Models, research, and application* (pp. 565-604). Timonium, MD: York Press.

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E., Munro, M. J. & MacKay, I. R. A. (1995). Factors affecting the strength of perceived foreign accent in a second language. *Journal of Acoustical Society of America*, *97*, 3125-3134.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, *25*, 169-186.

Flege, J.E., Bohn, O., Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*, 437-470.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In Schiller, Niels O., & Meyer Antje S. (eds.). *Phonetics and Phonology in Language Comprehension and Production.* Mouton de Gruyter.

Flege, J. E., Schirru, C. & MacKay, I. (2003) . Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*, 467-491.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J.I. Hualde (Eds.), *Laboratory Phonology 9.* Berlin: Mouton de Gruyter.

Gandour, J. T. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics, 11*, 149-175.

Gandour, J. T. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics, 12*, 235-261.

Grosjean, F. (1989). Neurolinguists, beaware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*, 3-15.

Grosjean, F. (1992). Another view of bilingualism. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp.51-63). Amsterdam: North-Holland.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.). *One Mind, Two Languages: Bilingual Language Processing* (pp. 1-22). Oxford: Blackwell.

Guion, S. G., Flege, J. E., & Loftin, J.D. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics. 28,* 27-42

Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second language. *Phonetica, 60*, 98-128.

Halle, M. (1959). *The sound pattern of Russian: A linguistic and acoustical investigation.* The Hague: Mouton. Second printing,1971.

Harada, T. (2003). L2 influence on L1 speech in the production of VOT. In M. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)* (pp. 1085-1088). Barcelona: Universitat Autónoma de Barcelona

Howie, J. (1976). *Acoustical studies of Mandarin vowels and tones.* London: Cambridge University Press.

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using multidimensional scaling. *Journal of the Acoustical Society of America, 97*, 553-562.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics, 50,* 93-107.

Kuhl, P. K., & Iverson, P. (1995). Linguistic experience and the "perceptual magnet effect". In W.Strange (Eds.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121-154). Baltimore: York Press.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606-608.

Lado, R. (1957). *Linguistics Across Cultures.* Ann Arbor: University of Michigan Press

Lin, Y. H. (1989). *Autosegmental Treatment of Segmental Processes in Chinese Phonology.* Unpublished PhD dissertation, University of Texas, Austin.

Major, R. (1992). Losing English as a first language. *The Modern Language Journal, 76*, 190-208.

Major, R. (1993). Sociolinguistic factors in loss and acquisition of phonology. In K. Hyltenstam & A. Viberg (Eds.), *Progression and regression in language: Sociocultural, neuropsychological and linguistic perspectives* (pp. 463-478). Cambridge: Cambridge University Press.

McRobbie, Z. (2003). Sound change in progress: An acoustic study of vowel variation in the speech of Hungarians in Vancouver, British Columbia. In Z. McRobbie & C. So (Eds.), *Proceedings of the 13th Conference of the Finno-Ugric Studies Association of Canada (pp. 83-91).* Burnaby: Simon Fraser University.

Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of Chinese Language Teachers Association. 24*, 49-65.

Munro, M. J. (1992). *Perception and production of English vowels by native speakers of Arabic.* Unpublished PhD Thesis. University of Alberta.

Munro, M. J., Derwing, T. M., & Flege, J. E. (1999). Canadians in Alabama: a perceptual study of dialect acquisition in adults. *Journal of Phonetics. 27*, 385-403.

Nearey, T., & Assmann, P. (1986). Modeling the role of inherent spectral change in vowel identification. *Journal of the Acoustical Society of America, 80,* 1297-138.

Nunnaly, J. (1978). *Psychometric theory.* New York: McGraw-Hill.

Pavlenko, A. (2000). L2 influence on L1 in late bilingualism. *Issues in Applied Linguistics, 11*(2), 175-205.

Peng, S. H. (1993). Cross-language influence on the production of Mandarin /f/ and /x/ and Taiwanese /h/ by native speakers of Taiwanese Amoy. *Phonetica, 50*, 245-260.

Piske, T. & MacKay, I. (1999) Age and L1 use effects on degree of foreign accent in English. *Proceedings of the 14th International Congress of Phonetic Sciences,* San Francisco, 1-7 August 1999, pp. 1433-1436.

Rose, P. (1987). Considerations in the normalization of the fundamental frequency of linguistic tone. *Speech Communication, 6*, 343-351.

Sancier, M.L. & Fowler, C. A. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics. 35*, 421–436.

Schiller, N. O., & Meyer A. S. (eds.). (2003). *Phonetics and Phonology in Language Comprehension and Production.* Mouton de Gruyter.

Shira, S. (2003). DurationFormants_F.praat. University of Washington. Retrieved from the Web January 10, 2007.

http: /home.myuw.net/ssetsuko/durationFormants_F.praat.

Stevens, K. N. (1999). *Acoustic Phonetics.* Cambridge, MA: The MIT Press (Current Studies in Linguistics, 30).

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of Acoustical Society of America. 88*, 97-100.

Wan, I-P. & Jaeger, J. (2003). The phonological representation of Taiwan Mandarin vowels: a psycholinguistic study. *Journal of East Asian Linguistics. 12*, 205-257.

Wang, X. (1997). *The acquisition of English vowels by Mandarin ESL learners: A study of production and perception.* Unpublished Master's Thesis, Simon Fraser University.

Wang, Y. (1995). *Voice-onset time in English stops: A comparison of productions by American English and Mandarin Chinese natives.* Unpublished Master's Thesis. Norwegian University of Science and Technology.

Wang, Y., Spence, M., Jongman, A., and Sereno, J. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America. 106*, 3649-3658.

Wang, Y., Jongman, A., and Sereno, J. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America. 113,* 1033-1043.

Weinreich, U. (1953). *Language in contact.* The Hague: Mouton.

Welby, P. (2003). Label-vowel. Praat. Ohio State Univeristy. Retrieved from the web January 10 2007. http://www.icp.inpg.fr/~Welby/PRAAT/label-vowel.praat

White, C. M. (1981). Tonal perception errors and interference from English Intonation. *Journal of Chinese Language Teachers Association. 16*, 27–56.

Wikipedia, the free encyclopedia. (n.d.). *Taiwanese (linguistics).* Retrieved June, 6, 2008, from http://www.wikipedia.org/wiki/Taiwanese_(linguistics).

Williams, L. (1980). Phonetic variation as a function of second-language learning. *Child Phonology. 2*, 185-215.

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

Wu, Z. J. (1992). *An Introduction to Modern Chinese Phonetics.* Beijing: Sinolingual.

Wu, Z., & Lin, M. (1989). *Shi Yan Yu Yin Xue Gai Yao (Introduction to Experimental Phonetics)*. Beijing: Gao Deng Jiao Yu Chu Ban She (Higher Education Press)

Yeni-Komshian, G. H., Flege, J. E. & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition. 3*, 131-49.