Journal of Phonetics 65 (2017) 60–76

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

L1–L2 interactions of vowel systems in young bilingual Mandarin-English children

Jing Yang [a,*], Robert A. Fox [b]

[a] Department of Communication Sciences and Disorders, University of Central Arkansas, Conway, AR 72035, USA
[b] Department of Speech and Hearing Science, The Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 16 August 2016
Received in revised form 28 May 2017
Accepted 2 June 2017

Keywords:
Mandarin-English bilingual children
L1–L2 interactions
Vowel production

ABSTRACT

This study examined the influence of L1 (Mandarin)–L2 (English) interactions on the organization of vowel systems and fine-grained spectral features of vowel productions in young bilingual Mandarin-English children. The participants included 39 children (15 bilinguals, 15 Mandarin monolinguals, and 9 English monolinguals) at 5–6 years of age. The bilingual children were divided into Bi-low (at the early stage of English learning with low proficiency in English) and Bi-high (highly proficient in English) groups. Each participant was recorded producing one set of Mandarin words containing /a, i, u, y, ɤ/ and/or one set of English words containing /i, ɪ, e, ɛ, æ, u, ʊ, o, ɑ, ʌ/. Formant frequencies at five temporal locations were measured. Both static (midpoint formant values) and dynamic (formant movement pattern, trajectory length) acoustic properties were examined. Bi-low children showed a strong effect of L1 on L2. The L1 features were maintained and transferred to the new phonetic system. Bi-high children produced L2 vowels in a near-native manner. Meanwhile, they tended to transfer some L2 features to their L1 and moved the L1 vowels closer to L2 vowels, which suggested an assimilatory process. Both static and dynamic spectral features were affected by L1–L2 interactions.

1. Introduction

Based on the common belief of “earlier is better,” a large body of research has revealed that early L2 speakers, who started to acquire an L2 at a younger age, can produce and differentiate L2 sounds better than late L2 speakers, who started to acquire an L2 in adulthood (e.g. Aoyama, Guion, Flege, Yamada, & Akahane-Yamada, 2008; Baker, Trofimovich, Flege, Mack, & Halter, 2008; MacLeod, Stoel-Gammon, & Wassink, 2009; Meador, Flege, & MacKay, 2000). Some researchers reported that bilinguals who started to acquire an L2 before puberty showed little or no difference in the production of the L2 sounds or the discrimination of L2 contrasts from the native speakers (Flege, MacKay, & Meador, 1999; Mack, 1989; MacLeod & Stoel-Gammon, 2005). In addition to the monolingual-like ability of L2, some researchers found that early bilingual speakers were able to retain monolingual-like ability in their L1 (MacLeod et al., 2009). For those studies that found no difference between early bilinguals and monolinguals of each language, the underlying assumption was that the early bilinguals could develop two separate phonetic systems independently.

However, many previous studies have suggested the interdependence of the two sub-systems in bilingual speakers (Grosjean, 1989; Keshavarz & Ingram, 2002; Paradis, 2001; Paradis & Genesee, 1996). Over the past few decades, the dynamic and continuing interaction between L1 and L2 has been intensively examined in adult speakers who started to learn an L2 at various ages (Baker & Trofimovich, 2005; Flege, Schirru, & MacKay, 2003; Kartushina, Hervais-Adelman, Frauenfelder, & Golestani, 2016). However, to what extent and in which manner the L1–L2 interactions affect the formation of phonetic systems in young bilingual children is not fully understood. So far, researchers have reported the development and organization of phonetic systems in English-learning children from diverse language backgrounds (e.g. Korean-English, Japanese-English, French-English, Arabic-English bilingual children) (Khattab, 2000; Lee & Iverson, 2012; Netelenbos & Li, 2013; Oh et al. 2011; Turner, Netelenbos, Rosen, & Li, 2015). Little research has been done to examine vowel production in young Mandarin-English

* Corresponding author at: Department of Communication Sciences and Disorders, University of Central Arkansas, 201 Donaghey Ave, Conway, AR 72035, USA.
E-mail address: jyang@uca.edu (J. Yang).

http://dx.doi.org/10.1016/j.wocn.2017.06.002
0095-4470/

bilinguals. The present study aims to expand previous studies by investigating the organization of vowel systems and the change of static and dynamic spectral features of vowel productions under the influence of L1–L2 interactions in young bilingual Mandarin-English children.

1.1. Mutual influence between L1 and L2 phonetic systems

According to the Speech Learning Model (SLM, Flege, 1995), a bilingual's L1 and L2 systems coexist in a common phonological space. Following this assumption, L1 and L2 would naturally influence each other. In particular, L1 and L2 sounds interact with each other at the allophonic level rather than the phonemic level. When the interactions between L1 and L2 are strong enough, the phonetic structures and the perceptual organization of both L1 and L2 in the bilinguals may change and show different patterns from those of monolinguals of each language. A large number of studies have evidenced that simultaneous bilinguals and early sequential bilinguals may still differ from monolinguals in their production and perceptual abilities (Højen & Flege, 2006; Pallier, Bosch, & Sebastian-Gallés, 1997; Sebastian-Gallés & Soto-Faraco, 1999; Tsukada et al., 2005).

There are two types of interaction effects: an effect of L1 on L2 and vice versa. As Watson (1991) pointed out, for sequential bilinguals who learn an L2 after a complete or relatively complete acquisition of L1, their native language (L1) is used as the base on which to establish the L2 phonetic system. In this case, bilinguals are likely to assimilate the phonetically similar L2 sounds into established L1 categories at the beginning of L2 acquisition. However, not every L2 sound is equally assimilated to an L1 sound category. The extent to which the L2 sound is assimilated to the L1 sound category is primarily determined by the phonetic-acoustic similarity between the L2 and L1 sounds (Perceptual Assimilation Model-PAM, Best, 1994, 1995; SLM, Flege, 1995). Following continuous exposure and immersion in L2, the influence of L1 on L2 is attenuated, and separate L2 sound categories may eventually be established.

Unlike the convergent findings regarding the effect of L1 on L2 in early stage L2 speakers, researchers hold divergent opinions on the influence of L2 on L1. Bohn and Flege (1992), for example, observed no change of L1 sounds as a function of L2 experience. By contrast, many other studies have reported a change in L1 as a result of L2 learning in both perception (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Hazan & Boulakia, 1993) and production (Flege, 1987; Flege & Hillenbrand, 1984; Harada, 2003; Kang & Guion, 2006; Mack, 1990). Among these studies, two opposite processes have been observed: dissimilatory processes and assimilatory processes. A dissimilatory process takes place when a new phonetic category in L2 has been established and speakers shift a phonetically similar L1 sound away from the L2 sound to make a contrast between the two sounds (Guion, 2003; Mack, 1990; Oh et al., 2011; Yusa et al., 2010). Guion (2003) found that Quichua (L1) vowels systematically raised and moved away from the similar L2 (Spanish) vowels in Quichua-Spanish bilinguals who had developed distinct vowel categories for the L2. However, researchers have also found instances of a shift in an L1 sound towards an acoustically similar L2 sound, which represents an assimilatory process (Chang, 2012; Flege, 1987). Flege (1987) found the forward movement of French (L1) /u/ as a result of influence from English (L2) /u/ that was located in a more fronted position relative to French /u/. Chang (2012) found that only six weeks' intensive immersion in Korean (L2) resulted in noticeable assimilatory modification of most English (L1) sounds in native English speakers. Kartushina et al. (2016) also reported a drift of native vowels (French vowel /ø/ and /y/) to non-native vowels (Russian /ɨ/) after short-term visual articulatory feedback training in native French speakers.

Efforts have been made to uncover the underlying driving forces of dissimilatory and assimilatory processes of L1 sounds as a function of L2 immersion (Chang, 2012, 2013). Although no consensus has been reached regarding this, researchers have shown that the magnitude and direction of L1–L2 interaction in bilinguals are highly correlated with the starting age of L2 learning and the amount of L2 experience. Baker and Trofimovich (2005) examined vowel production in Korean-English bilinguals differing in age of exposure to L2 (early and late) and the amount of experience with L2 (+1 and +7 years) in comparison to monolingual speakers. They found that late bilinguals' L1 (Korean) vowels showed no differences from monolingual Korean speakers while their L2 (English) vowels were quite different from those produced by monolingual English speakers as a result of the influence of their L1. However, for the early bilinguals, the inexperienced speakers (+1 year experience with L2) maintained their L1 (Korean) vowels, but they were less likely to separate the L2 (English) vowels. The experienced speakers (+7 years' experience with L2) better separated the vowel categories in the L2, but they differed from monolingual speakers of each language in certain vowels. These findings demonstrated that early bilinguals, especially experienced early bilinguals, showed different features from monolinguals of each language as a result of bi-directional influence between their L1 and L2. By contrast, late bilinguals showed a unidirectional influence of their L1 on the L2.

In addition to the bi-directional interaction, researchers also reported distinct patterns of phonetic drift for L2 sounds in early bilinguals and late bilinguals. For example, Flege et al. (2003) compared the production of English /eɪ/ and Italian /e/ by Italian (L1)–English (L2) bilinguals varying in the age of learning (early and late) and the amount of English usage (low and high). The English /eɪ/ produced by native English speakers is characterized by a greater formant movement relative to a typical Italian /e/. The authors found that whereas the early-low bilinguals tended to dissimilate the L2 sound from the L1 counterpart by overshooting the English /eɪ/ with greater formant movement, the late bilinguals always assimilated the L2 sound to similar L1 sound by undershooting English /eɪ/ with less formant movement. These findings were, to a large extent, based on adult and adolescent L2 speakers. Previous studies have shown continuing refinement of acoustic properties in children's vowel production (Yang & Fox, 2013) and the privilege of neural plasticity in young children as opposed to adults in language learning (e.g. Stiles, 2000). In the present study, we were interested in how the L1–L2 interactions would be manifested in young bilingual children.

1.2. Organization of L1 and L2 vowel systems in bilingual children

To date, the separation of phonetic systems in bilingual children has been documented in bilingual children from different language backgrounds (Khattab, 2000; Lee & Iverson, 2012; Netelenbos & Li, 2013; Turner et al., 2015; Yang, Fox, & Jacewicz, 2015). Among these studies, Lee and Iverson (2012) investigated the formation of vowel categories in Korean-English bilingual children. The authors recruited two groups of bilingual children (5-year-olds and 10-year-olds) and compared the formant values of vowel productions in each language with corresponding age-matched monolingual children. The results revealed that both 5- and 10-year-old bilingual children produced Korean and English similarly to corresponding monolingual children. Children as young as five years of age could separate the two vowel systems in their speech production. In addition, the authors identified both assimilatory and dissimilatory movements of Korean and English vowels in the 10-year-old bilingual children.

In another longitudinal study on the vowel development of Japanese-English bilingual speakers, Oh and colleagues (2011) compared the vowel duration and formant values between English-learning native Japanese speakers (adults and children) and native English speakers (adults and children) over a one-year-period. The results indicated that while the Japanese adults showed little change in their production of English vowels, the Japanese children improved their production of the new English vowels and produced them in a more native-like manner. Meanwhile, the children's production of native (Japanese) vowels /i/, /a/ and /u/ changed following the mastery of non-native (English) vowels /i/, /ɪ/, /ɑ/, /ʌ/, /ʊ/. Moreover, both assimilation and dissimilation processes were observed in the children's production of English and Japanese vowels, which suggested the more malleable L1 system in young L2 learners.

These studies, together with the abovementioned bilingual studies about adult L2 speakers, suggest that young bilingual children are more likely to establish a native-like phonetic system than late L2 speakers in a shorter period. Meanwhile, by virtue of the non-stabilized L1 system, the acquisition of a new language is more likely to generate noticeable changes in their native language. However, since most previous studies on bilingual children's speech production chose participants with a relatively long period of L2 immersion, very few examined the starting stage of L2 learning in young children. In addition, as the majority of bilingual studies focused on the vowel production represented by the midpoint formant frequency values, little research has been conducted to investigate the vowel dynamic spectral features in young bilingual children. Addressing these gaps, the present study recruited two groups of young sequential bilingual Mandarin-English children: one group had lived in the U.S. less than six months; one group had at least two years of residency in the U.S. and was highly immersed in the English language. The comparison of novice and proficient bilingual children enables us to better track the process of speech development and the separation of L2 from L1 in bilingual children.

In addition to conventional midpoint formant frequency values, the present study compared the dynamic spectral features between bilingual children and monolingual children. As revealed by some previous work (e.g. Harrington & Cassidy, 1994; Hillenbrand, Getty, Clark, & Wheeler, 1995; Nearey, 2013; Nearey & Assmann, 1986), even monophthongs show systematic patterns of spectral change regardless of phonetic context. This type of spectral change characterizes the inherent feature of vowel quality. However, although this property has been addressed in native English children (Assmann & Katz, 2000; Assmann, Nearey, & Bharadwaj, 2013; Jacewicz, Fox, & Salmons, 2011a), it has rarely been examined in bilingual children. The present study seeks to add to the body of knowledge of speech production in bilingualism by thoroughly examining both static and dynamic features of vowels in both languages. This type of research can help us better understand to what extent the L1–L2 interactions affect the detailed fine spectral features of vowel production in a young bilingual population.

1.3. Mandarin and English vowel systems

Mandarin and English belong to different language families and show systematical differences in many aspects of the phonetic systems. Mandarin is a tone language characterized by a relatively simple syllable structure in which a maximum number of four phonemes can occur. Regarding the vowel system, a commonly cited point of view states that Mandarin contains eight single vowels /i, ɿ, ʅ, y, u, a, ɤ, o/ (Lin & Wang, 2001). Among these eight vowels, ɿ and ʅ, namely the high front apical vowel and the high back apical vowel, are two allophonic variants of the high front vowel phoneme /i/, which occur after alveolar and retroflex affricates/fricatives /ts, tsʰ, s/ and /tʂ, tʂʰ, ʂ/, respectively. The vowel /o/ is phonetically realized as a diphthong /uo/. Thus, it is considered to be a diphthong rather than a monophthong (Zee, 2001). Duanmu (2007) proposed that Mandarin has five basic vowel phonemes /a, i, u, y, ə/ with each phoneme representing several variants (shown in Fig. 1). In contrast, English has a relatively large inventory of

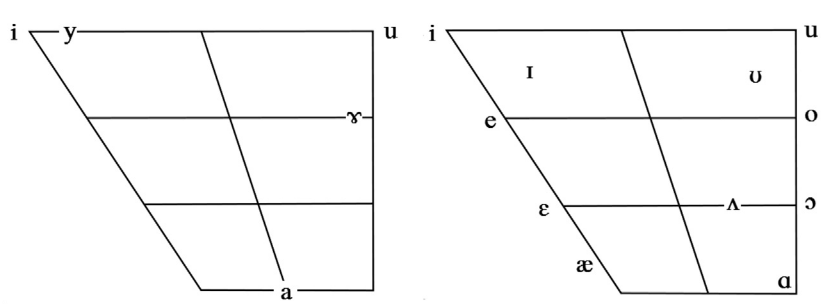

Fig. 1. Vowel charts showing the monophthongal vowel phonemes in Mandarin (left) and American English (right).

monophthongal vowel phonemes regardless of the regional variations. The bilingual and monolingual English participants in the present study were recruited from the central Ohio region in which Midland English is spoken (Jacewicz, Fox, & Salmons, 2011b; Labov, Ash, & Boberg, 2006). The vowel system of Midland English is composed of 11 monophthongal vowel phonemes /i, ɪ, e, ɛ, æ, u, ʊ, o, ʌ, ɔ, ɑ/ (shown in Fig. 1). It is noteworthy that the English vowels /e/ and /o/ are commonly realized as /eɪ/ and /oʊ/ and many speakers of Midland English demonstrated an on-going phonological process of low-back merger of /ɔ/ and /ɑ/ (Clopper, Pisoni, & De Jong, 2005; Labov et al., 2006).

Unlike the quadrilateral vowel space in English, Mandarin vowel space is shaped like a triangle with the vowels /a, i, u/ occupying the three corner positions. English has a tense-lax distinction that shows systematic differences in both temporal and spectral characteristics, whereas Mandarin lacks this feature. In addition to the monophthongs, Mandarin has a rich number of compound vowels. In particular, there are nine diphthongs /aɪ, aʊ, eɪ, uo, ua, iɛ, ia, yɛ, oʊ/ and four triphthongs /iaʊ, ioʊ, uaɪ, ueɪ/ commonly used in Mandarin. By contrast, English has only three phonemic diphthongs /aɪ, aʊ, ɔɪ/. Despite the apparent cross-language differences, some vowels such as the monophthongs /i/, /u/ and diphthongs /aɪ, aʊ, eɪ, oʊ/ are shared by both languages and demonstrate acoustic–phonetic similarities. Comparison of the shared vowels allows us to closely examine the phonetic drift of one language under the influence of the other.

1.4. Purpose and predictions of the present study

To date, only a few studies have been done on the vowel production of Mandarin-English bilingual adults (e.g. Chen, 2006; Chen, Robb, Gilbert, & Lerman, 2001; Jiang, 2008; Liao, 2006; Wang, 1997) and heritage speakers of Mandarin (Chang, Yao, Haynes, & Rhodes, 2011). The effects of L1–L2 interactions on the detailed acoustic properties of vowel production in young bilingual Mandarin-English children remain unexplored. The primary goal of the present study, then, is twofold: (1) to examine the extent to which the vowel structures and fine-grained spectral features of vowel productions in bilingual children are affected by the L1–L2 interactions; and (2) to identify the specific mechanism of phonetic drift used by the bilingual children in forming and reconstructing the vowel systems.

To answer these questions, two groups of early bilingual Mandarin-English children (5–6 years old) and two groups of monolingual correspondents were recruited. Both groups of bilingual children were sequential bilinguals (acquiring Mandarin earlier than English). One group was highly proficient in English (Bi-high) and the other group was less proficient in English (Bi-low). For the Bi-low children who were at the beginning stage of English learning, there were three possible outcomes for their English vowels. First, if bilingual children tried to separate the two vowel systems from the beginning but had not established native-like vowel categories for their L2, they might overshoot (dissimilate) the L2 vowels to realize the phonetic contrast between L1 and L2. However, based on the findings of Yang et al. (2015), emergent bilingual Mandarin-English children tend to follow the speech learning process of adult L2 learners and use the native language (Mandarin) as the basic frame to form the English vowel categories. In this case, the dissimilatory movement of L2 would be less likely to happen. Instead, we expected a second possible outcome: the Bi-low children's English vowels might be assimilated to close Mandarin vowel categories. They would not establish fully separate vowel categories for English. The acoustic features of English vowels in the Bi-low children would be different from those of the English monolingual children. Third, given the abundant evidence showing that young children can develop native-like phonetic patterns in an L2 in a relatively short period (Oh et al., 2011) and can form completely separate vowel categories as monolinguals (Lee & Iverson, 2012), it would also be possible that the effect of L1 on L2 in the Bi-low children would be limited and they might show little or no difference from the English monolinguals.

Regarding the effect of L2 on L1, previous studies revealed the effect of L2 on L1 in adult learners within a short period of immersion (Chang, 2012) or after a short-term visual auditory feedback training (Kartushina et al., 2016). Given the more malleable phonetic system in young children, we expected that even the Bi-low children would show some effect of L2 on L1. According to the rule of SLM, the dissimilatory process will not occur unless independent L2 sound categories have been established. As we hypothesized that the Bi-low children would not acquire a monolingual-like vowel system for their English, we expected that dissimilation between L1 and L2 vowels would not occur. Instead, the Bi-low children might adopt an assimilatory mechanism to shift L1 vowels toward close L2 vowels. However, if the Bi-low children successfully established a separate English vowel system, it would be possible that they might shift the L1 vowels away from the L2 vowels to maintain the phonetic contrasts between L1 and L2.

The Bi-high children, in contrast, had been deeply immersed in an English environment for a longer time. Previous studies revealed that L2 learners who started at an early age and had a long time of immersion could develop monolingual-like ability in speech production and perception (Lee & Iverson, 2012; MacLeod et al., 2009). Therefore, we expected that the Bi-high children would have developed monolingual-like ability in producing English vowel features. The effect of L1 on L2 would be attenuated or diminished. However, their long-term immersion in L2 would be expected to influence their L1. Since the Bi-high children had established independent L2 vowel categories, following the SLM rules, it would be likely that they would adopt a dissimilatory process to shift the similar L1 vowels away from the acquired L2 vowel categories to maintain the two separate phonetic systems in L1 and L2. However, many previous studies have shown both assimilation and dissimilation mechanisms in young bilingual children (Lee & Iverson, 2012; Oh et al., 2011). Therefore, we could not rule out the possibility of assimilatory phonetic drift of L1 vowels to L2 vowels in the Bi-high children.

2. Methods

2.1. Speakers

The participants included 39 children: 15 sequential Mandarin-English bilingual children aged 5–6 years old, 15 age-matched Mandarin monolingual (M-mono) children, and nine age-matched English monolingual (E-mono) children. The bilingual Mandarin-English children were divided into two groups based on their proficiency in English (see Table 1 for detailed information). The Bi-high group included eight children (six girls and two boys, Mean age = 67.1 months, SD = 3.3 months) born and raised in the U.S. (central Ohio) (average LOR in U.S. = 60.5 months). These children were raised in a near-monolingual environment and learned Mandarin as their first language from family contact and interaction with individuals in the local Mandarin-speaking community. Meanwhile, these children had a very limited amount of exposure to English before the time when they enrolled in English daycare or kindergarten at about three years of age (average AOL in English = 35.5 months). By the time of recording, these children had extensive experience with English for about three years. The Bi-low group included seven children (three girls and four boys, Mean age = 68.4 months, SD = 4.0 months) born and raised in China who had lived in the U.S. (central Ohio) for less than 6 months (average LOR in U.S. = 3.9 months). Compared to the Bi-high children, the Bi-low children started to learn English at an older age (average AOL in English = 55.4 months) and had less exposure to English, although some Bi-low children had started to learn English since the time when they enrolled in the kindergarten in China. Regarding the dialect background of bilingual children's parents, 13 out of 15 children had both parents coming from northern dialect regions of China. The other two children had one parent from the northern dialect regions. Mandarin was the daily-life language used in all bilingual children's families. The 15 age-matched M-mono children were native Mandarin speakers born and raised in the Beijing area (seven girls and eight boys, Mean age = 69.2 months, SD = 9.0 months). The nine E-mono children were native English speakers born and raised in the central Ohio region (three girls and six boys, Mean age = 66.7 months, SD = 6.3 months). All children were reported as having no speech and language disorders.

2.2. Speech materials

The recording materials for bilingual children consisted of two word lists (shown in Table 2): 10 Mandarin disyllabic words and 20 English monosyllabic or disyllabic words that covered the monophthongal vowel phonemes in each language. Following Duanmu's (2007) viewpoint, the five basic Mandarin vowel phonemes /a, i, u, y, ɤ/ were selected as the L1 target vowels. Note that the present study used the symbol /ɤ/ rather than /ə/ to represent the standard usage of the mid-high unrounded back vowel. As for the English vowel set, due to the merger of /ɔ/ and /ɑ/ in most dialects of American English, vowel /ɔ/ was not selected for use in the present study. Therefore, the ten American English monophthongal vowel phonemes /i, ɪ, e, ɛ, æ, u, ʊ, o, ɑ, ʌ/ were selected as English target vowels. In all disyllabic words, the target vowel appeared in the first syllable. The selection of both Mandarin and English words was based on word familiarity, word frequency, and picturability. The consonants preceding the English vowels were all stops except for /f/ in 'feet' whereas the consonants following the English vowels were all voiceless stops or fricative /s/. As for Mandarin, all tested words started

Table 1
Characteristics of subgroups of bilingual Mandarin-English children: subject group, subject number, gender, age, length of residency in the U.S., age of learning in English, percentage of daily usage of English, age of learning in Mandarin, percentage of daily usage of Mandarin.

Group	SubNum	Gender	Age	LOR_in_U.S. (in months)	AOL_E	E_usage	AOL_M	M_usage
Bi-high	1	F	5;11	58	3;6	50%	Since birth	50%
	2	F	5;7	67	4	50%	Since birth	50%
	3	M	5;5	65	3;7	80%	Since birth	20%
	4	F	5;10	70	2;1	70%	Since birth	30%
	5	F	5;10	47	3	70%	Since birth	30%
	6	F	5;6	66	2;6	70%	Since birth	30%
	7	M	5;1	45	2	70%	Since birth	30%
	8	F	5;7	66	3	70%	Since birth	30%
Bi-low	1	F	5;10	5	4;6 (in China), 5;5 (in U.S.)	40%	Since birth	60%
	2	M	5;8	3	4 (in China); 5;5 (in U.S.)	40%	Since birth	60%
	3	M	5;8	3	5;5 (in U.S.)	30%	Since birth	70%
	4	M	5;2	1.5	4 (in China); 5 (in U.S.)	30%	Since birth	70%
	5	F	5;6	6	3 (in China), 5 (in U.S.)	20%	Since birth	80%
	6	F	6;3	5	5;11 (in U.S.)	20%	Since birth	80%
	7	F	5;10	4	5;6 (in U.S.)	40%	Since birth	60%

Table 2
Word lists used to collect Mandarin (left) and English (right) speech samples.

Vowel	Pinyin	IPA	Gloss	Vowel	Word	IPA	Word	IPA
a	d**à** xiàng	/t**a** ɕiaŋ/	elephant	i	feet	/f**i**t/	geese	/g**i**s/
	d**à** suàn	/t**a** suan/	garlic	ɪ	kiss	/k**ɪ**s/	kitchen	/k**ɪ**tʃɪn/
i	p**í** qiú	/pʰ**i** tɕʰioʊ/	ball	e	cake	/k**e**k/	tape	/t**e**p/
	b**í** zi	/p**i** tsɿ/	nose	ɛ	desk	/d**ɛ**sk/	pet	/p**ɛ**t/
u	t**ù** zi	/tʰ**u** tsɿ/	rabbit	æ	cat	/k**æ**t/	bat	/b**æ**t/
	p**ú** tao	/pʰ**u** tʰɑʊ/	grape	u	boot	/b**u**t/	goose	/g**u**s/
y	j**ú** zi	/tɕ**y** tsɿ/	orange	ʊ	book	/b**ʊ**k/	cookie	/k**ʊ**kɪ/
	y**ú** tóu	/**y** tʰoʊ/	fish head	o	coat	/k**o**t/	boat	/b**o**t/
ɤ	g**ē** ge	/k**ɤ** kɤ/	older brother	ɑ	box	/b**ɑ**ks/	stop	/st**ɑ**p/
	g**ē** zi	/k**ɤ** tsɿ/	pigeon	ʌ	cup	/k**ʌ**p/	duck	/d**ʌ**k/

with stop consonants except for the words containing /y/ which had either an initial zero (no preceding consonant) or the affricate /tɕ/ due to phonotactic constraints in Mandarin syllable structure. In addition, the third tone was avoided for Mandarin words to reduce the tone-vowel interaction that is mainly represented in tone 3 (Chang, 2010; Hoole & Hu, 2004).

2.3. Procedures

The data collection for all bilingual children and English monolingual children was completed in central Ohio, while that for Mandarin monolingual children was completed in Beijing, China. Prior to the recording, a questionnaire covering the demographic information, information about participants' language learning and language usage was completed by their parents.

There were two recording sessions for each bilingual subject. Mandarin words were produced in the first session and English words were produced in the second session after a 15–20 min break. The experimenter (a fluent bilingual Mandarin-English speaker) interacted with the speakers in Mandarin during the Mandarin sessions and English during the English sessions. For each monolingual speaker, only one recording session in their native language was conducted. The recording session for each language consisted of two recording blocks. In each block, a randomly ordered list of test words was produced one time by each participant. The same random order was used in both blocks for each language.

A visual-auditory word repetition task was used to collect speech samples under control of a custom MATLAB program. During the recording period, each speaker was seated facing a laptop computer in a quiet room. Pictures containing target words were presented on the computer screen followed by audio prompts produced by a native adult speaker. In particular, the Mandarin audio prompts were produced by a young female speaker of Mandarin and the English audio prompts were produced by a young female native speaker of Ohio English. The participants were then asked to repeat each word immediately after the prompt. The present study adopted a visual-auditory word repetition task because it ensured a better control of stimulus presentation and the target words were elicited in the designed manner (Edwards & Beckman, 2008). During the recording sessions, a Shure SM10A head-mounted microphone was situated approximately 1-inch from the subject's mouth. Speech samples were recorded directly onto a hard drive disk with a 16-bit quantization rate and 44.1 kHz sampling rate.

2.4. Acoustic measurements

The landmark locations of vowel onset and offset were marked by hand and determined primarily on the basis of the waveform, accompanying with a visual check of the spectrogram (Kent & Read, 1992). Vowel onset was defined at the point of zero crossing of the first glottal pulse following stop release or cessation of frication. For the Mandarin word 'yú' (fish) that has no word-initial consonant, the vowel onset was defined as the start of the first clear glottal pulse. Vowel offset of Mandarin vowels was set at the point of the zero crossing point of the last glottal pulse extending through F1 and F2. The offset of English vowels was set at the zero crossing point close to the significant decrease in amplitude due to the stop closure or was marked at the point prior to the onset of frication. A reliability check was implemented on all landmark locations and extraction of formant trajectories. The uncertain measurements, approximately 2% of the total data, were double checked. Any discrepancies were resolved between the two authors following discussion.

2.4.1. Formant frequencies

All tokens were first down-sampled to 11.025 kHz. A spectrographic analysis program TF32 (Milenkovic, 2003) was then used to extract the frequency values of the first two formants, F1 and F2, at five equidistant temporal locations (20–35–50–65–80% point) in order to capture the dynamic spectral change of the vowel movement (Fox & Jacewicz, 2009; Jacewicz, Fox, & Salmons, 2011c). According to previous studies on the development change of vocal tract anatomy, the sexual dimorphism of vocal tract structures does not occur before 7 years of age (Vorperian et al., 2005). The effect of vocal tract length on the vowel formant frequency values between sexes during the pre- and peri-puberty period is small (Lee & Iverson, 2009). In the present study, the speakers across all four groups were in a narrow age range younger than 7. Therefore, the effect of different vocal tract lengths on the formant frequencies was expected to be small (Vorperian & Kent, 2007). Thus, the original formant frequency values (not normalized values) were used for further calculation and analysis.

2.4.2. Trajectory length

Trajectory length (TL) was derived on the basis of formant frequency values. TL is the sum of the Euclidean distances (in Hz) of the four separate vowel sections between the 20% and 80% points (Fox & Jacewicz, 2009),

$$TL = \sum_{n=1}^{4} VSL_n \tag{1}$$

where the length of each vowel section (VSL) is calculated based on the formula:

$$VSL_n = \sqrt{(F1_{n+1} - F1_n)^2 + (F2_{n+1} - F2_n)^2} \tag{2}$$

A longer TL reflects a greater change in formant frequency values and a greater magnitude of formant movement. This measure provides a detailed evaluation of the extent to which the formant trajectory changes over the course of vowel duration between the selected time locations. This measure is particularly helpful in capturing the formant movement of the curved formant trajectories.

2.5. Statistical analysis

The mid-point formant frequency values and vowel trajectory length (TL) were subject to ANOVA tests. The results on the Mandarin vowel system and English vowel system are presented in Section 3.1 and 3.2. A two-way repeated measures ANOVA was conducted for each acoustic measure among the two bilingual groups and monolingual group, with vowel quality as the within-subject factor and language experience as the between-subject factor. Because the change of formant values and TLs as a function of vowel quality was expected

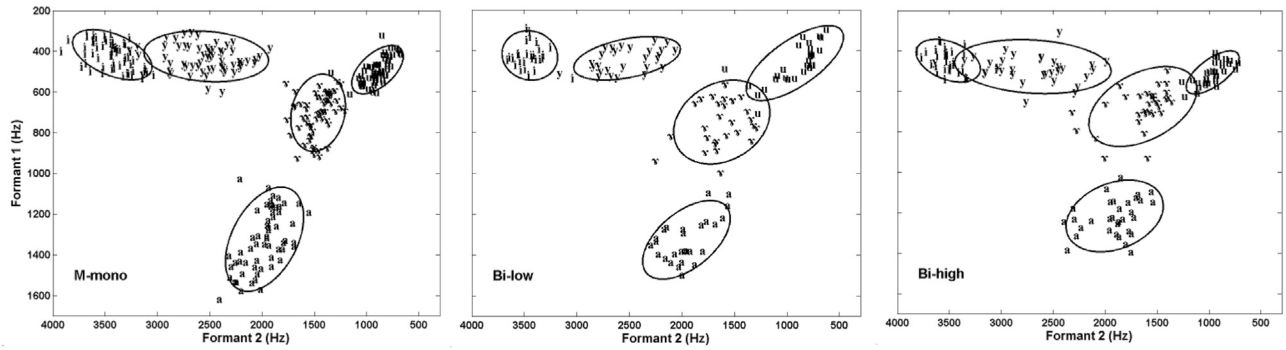

Fig. 2. Scatter plots of midpoint formant frequency values for individual Mandarin vowel categories in each group of children. Each ellipse encircles approximately 95% of the samples in each Mandarin vowel category.

and has been reported in previous studies, it was not of interest in the present study. When the main effect of language experience and/or interaction effect yielded a significant result, one-way ANOVAs with Bonferroni post hoc (adjusted for multiple comparisons among the three groups) tests were conducted to examine the differences of the acoustic measures among the three groups of children for each vowel. The results on the shared Mandarin and English vowels are presented in Section 3.3. A one-way ANOVA or one-way repeated measures ANOVA was directly used to examine how similar or different each Mandarin-English shared vowel pair was within each group of children.

3. Results

3.1. Mandarin (L1) vowel system

3.1.1. Midpoint F1 by F2 vowel space

Following the traditional approach of using midpoint formant frequency values to characterize the basic acoustic features of vowel production, the formant values at the midpoint locations for all vowels were measured and plotted in the F1 × F2 vowel space. To better describe and quantify the distribution of each vowel category, vowel ellipses were plotted for individual vowel categories (shown in Fig. 2) following the method used in Zhou and Xu (2008). Each ellipse encircled approximately 95% of the samples in each Mandarin vowel category. The ellipse area of each vowel category was calculated (shown in Table 3). Both groups of bilingual children clearly separated the five Mandarin vowels in the acoustic vowel space. However, the Bi-high children showed greater variation in their production of /y/ and /ɤ/ and some overlap between /i/ and /y/ in comparison to the Bi-low and M-mono children. Also, the Bi-high children produced /a/ at a higher position than the other two groups. Note that the size of the ellipse for the vowel /i/ in the Bi-high children was smaller than those in the other two groups, which indicated less variation of this vowel in the Bi-high children.

Table 3
The area of ellipses of each Mandarin vowel in each group of children (in kHz^2).

	a	i	u	y	ɤ
Bi-low	0.211	0.101	0.193	0.155	0.287
Bi-high	0.250	0.095	0.063	0.312	0.292
M-mono	0.272	0.128	0.076	0.231	0.155

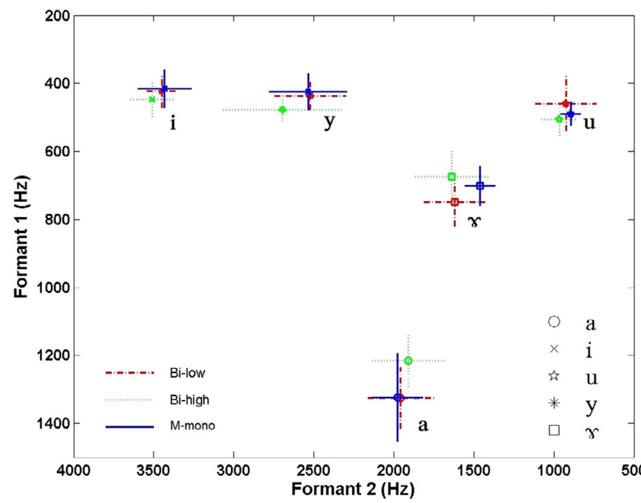

Fig. 3. Dispersion of Mandarin vowels for each group of children plotted on the basis of the means and standard deviations of midpoint formant frequency values.

Fig. 3 shows the means and standard deviations of midpoint formant frequency values for the five Mandarin vowels in each group of children. Although the two-way repeated measures ANOVAs revealed no significant main effect of language experience on F1 or F2, when both compared to the M-mono children, the Bi-low and Bi-high children still showed apparent differences in the relative positions of individual vowels. To better quantify the difference between the two groups of bilingual children, the quadratic distance (also called Mahalanobis distance, McLachlan, 1999) in the F1 × F2 vowel space was used to measure the separation of vowel categories between the Bi-low and M-mono children as well as the Bi-high and M-mono children for each vowel. As shown in Table 4, all vowels except for /u/ showed greater separation between the Bi-high and M-mono children than between the Bi-low and M-mono children.

Table 4
Quadratic (Mahalanobis) distance showing the separation between each vowel category in Mandarin monolinguals and corresponding vowel category in each group of bilingual children.

	a	i	u	y	ɤ
Bi-low to M-mono	0.1429	0.1857	0.7375	0.2081	0.9702
Bi-high to M-mono	0.8441	0.9335	0.6018	1.0456	1.1841

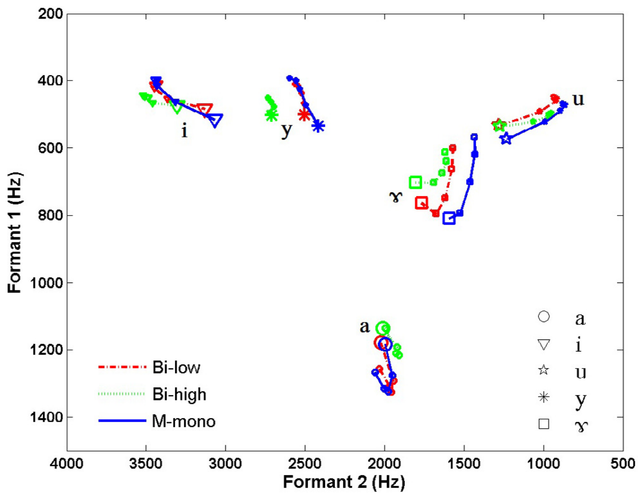

Fig. 4. Vowel spectral change plotted on the basis of formant frequencies at five time locations (the 20–35–50–65–80% points) over the course of vowel duration for five Mandarin vowels in each group of children. The larger size symbols represent the 80% point.

Table 5
Summary of one-way ANOVA results showing the effect of L2 experience on trajectory lengths (TLs) of each Mandarin vowel (means and standard errors are in Hz).

	Bi-low M (SE)	Bi-high M (SE)	M-mono M (SE)	*p*	η_p^2	Bonferroni post hoc significant difference
a	398 (13)	342 (19)	413 (31)	0.236	0.101	
i	447 (28)	330 (45)	514 (27)	0.002	0.375	Bi-high < M-mono
u	510 (40)	422 (46)	493 (26)	0.242	0.1	
y	372 (36)	291 (26)	411 (25)	0.020	0.251	Bi-high < M-mono
ɤ	449 (28)	351 (21)	543 (22)	0.015	0.267	Bi-high < M-mono

3.1.2. Spectral change in Mandarin vowels

While the vowel dispersion pattern of the midpoint formant frequency values represents a traditional approach to describe the general distribution of relative positions in the acoustic vowel space, it does not allow us to capture the nature of vowel quality change over time. As mentioned earlier, measurements made at multiple time locations (the 20%, 35%, 50%, 65%, and 80% points) provide an estimate of the extent of formant movement over the central 60% portion of vowel duration. Based on the multiple measurements, the direction and the magnitude of the formant movement can be described. As shown in Fig. 4, both Bi-low and Bi-high children were similar to the M-mono children in the directions of formant movement for most of the Mandarin vowels. However, the Bi-high children showed different shapes and/or magnitudes of formant movement than the M-mono and Bi-low children for the vowels /i/, /y/, /a/ and /ɤ/. In particular, the Bi-high children showed forward movement with less formant movement for the vowels /i/, /y/ and /ɤ/. For the vowel /a/, the formant trajectory of the Bi-high children was located at a higher position with lower F1 than those in the other two groups.

Table 5 shows the means and standard errors of the TLs for each Mandarin vowel in all three groups of children. The Bi-low children produced TLs similar to those of the M-mono children for all five vowels. The Bi-high children, contrarily, produced shorter TLs than the Bi-low and M-mono children. The two-way repeated measures ANOVA yielded a significant main effect of vowel ($F(4,108) = 7.078$, $p < 0.0001$, $\eta_p^2 = 0.208$) and language experience ($F(2,27) = 10.430$, $p < 0.0001$, $\eta_p^2 = 0.436$) but no significant interaction effect. As the main effect of language experience was of particular interest in the present study, one-way ANOVAs and Bonferroni post hoc tests were implemented, which returned no significant differences in TLs between the Bi-low and M-mono children for any vowel. However, for the vowels /i, y, ɤ/, the Bi-high children showed significantly shorter TLs than the M-mono children. These results indicated that language experience in English produced a significant change in the vowel dynamics of the Bi-high children in their native vowels. However, the short period of immersion in English did not affect the Bi-low children's vowel dynamic features in their Mandarin vowels.

3.2. English (L2) vowel system

3.2.1. Midpoint F1 by F2 vowel space

Fig. 5 shows the scatter plots of midpoint F1 and F2 values for the English vowels in each group of children. Following the same method used for Mandarin vowels, an ellipse was plotted for each English vowel category. The E-mono and Bi-high children showed some overlap among the English vowels. However, the Bi-low children showed a greater degree of overlap than the other two groups in the F1 by F2 vowel space. In general, the production of the ten English vowels in the Bi-low children can be categorized into four clusters in the acoustic vowel space: /i, ɪ, e/, /ɛ, æ/, /u, ʊ, o/ and /ɑ, ʌ/. These highly overlapped vowel clusters indicated that the Bi-low children did not clearly separate these English vowels in the acoustic

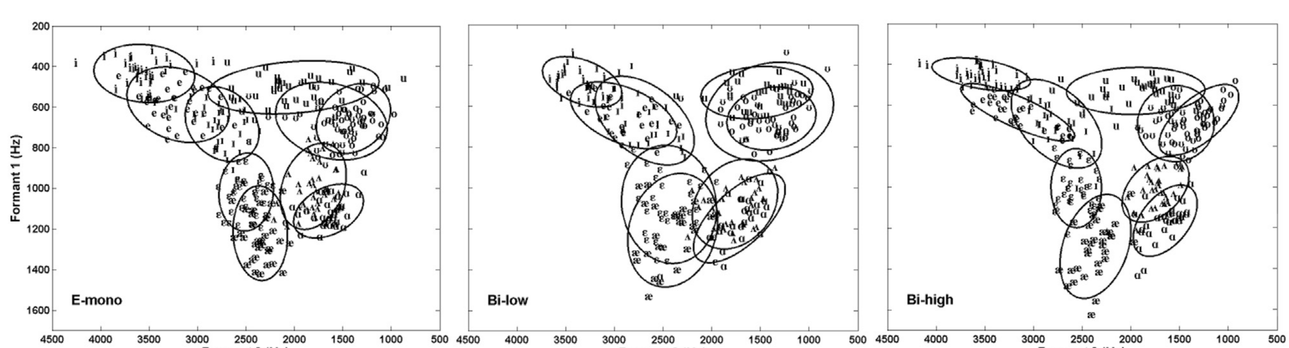

Fig. 5. Scatter plots of midpoint formant frequency values for individual English vowel categories in each group of children. Each ellipse encircles approximately 95% of the samples in each English vowel category.

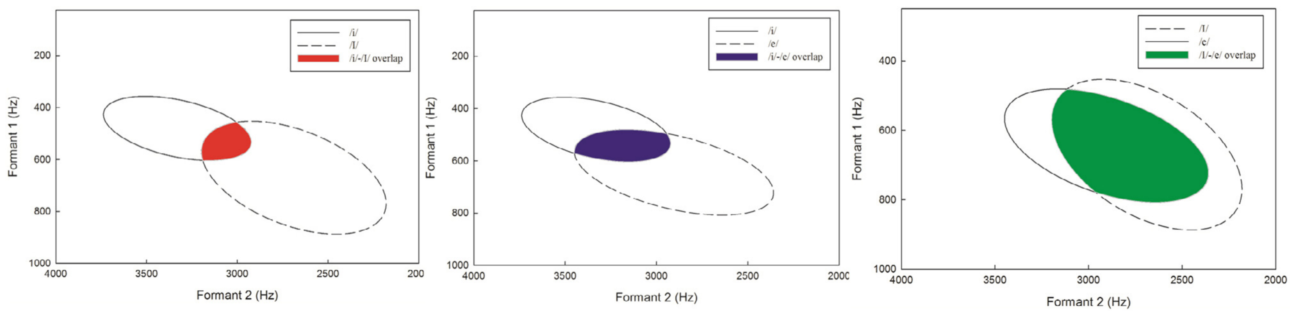

Fig. 6. Illustration of overlap area between /i/ and /ɪ/ (left), /i/ and /e/ (middle), and /ɪ/ and /e/ (right) in the Bi-low children.

Table 6
The area of the ellipse for each English vowel in each group of children (in kHz^2).

	i	ɪ	e	ɛ	æ	u	ʊ	o	ɑ	ʌ
Bi-low	0.143	0.309	0.246	0.443	0.399	0.230	0.492	0.220	0.245	0.294
Bi-high	0.114	0.280	0.150	0.158	0.276	0.258	0.231	0.169	0.157	0.158
E-mono	0.234	0.214	0.299	0.169	0.206	0.355	0.340	0.202	0.139	0.219

space. As shown in the right panel, the Bi-high children showed better separation among the vowel categories, similar to that of the E-mono children. However, the Bi-high children still demonstrated differences from the E-mono children. In particular, the /ɪ/ vowels produced by the Bi-high children were more scattered than those produced by the E-mono children. Moreover, the vowels /u/ and /ʊ/ produced by the Bi-high children showed less dispersion than those produced by the E-mono children.

To examine the extent to which the vowel categories were separated, for each group, the proportion of the overlap of selected intersecting vowel pairs (i-ɪ, i-e, ɪ-e, u-ʊ, u-o, ʊ-o, ɛ-æ, ɑ-ʌ) was calculated (e.g. shown in Fig. 6). First, the ellipse area of each vowel category was calculated (shown in Table 6). Then, the overlap area of each selected vowel pairs was calculated (shown in Table 7), similar to the method of SOAM 2-D in Wassink (2006). Finally, the proportion of the overlap to each of the two involved vowel categories was calculated (shown in Table 8). The overlap areas for most vowel pairs in the Bi-low group were substantially greater than those in the other two groups of children. Moreover, the proportions of overlap to each vowel category in the Bi-low children were greater than those in the other two groups. These results indicated that the Bi-low children were less likely to separate the vowel categories in the acoustic space.

Table 7
The area of intersection between two ellipses for eight selected English vowel pairs in each group of children (in kHz^2).

	i-ɪ	i-e	ɪ-e	u-ʊ	u-o	ʊ-o	ɛ-æ	ɑ-ʌ
Bi-low	0.030	0.052	0.190	0.206	0.089	0.220	0.299	0.194
Bi-high	0.004	0.006	0.082	0.057	0.006	0.114	0.024	0.044
E-mono	0.000	0.095	0.095	0.096	0.035	0.171	0.065	0.071

Fig. 7 shows the means and standard deviations of midpoint formant frequency values for the English vowels in each group of children. The two-way repeated measures ANOVAs revealed no significant main effect of language experience on F1 or F2 values. However, evident positional variation can be observed in the Bi-low children relative to the E-mono children. The Bi-high children also displayed slight positional changes for certain vowels when compared to the E-mono children. The calculation of quadratic distance (shown in Table 9) revealed that the distance between the Bi-low children and the E-mono children was greater than the distance between the Bi-high children and the E-mono children for most English vowels except for /o/, /ɑ/ and /æ/. This finding indicated that the Bi-low children's production of English vowels showed more acoustic deviations from the E-mono children than did the Bi-high children's productions.

3.2.2. Spectral change in English vowels

Fig. 8 shows the formant trajectories of the ten English vowels in each group of children. Among the three groups of children, the Bi-high children showed similar formant movement patterns to the E-mono children for the majority of the English

Table 8
The proportion of the overlapping area to the overall area of each vowel category for the eight selected vowel pairs in each group of children (e.g. for the vowel pair i-ɪ, O/i represents the proportion of the area of i-ɪ overlap to the area of i vowel ellipse). This measurement helps us quantify the separation of vowel categories in the acoustic space for each group of children.

	i-ɪ		i-e		ɪ-e		u-ʊ		u-o		ʊ-o		ɛ-æ		ɑ-ʌ	
	O/i	O/ɪ	O/i	O/e	O/ɪ	O/e	O/u	O/ʊ	O/u	O/o	O/ʊ	O/o	O/ɛ	O/æ	O/ɑ	O/ʌ
Bi-low	21%	10%	37%	21%	61%	77%	89%	42%	39%	40%	45%	100%	68%	75%	79%	66%
Bi-high	4%	1%	5%	4%	29%	55%	22%	25%	2%	4%	49%	68%	15%	9%	28%	28%
E-mono	0	0	41%	32%	44%	32%	27%	28%	10%	18%	51%	85%	38%	31%	51%	32%

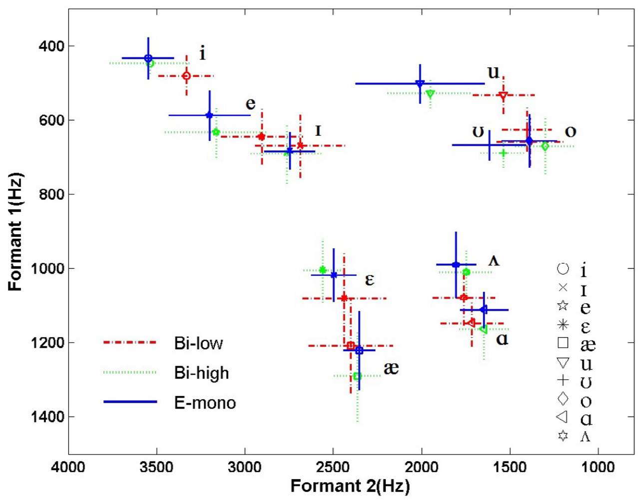

Fig. 7. Dispersion of English vowels for each group of children plotted on the basis of the means and standard deviations of midpoint formant frequency values.

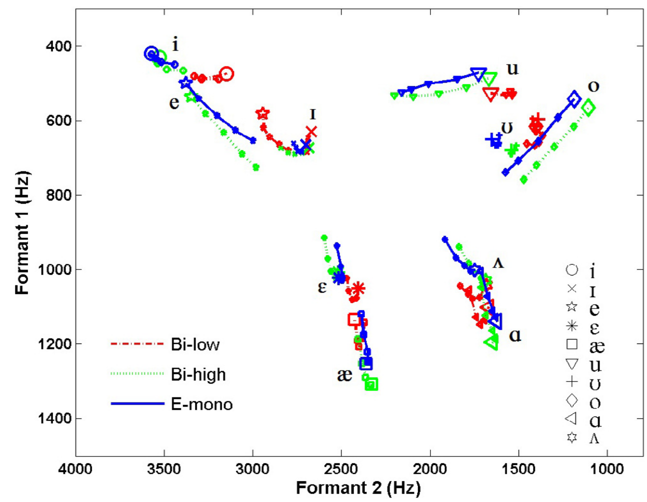

Fig. 8. Vowel spectral change plotted on the basis of formant frequencies at five time locations (the 20–35–50–65–80% points) over the course of vowel duration for 10 English vowels in each group of children. The larger size symbols represent the 80% point.

vowels. However, the Bi-low children differed from the Bi-high and E-mono children for all English vowels in terms of the relative position, direction and/or magnitude of the formant trajectories. To better compare the formant trajectories across the three groups of children, the ten English vowels were divided into four classes each of which is shown in Fig. 9.

For the high front vowel group, /i, ɪ, e/, the formant trajectories of /i/ and /e/ were located in a lower and further back position in the Bi-low children than in the Bi-high or E-mono children. In addition, the vowel /e/ showed much less formant movement in the Bi-low children than in the other two groups. Regarding the vowel /ɪ/, all three groups of children showed a different pattern of formant movement. The E-mono children showed little change in F2 but an increase followed by a decrease in F1. The Bi-high children showed a decrease in F2 but little change in F1. The Bi-low children showed no change in F2 but a decrease in F1.

For the low front vowel group /ɛ/ and /æ/, the formant trajectories of these two vowels were much closer to each other in the Bi-low children than in the Bi-high and E-mono children. This pattern indicated that the Bi-low children did not separate these two vowel categories to the same extent as the Bi-high or E-mono children. The vowel /æ/ in the Bi-high and E-mono children showed an increase in F1 with a slight decrease in F2. However, the formant trajectory in the Bi-low children showed an increase in F1 followed by a decrease. Thus, the direction of formant trajectory in the Bi-low children was opposite to those in the Bi-high and E-mono children.

For the high back vowel group /u, ʊ, o/, both Bi-high and E-mono children produced the vowels /u/ and /o/ with substantial formant movement. However, the Bi-low children produced these two vowels with little formant movement. In addition, the formant trajectory of /u/ in the Bi-low children was located further back and moved in an opposite direction from that in the Bi-high or E-mono children. For the vowel /ʊ/, all three groups of children showed little formant change. However, this vowel produced by the Bi-low children was located at a further back and higher position than the other two groups.

For the low back vowel group /ʌ/ and /ɑ/, the formant trajectories of these two vowels were much closer to each other in the Bi-low children than in the other groups. In addition, the formant trajectory of /ʌ/ in the Bi-low children was located at a lower position with less change in F1 in comparison to the Bi-high and E-mono children. The formant trajectory of /ɑ/ showed an increase in F1 and a decrease in F2 in the Bi-high and E-mono children. However, in the Bi-low children, /ɑ/ was fronted and the formant trajectory moved in a different direction from the other two groups.

Table 10 shows the means and standard errors of TLs for English vowels in each group of children. Compared to the Bi-low children, the Bi-high children were more similar to the E-mono children on the TLs for most English vowels. However, the Bi-high children still showed differences from the E-mono children in that the Bi-high children produced longer TLs than the E-mono children for the vowels /i, ɪ, æ, u, ɑ/. The two-way repeated measures ANOVA yielded a significant main effect of vowel ($F(9,189) = 18.147$, $p < 0.0001$, $\eta_p^2 = 0.464$) and a significant interaction effect of vowel by language experience ($F(18,189) = 5.432$, $p < 0.0001$, $\eta_p^2 = 0.341$). One-way ANOVAs were then conducted (results summarized in Table 6) to examine the difference of TL among three groups of children for individual vowels. The results showed a complex pattern of group differences. In particular, the Bi-low children produced significantly longer TLs for the vowels /i/, /ɪ/, /ɛ/, /æ/, /ɑ/ and shorter TLs for the vowel /o/ than the E-mono children. These

Table 9
Quadratic (Mahalanobis) distance showing the separation between each English vowel category in English monolinguals and corresponding vowel category in each group of bilingual children.

	i	ɪ	e	ɛ	æ	u	ʊ	o	ɑ	ʌ
Bi-low to E-mono	1.0593	0.3688	1.1228	0.583	0.2905	1.4408	0.8222	0.0464	0.4379	0.9354
Bi-high to E-mono	0.2322	0.1426	0.5771	0.4614	0.5442	0.4877	0.3748	0.6313	0.7627	0.5127

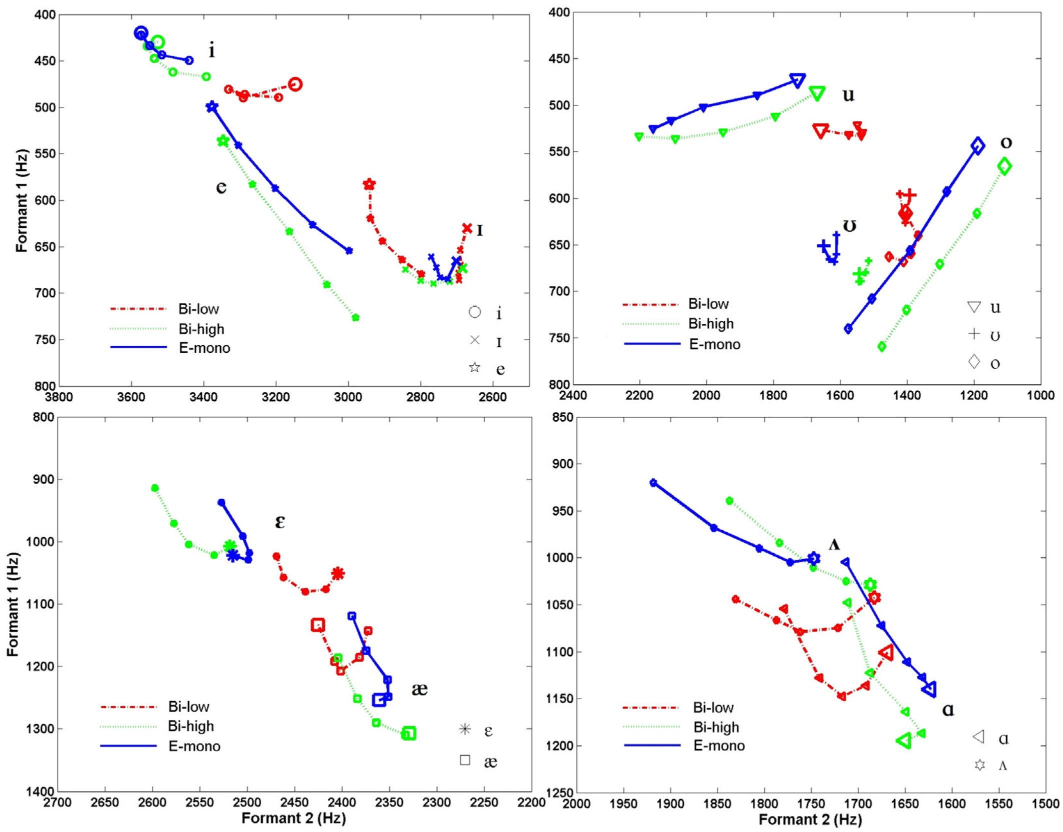

Fig. 9. Illustration of English vowel spectral change for the four subsets of vowels in each group of children. Note that the scale of each panel is different. The larger size symbol represents the 80% point.

Table 10
Summary of statistical results showing the effect of L2 experience on trajectory length (TLs) of each English vowel (means and standard errors are in Hz).

	Bi-low M (SE)	Bi-high M (SE)	E-mono M (SE)	p	η_p^2	Bonferroni post hoc significant difference
i	499 (52)	300 (33)	254 (20)	<0.001	0.56	Bi-high < Bi-low, E-mono < Bi-low
ɪ	268 (23)	253 (16)	181 (9)	0.002	0.456	E-mono < Bi-low, E-mono < Bi-high
e	345 (24)	459 (48)	457 (51)	0.169	0.156	
ɛ	319 (33)	250 (17)	240 (14)	0.035	0.274	E-mono < Bi-low
æ	352 (31)	274 (25)	247 (19)	0.022	0.306	E-mono < Bi-low
u	325 (37)	595 (72)	517 (108)	0.106	0.193	
ʊ	246 (32)	203 (18)	196 (15)	0.237	0.128	
o	244 (18)	458 (36)	464 (32)	<0.001	0.6	Bi-low < Bi-high, Bi-low < E-mono
ɑ	318 (17)	295 (21)	240 (10)	0.007	0.379	E-mono < Bi-low
ʌ	282 (16)	251 (25)	264 (16)	0.563	0.053	

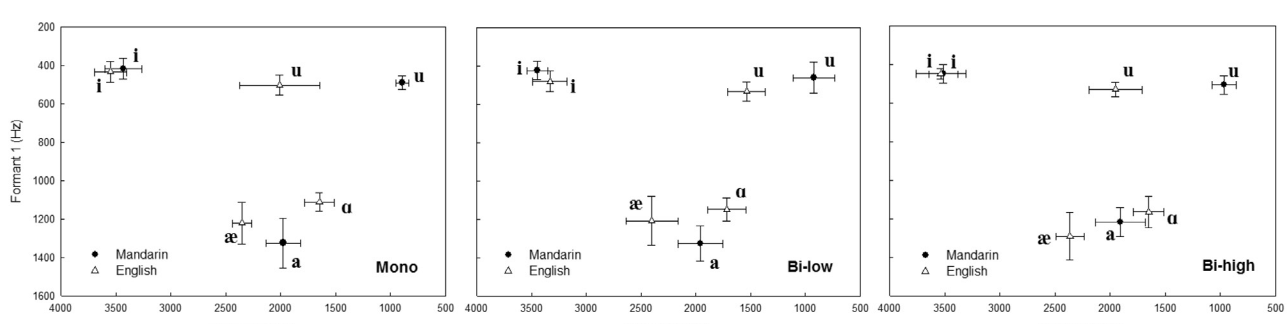

Fig. 10. Means and standard deviations of midpoint formant frequency values for Mandarin and English shared vowels in each group of children.

differences between the Bi-low and E-mono children demonstrated that the young children learning English over a short period (less than six months in the present study) had not developed native-like patterns of vowel spectral change as monolingual English children. On the other hand, while the Bi-high children showed limited differences from the E-mono children on most English vowels, they produced significantly longer TLs for the vowel /ɪ/ than the E-mono children. This finding suggested that even though bilingual children were highly proficient in English and had used English for a long time, they still differed from monolingual peers in certain aspects of acoustic properties for certain vowels.

3.3. Shared vowels

Included among the five Mandarin vowels and ten English vowels are several vowel pairs that have similar phonetic quality. In order to closely examine the extent to which the L2 experience affects bilingual children's production of these shared vowels, Mandarin /a, i, u/ and English /ɑ, i, u, æ/ were selected for further analysis. There was no single English vowel matched with Mandarin /a/. Preliminary analysis showed that the vowels /ɑ/ and /æ/, rather than /ʌ/, were more phonetically similar to Mandarin /a/; therefore, the vowels /ɑ, æ/ were chosen as comparison tokens. Note that because each monolingual speaker just produced the words in their native language, the shared vowels of the Mono group in this section were separately selected from monolingual Mandarin and monolingual English children.

3.3.1. Midpoint vowel dispersion of shared vowels

The mean values of midpoint formant frequencies for the selected shared vowels in each group were plotted in Fig. 10. As shown in the left panel for the Mono children, Mandarin /i/ and English /i/ were located close to each other while Mandarin /u/ and English /u/ showed a greater separation in the acoustic vowel space. Mandarin /a/ was located in a more central and lower position relative to the two English counterparts. One-way ANOVAs were conducted to compare each formant (F1 and F2) at the midpoint for each vowel pair between the English monolingual and Mandarin monolingual children. The results revealed no significant difference between Mandarin and English /i/ for either F1 or F2. However, English /u/ showed a significantly higher F2 than Mandarin /u/ ($F(1,22) = 138.445$, $p < 0.0001$, $\eta_p^2 = 0.863$). For Mandarin /a/ and English /ɑ/ and /æ/, the results showed significant differences among these three vowels for both F1 ($F(2,30) = 11.208$, $p < 0.0001$, $\eta_p^2 = 0.428$) and F2 ($F(2,30) = 60.652$, $p < 0.0001$, $\eta_p^2 = 0.802$). In particular, Mandarin /a/ showed significantly higher F1 than English /ɑ/ and all three vowels were significantly different from each other on the F2.

One-way repeated measures ANOVAs were used to examine F1 and F2, respectively, for each vowel pair in the two groups of bilingual children. As shown in the middle panel for the Bi-low children, unlike the Mono children, Mandarin /i/ was located in a further high front position than English /i/. This observation was confirmed by a significantly lower F1 ($F(1,6) = 28.162$, $p = 0.002$, $\eta_p^2 = 0.824$) and a significantly higher F2 ($F(1,6) = 6.526$, $p = 0.043$, $\eta_p^2 = 0.521$) in Mandarin /i/ than in English /i/. For the vowel pair of Mandarin /u/ vs. English /u/, similar to the Mono children, English /u/ was located in a more fronted position than Mandarin /u/ (F2: ($F(1,6) = 42.768$, $p = 0.001$, $\eta_p^2 = 0.877$). However, English /u/ was located at a back position and was closer to Mandarin /u/ compared to that in Mono children. The Euclidean distance between the two /u/ vowels in the Bi-low children was only 616 Hz while that in the Mono children was 1115 Hz. This difference evidenced the approximation of English /u/ toward Mandarin /u/ in the Bi-low children. For Mandarin /a/ and English /ɑ/ and /æ/, the Bi-low children showed similar organization to Mono children. The repeated measures ANOVAs yielded significant differences among the three vowels for both F1 ($F(2,12) = 12.746$, $p = 0.001$, $\eta_p^2 = 0.680$) and F2 ($F(2,12) = 60.429$, $p < 0.001$, $\eta_p^2 = 0.910$).

As shown in the right panel for the Bi-high children, Mandarin /i/ and English /i/ overlapped completely. No significant difference was yielded for either F1 or F2. For the pair of Mandarin /u/ and English /u/, the Bi-high children showed a large separation between these two vowels in the acoustic space (F2: ($F(1,7) = 125.161$, $p < 0.0001$, $\eta_p^2 = 0.947$), similar to the Mono children. As for the vowels of Mandarin /a/ and English /æ/ and /ɑ/, the statistical results showed significant differences on both F1 ($F(2,14) = 7.067$, $p = 0.008$, $\eta_p^2 = 0.502$) and F2 ($F(2,14) = 84.062$, $p < 0.0001$, $\eta_p^2 = 0.923$). However, different from the vowel dispersion pattern in Mono or Bi-low children, Mandarin /a/ moved up and showed a greater approximation to the two English counterparts along the F1 axis. The post hoc analysis revealed no significant difference between Mandarin /a/ and English /æ/ or /ɑ/ in F1 values.

3.3.2. Spectral change in shared vowels

Fig. 11 shows the formant trajectories of the shared vowels produced by monolingual and bilingual children. In the panel

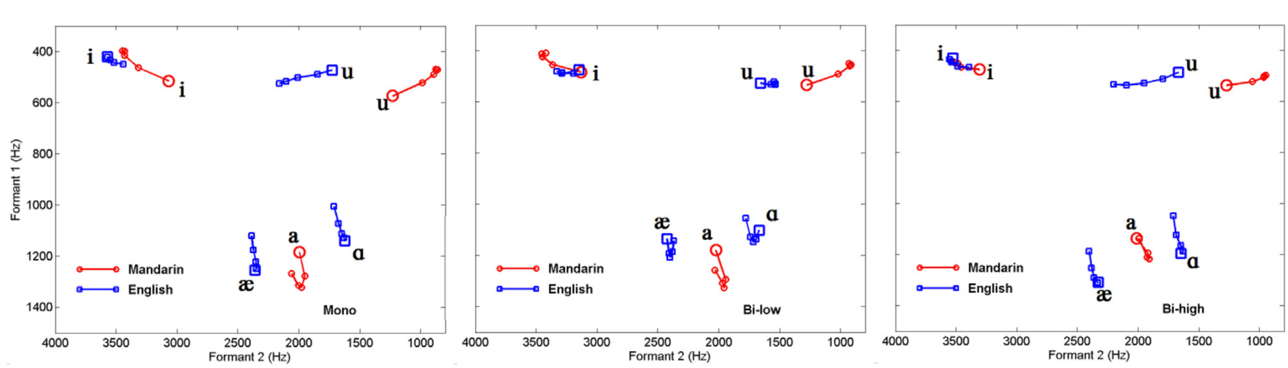

Fig. 11. Vowel spectral change of Mandarin and English shared vowels for each group of children. The larger size symbols represent the 80% point.

for the Mono children, the formant trajectories of all three Mandarin vowels moved in an opposite direction from their English counterpart vowels. In addition, the magnitudes of formant movement of Mandarin /i/ and /a/ were larger than those of English counterparts. These differences in vowel spectral change cannot be seen when looking only at the midpoint formant frequency values. The one-way ANOVA on the TLs of each vowel pair between the M-mono and E-mono children revealed a significant longer TLs in Mandarin /a/ than in English /æ/ and /ɑ/ ($F(2,32) = 14.672$, $p < 0.001$, $\eta_p^2 = 0.494$) as well as significantly longer TLs in Mandarin /i/ than in English /i/ ($F(1,22) = 46.557$, $p < 0.001$, $\eta_p^2 = 0.679$). The significantly different TLs between Mandarin /i/ and English /i/ suggested that even though Mandarin /i/ and English /i/ showed an approximation of midpoint formant values in the acoustic space, they were still different in the magnitude of spectral change.

The Bi-low children maintained the formant movement patterns for their Mandarin vowels. As for their English vowels, the formant trajectories moved in a different direction from those in the E-mono children, but in the same direction as those in the M-mono children. The consistent direction of formant trajectories between Mandarin and English shared vowels in the Bi-low children suggested that these children were transferring some features of their L1 vowels to the L2 vowels. For the Bi-high children, the direction of formant trajectories in both Mandarin and English vowels conformed to those of monolingual children. However, their Mandarin vowels still showed differences from the M-mono children in either the magnitudes of the trajectories or the relative positions of the vowels. Specifically, their Mandarin /i/ showed much less formant movement than was found for the M-mono children, which might be affected by less formant movement of English /i/. These results indicated that the Bi-high children's Mandarin vowels were affected by their English and showed different formant movement patterns from the M-mono children.

One-way repeated measures ANOVAs on the TLs for the two groups of bilingual children revealed no significant differences between Mandarin /a/ and English /æ/ and /ɑ/, nor between Mandarin /i/ and English /i/. However, for Mandarin /u/ and English /u/, the Bi-low children produced significantly shorter TLs for English /u/ than for Mandarin /u/ ($F(1,6) = 25.334$, $p = 0.002$, $\eta_p^2 = 0.809$) but the Bi-high children produced significantly longer TLs for English /u/ than for Mandarin /u/ ($F(1,7) = 27.586$, $p = 0.001$, $\eta_p^2 = 0.798$). These results indicated that both groups of bilingual children differed from monolingual children in the magnitude of spectral change, which suggests the influence of language experience on the vowel properties of L1 and L2 in young bilingual children.

4. General discussion

The dynamic interaction between L1 and L2 is always a core issue in bilingualism research. This study was designed to examine the influence of L1–L2 interactions on the structure of vowel systems in young sequential bilingual children. In particular, our research questions were, first, how the vowel structures of bilingual Mandarin-English children differ from those of the corresponding monolingual children during the process of L2 learning; and second, what kind of underlying adaptive mechanism is used by bilingual children to restructure their vowel systems as a consequence of the L1–L2 interactions?

4.1. Bidirectional interaction in young Mandarin-English bilingual children

In the present study, we observed interaction effects going in both directions in young bilingual children. However, the strength and size of the bi-directional effect were determined by different stages of L2 learning in the bilingual children. Consistent with our prediction, the Bi-low children showed a strong effect of L1 on their production of L2 vowels. Similar to inexperienced adult L2 learners (Baker & Trofimovich, 2005; Oh et al., 2011), these young sequential bilingual children also adopted the phonetic system of their native language as the base to build the new sound systems for the L2 during the early stage of L2 learning. Mandarin has a smaller inventory of monophthongs than English. The overlapped English vowel clusters and greater variation in each English vowel category in the Bi-low children indicated that this group of children had not established separate acoustic categories for English vowels.

The influence of L1 on L2 in early stage L2 learners is not only manifested on their production of L2 sounds but also represented in their perceptual structure of L2 sounds. Sun and van Heuven (2007) reported data on the perceptual assimilation of English vowels by Mandarin listeners. Listeners were presented with a set of English vowels embedded in the phonetic structure of /hVba:/ and were asked to select one Mandarin vowel that was most similar to the vowel they heard. The results revealed that English vowels /i/, /ɪ/, and /e/ were perceptually assimilated into Mandarin /eɪ/, English /ɛ/ and /æ/ were identified as Mandarin /a/, English /ʊ/ and /oʊ/ were categorized into Mandarin /o/ and English /ɔ/, /ɑ/, /ʌ/ were highly confused and grouped into Mandarin /a/. The perceptual assimilation pattern in their study highly matched the four groups of English vowels produced by the Bi-low children in the present study, which evidenced the linkage between perception and production in L2 acquisition. In addition, the Bi-low children produced English shared vowels with the same formant movement direction as their Mandarin counterparts. The TLs of English shared vowels produced by the Bi-low children were more similar to the Mandarin vowels and were significantly longer than those in the Bi-high and E-mono children. These findings evidenced that the Bi-low children transferred the spectral change pattern in their native language to the second language.

Unlike the Bi-low children, there was much less influence of L1 on L2 in the Bi-high children. Both the static and dynamic vowel spectral features in the Bi-high children showed no significant difference from the E-mono children even though slight positional changes of certain vowels and longer TLs of English /ɪ/ in the Bi-high children relative to the E-mono children can be observed. This finding suggested that the Bi-high children had established a separate L2 vowel system. Although the Bi-high children produced English /ɪ/ with TLs similar to the Bi-low children (which may be associated with the longer TL of the acoustically similar Mandarin vowel /i/), we still lacked solid evidence to show the influence of L1 on L2 in the Bi-high children.

Inconsistent with our predictions, there was no solid evidence to show the effect of L2 learning on L1 in the Bi-low children. Their Mandarin vowel system showed no statistical difference from the M-mono children even though slight positional variation was observed in their production of the vowel /ɤ/. However, the Bi-high children produced their native (Mandarin) vowels /a, /i/, /y/ and /ɤ/ differently from the M-mono children in both static and/or dynamic acoustic characteristics, which provided evidence to support the influence of L2 on L1. Among these four Mandarin vowels, /a/ and /i/ have phonetically similar English counterparts. We observed that the Bi-high children raised Mandarin /a/ close to English /ɑ/ and /æ/. In addition, the Bi-high children transferred the dynamic feature of English /i/ to the Mandarin counterpart and produced Mandarin /i/ with much shorter TLs compared to those in the M-mono children. These findings indicate that the effect of L2 on L1 may not only be manifested on static acoustic features. Vowel dynamic spectral characteristics are also subject to the interaction effect.

For the other two Mandarin vowels /y/ and /ɤ/ that have no phonetically similar English counterparts, the Bi-high children showed forward positional change of Mandarin /y/ relative to the M-mono children. Notice that the formant trajectory of Mandarin /y/ in the Bi-high children was characterized by less formant movement. We assumed that the fronted and static /y/ in the Bi-high children might be affected by their English high front vowels. Mandarin /y/ is produced with protruded lips that cause a relatively back tongue placement along the vocal tract in comparison to high front vowel /i/. English has two high front vowels /i/ and /ɪ/ but no vowel in a high center position. In addition, these English vowels have a generally smaller amount of formant movement than Mandarin vowels. Regarding the vowel /ɤ/, the Bi-high children produced this vowel with less amount of spectral change and tended to move this vowel to a central position. The centralization was also observed in the Bi-low children. We assumed that English vowels /ε/ and /ʌ/ might play a role in the centralization of Mandarin /ɤ/. Mandarin lacks vowels in the mid-low position whereas the two mid-low vowels /ε/ and /ʌ/ are located in a relatively central position in the acoustic vowel space. It is likely that these two acoustically centralized vowels in English might interfere with the production of Mandarin /ɤ/ in the bilingual children.

Chang (2012, 2013) compared the phonetic drift of L1 in inexperienced and experienced Korean-learning English speakers. He found that while inexperienced Korean learners exhibited relatively pronounced shifts in their production of English stops and vowels, experienced Korean learners showed less phonetic drift. The author proposed that it could be the novelty effect on L1 phonetic drift in adult L2 learners. In the present study, no evident phonetic drift of L1 sounds was found in the Bi-low children. However, the Bi-high children presented substantial drift of static and dynamic vowel spectral features in their L1 vowels. These findings are contradictory to the findings in Chang (2012, 2013).

We assumed that regardless of the different target languages examined in the present study from Chang's studies, the differences between our findings and those of Chang (2012, 2013) might be partially related to the distinctive L2 learning strategies in children vs. adults. Previous studies have suggested that adults outperform children in the production of L2 sounds (Oh et al., 2011) and perception of L2 vowel contrasts (Baker et al., 2008) at the beginning stage of L2 learning. Considering that the adult learners in Chang (2012, 2013) received intensive Korean (L2) training in the classroom, they were likely to be highly motivated and cognitively concentrated in L2 learning at the initial stage. In addition, given the evidence showing the novelty effect on language learning in adults (Chang, 2013), the extensive L2 training in the adults was able to affect their L1 within such a short-term immersion. In contrast, although young children can successfully develop a native-like language ability for a new language in a short period, they receive different types of language input and are immersed in a different language learning environment than adults (Jia & Aaronson, 2003). Moreover, an earlier study on phonetic separation in an emergent bilingual Mandarin-English child found that young bilingual children followed a certain developmental path in forming the phonetic system in an L2 (Yang et al., 2015). The Bi-low children in the present study had only lived in the U.S. for less than six months, which might not be long enough for them to fully establish a separate English vowel system. Therefore, the developing L2 system might not be stable and strong enough to exert influence on their L1.

4.2. Assimilatory mechanism in young bilingual Mandarin-English children

Regarding the mechanisms of phonetic movement adopted by these bilingual children, unlike previous studies that revealed the deviation of L1 sounds from similar L2 sounds (e.g. Yusa et al., 2010) or both assimilatory and dissimilatory processes in young bilingual children (Lee & Iverson, 2012; Oh et al., 2011), the present study provides evidence in favor of the assimilatory movement for both L1 and L2 sounds. As to the assimilatory movement of L2 sounds to L1 sounds, consistent with previous studies which found adaptive usage of native phonetic features in building the L2 sound system in adult L2 learners (e.g. Chen et al., 2001; Liao, 2006; Wang, 1997), the present study revealed that young bilingual children also borrowed phonetic features from L1 to form the new vowel structure. For example, the Bi-low children produced English /u/ with a further back position closer to their Mandarin /u/ in comparison to the Bi-high children and E-mono children. In addition, the formant trajectories of the English vowels /i/ and /u/ in the Bi-low children moved in the same direction as their Mandarin counterparts but opposite to those in the Bi-high and E-mono children. These features in the Bi-low children suggest that the assimilatory mechanism is manifested not just in the shift of relative position in the acoustic space but also in the pattern of spectral change.

As posited in SLM, once the L2 sound categories are successfully established, L2 speakers tend to dissimilate the L1 sounds from the L2 sounds to maintain the phonetic contrast between the two languages. In the present study, the Bi-high children had fully acquired English vowel systems. However, divergent from the assumption in the model, the Bi-high children showed an assimilatory movement of L1 sounds to close L2 sounds instead of shifting their Mandarin (L1) vowels away from the English vowels. In particular, the Bi-high children shifted Mandarin /a/ to a higher position closer to English counterparts. They also produced Mandarin /i/ with a shorter formant trajectory, consistent with the magnitude of the formant trajectory of the English /i/. These assimilatory movements were consistent with the results reported in Chang (2012) and Kartushina et al. (2016) for adult L2 learners.

Our tentative thought was that this assimilatory movement might be associated with the amount of language usage and malleability of the phonetic system in the young bilingual children. Bilingual speakers who acquire two languages simultaneously and use both languages equally in their daily life are more likely to develop two monolingual-like contrastive phonetic systems (e.g. MacLeod & Stoel-Gammon, 2005). By contrast, the sequential bilingual children in the present study did not have an equal amount of usage and input in both languages. The Bi-low children used Mandarin as their primary language and spent most of their time speaking Mandarin with their family members and playmates. As they started to learn an L2, the established native language (Mandarin) affected the new language (English) and caused the approximation of the L2 to the L1. On the contrary, the Bi-high children preferred English as their dominant language and spoke English even at home with their parents and playmates. Meanwhile, their usage of Mandarin dramatically decreased. Many of these Bi-high children rarely used Mandarin in their daily life. Earlier research has shown that young children shift their dominant language from L1 to L2 in the first couple of years of arrival (Jia & Aaronson, 2003) and the proficiency of L1 decreases as a function of increased L2 proficiency during the process of L1 loss (Major, 1992). A great amount of empirical evidence suggests that as young Mandarin (L1)-English (L2) bilingual children grow up, they become heritage speakers of Mandarin and speak mainly English in their daily life outside the home. Chang et al. (2011) found that although heritage speakers of Mandarin maintained Mandarin-English contrasts, the heritage speakers with low exposure to Mandarin performed similarly, to some extent, to native English speakers who learned Mandarin as an L2. Based on this finding, it is reasonable to assume that the Bi-high children performed more like native English speakers than like native Mandarin speakers in their speech production. On the other hand, as suggested by Jiang (2008), "the L1 phonetic system established in childhood does not remain static" (pp. iv). Earlier studies have shown that a listener's perception of speech is shaped by their linguistic experience (e.g. Miyawaki et al., 1975) and the underlying perceptual structure can be changed by continuing immersion and exposure to a second language (Yang & Fox, 2014). The extensive immersion to English and reduced experience in Mandarin may cause a gradual change in the original phonetic representation of L1 sounds in young bilingual Mandarin-English children. The joint influence of all these factors on the Bi-high children as a whole might cause the assimilatory shift of unstable L1 sounds toward the close L2 sounds.

Before we draw conclusions, two other observations are worth further discussion. First, as shown in Fig. 7, the Bi-high children showed greater positional separation for certain English vowel categories such as /i/ and /e/ as well as /ɛ/ and /æ/ in comparison to the E-mono children. Flege et al. (2003) found "overshooting" of the formant movement English (L2) /e/ relative to Italian (L1) /e/ in early bilinguals with low L2 proficiency. Similarly, the Bi-high children in the present study might "overshoot" the production of these English vowels to emphasize the separation of English vowel categories. However, we still lacked convincing evidence to show this pattern in the Bi-high children because on one hand, no significant difference was obtained between the Bi-high and E-mono children with regard to the relative position of English vowels. On the other hand, some other English vowels (e.g. e-ɪ) showed even lesser degree of separation in the Bi-high children relative to the E-mono children.

Second, it is worth noting the difference of language input the Bi-low and Bi-high children received. Although both groups of bilingual children had Mandarin-speaking parents, the Bi-high children's parents had lived in the U.S. for a long period while the Bi-low children's parents were newcomers to the U.S. As pointed out by previous studies, adult learners' phonetic systems in L1 are also subject to the influence of L2 (Chang, 2012, 2013; Flege et al., 2003; Guion, 2003). Thus, the Mandarin used by the Bi-low children's parents likely systematically differed from the Mandarin spoken by the Bi-high children's parents (Chang et al., 2011; Polinsky & Kagan, 2007). As the L1 input received by the Bi-high children was likely "colored" by the L2, the Bi-high children's production of L1 vowels was more likely to differ from homeland monolingual norms.

5. Conclusion

In sum, our findings suggest that when sequential bilingual children acquire a second language, they maintain their L1 acoustic features and transfer their L1 features to the new phonetic system at the beginning stage of L2 learning. After bilingual children have been exposed to English for a relatively long period, the L2 phonetic systems are well established. The bilingual children are less likely to retain the acoustic properties in their native language while they develop a native-like ability in L2. In particular, the influence of L2 on their L1 production is presented in the form of assimilatory process that causes their L1 vowels to shift toward similar L2 vowels. Although informative, this present study had a small number of participants, especially the bilingual children. A future study with more participants is needed to verify the current findings. In addition, as the present study only focused on monophthongal vowel phonemes, it would also be of interest to examine the production of compound vowels, especially the similar diphthongs produced by bilingual Mandarin-English children. This would better enable us to observe the adaptive mechanism the bilingual children used in organizing the phonetic systems. It is noteworthy that the participants in the present study were young children whose language development was not complete at the age when the study was conducted. It is very likely that the acoustic–phonetic properties of their vowel production will continue changing as their language experiences change. Therefore, future research is needed to trace the process of language development in L1 and L2 by young bilingual children. This type of longitudinal investigation will further our understanding of the dynamic interaction between L1 and L2 during the ongoing process of speech development in bilingual children.

Acknowledgments

This research was supported, in part, by the Ruth Becky Irwin Fund in the Department of Speech and Hearing Science at The Ohio State

University. Part of this study was reported in INTERSPEECH 2014. The authors would like to thank Prof. Erika Levy and the four anonymous reviewers for their valuable and constructive comments to improve the quality of the paper. The authors would also like to thank the children and their parents for participating in this study.

References

Aoyama, K., Guion, S., Flege, J. E., Yamada, T., & Akahane-Yamada, R. (2008). The first years in an L2 speaking environment: A comparison of Japanese children and adults learning American English. *International Journal of Applied Linguistics, 46*, 61–90.

Assmann, P. F., & Katz, W. F. (2000). Time-varying spectral change in the vowels of children and adults. *Journal of the Acoustical Society of America, 108*(4), 1856–1866.

Assmann, P. F., Nearey, T. M., & Bharadwaj, S. V. (2013). Developmental patterns in children's speech: Patterns of spectral change in vowels. In G. S. Morrison & P. F. Assmann (Eds.), *Vowel inherent spectral change* (pp. 199–230). NY: Springer.

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

Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child-adult differences in second-language phonological learning: The role of cross-language similarity. *Language and Speech, 51*, 317–342.

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.

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). Timonium, MD: York.

Bohn, O.-S., & Flege, J. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14*, 131–158.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America, 54*, 421–428.

Chang, C. Y. (2010). *Dialect differences in the production and perception of mandarin chinese tones*. (Doctoral Thesis) Retrieved from ProQuest Dissertations and Theses. (UMI No. 3435661).

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics, 40*, 249–268.

Chang, C. B. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics, 41*(6), 491–504.

Chang, C. B., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *Journal of the Acoustical Society of America, 129*(6), 3964–3980.

Chen, Y. (2006). Production of tense-lax contrast by Mandarin speakers of English. *Folia Phoniatrica et Logopaedica, 58*(4), 240–249.

Chen, Y., Robb, C., Gilbert, H., & Lerman, J. (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics, 15*(6), 427–440.

Clopper, C. G., Pisoni, D. B., & De Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America, 118*(3), 1661–1676.

Duanmu, S. (2007). *The phonology of standard Chinese*. NY: Oxford University Press.

Edwards, J., & Beckman, M. E. (2008). Methodological questions in studying consonant acquisition. *Clinical linguistics & phonetics, 22*(12), 937–956.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*(1), 47–65.

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

Flege, J., MacKay, I., & Meador, D. (1999). Native Italian speakers' production and perception of English vowels. *Journal of the Acoustical Society of America, 106*, 2973–2987.

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

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. 233–273). Timonium, MD: York Press.

Fox, R. A., & Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. *Journal of the Acoustical Society of America, 126*, 2603–2618.

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

Guion, S. (2003). The vowel systems of Quichua-Spanish bilinguals: An investigation into age of acquisition effects on the mutual influence of the first and second languages. *Phonetica, 60*, 98–128.

Harada, T. (2003). L2 influence on L1 speech in the production of VOT. In M.-J. Sole, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th international congress of phonetic sciences* (pp. 1085–1088). Barcelona, Spain: Causal Productions.

Harrington, J., & Cassidy, S. (1994). Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English. *Language and Speech, 37*, 357–373.

Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech, 36*, 17–38.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*(5), 3099–3111.

Højen, A., & Flege, J. (2006). Early learners' discrimination of second-language (L2) vowels. *Journal of the Acoustical Society of America, 119*, 3072–3084.

Hoole, P., Hu, F. (2004). Tone-vowel interaction in standard Chinese. *International symposium on tonal aspects of languages: With emphasis on tone languages*. Retrieved from http://sprosig.isle.illinois.edu/tal2004/tal2004-Beijing/Hoole-Hu.pdf on March 21, 2016.

Jacewicz, E., Fox, R. A., & Salmons, J. (2011a). Regional dialect variation in the vowel systems of typically developing children. *Journal of Speech, Language and Hearing Research, 54*, 448–470.

Jacewicz, E., Fox, R. A., & Salmons, J. (2011b). Vowel change across three age groups of speakers in three regional varieties of American English. *Journal of Phonetics, 54*, 448–470.

Jacewicz, E., Fox, R. A., & Salmons, J. (2011c). Cross-generational vowel change in American English. *Language Variation and Change, 23*(1), 45–86.

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics, 24*(1), 131–161.

Jiang, H. (2008). *Effect of L2 phonetic learning on L1 vowels*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. AAT NR58514).

Kang, K.-H., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *Journal of the Acoustical Society of America, 119*(3), 1672–1683.

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2016). Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training. *Journal of Phonetics, 57*, 21–39.

Kent, R. D., & Read, C. (1992). *The acoustic analysis of speech*. San Diego, CA: Singular Publishing.

Keshavarz, M., & Ingram, D. (2002). The early phonological development of a Farsi-English bilingual child. *International Journal of Bilingualism, 6*, 255–269.

Khattab, G. (2000). VOT production in English and Arabic bilingual and monolingual children. *Leeds Working Papers in Linguistics, 8*, 95–122.

Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English: Phonology and phonetics*. Berlin: Mouton de Gruyter.

Lee, S., & Iverson, G. K. (2009). Vowel development in English and Korean: Similarities and differences in linguistic and non-linguistic factors. *Speech Communication, 51*(8), 684–694.

Lee, S. A. S., & Iverson, G. K. (2012). Vowel category formation in Korean-English bilingual children. *Journal of Speech, Language, and Hearing Research, 55*(5), 1449–1462.

Liao, J. S. (2006). *English vowel production of mandarin speakers*. (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3335222).

Lin, T., & Wang, L. (2001). *Yuyinxue Jiaocheng (A course in phonetics)*. Beijing: Beijing Daxue Chubanshe (Peking University Press).

Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception and Psychophysics, 46*, 187–200.

Mack, M. (1990). Phonetic transfer in a French-English bilingual child. In P. H. Nelde (Ed.), *Language attitudes and language conflict* (pp. 107–124). Bonn, Germany: Dümmler.

MacLeod, A. A., & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders, 3*, 118–127.

MacLeod, A. A., Stoel-Gammon, C., & Wassink, A. B. (2009). Production of high vowels in Canadian English and Canadian French: A comparison of early bilingual and monolingual speakers. *Journal of Phonetics, 37*, 374–387.

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

McLachlan, G. J. (1999). Mahalanobis distance. *Resonance, 4*(6), 20–26.

Meador, D., Flege, J. E., & MacKay, I. R. A. (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition, 3*, 55–67.

Milenkovic, P. (2003). "TF32 software program" (University of Wisconsin, Madison, WI), http://userpages.chorus.net/cspeech/ (Last viewed 19 January 2017).

Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*(5), 331–340.

Nearey, T. M., & Assmann, P. F. (1986). Modeling the role of inherent spectral change in vowel identification. *Journal of the Acoustical Society of America, 80*, 1297–1308.

Nearey, T. M. (2013). Vowel inherent spectral change in the vowels of North American English. In G. S. Morrison & P. F. Assmann (Eds.), *Vowel inherent spectral change* (pp. 49–85). Berlin Heidelberg: Springer.

Netelenbos, N., & Li, F. (2013). The production and perception of voice onset time in English-speaking children enrolled in a French-immersion program. *Proceedings of the 14th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, Lyon, France, pp. 2380–2383.

Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2011). A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of Phonetics, 39*, 156–167.

Pallier, C., Bosch, L., & Sebastian-Gallés, N. (1997). A limit on behavioral plasticity in speech perception *Cognition, 64*(3), B9–B17.

Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism, 5*, 19–38.

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition, 18*, 1–25.
Polinsky, M., & Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass, 1*(5), 368–395.
Sebastian-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition, 72*(2), 111–123.
Stiles, J. (2000). Neural plasticity and cognitive development *Developmental Neuropsychology, 18*(2), 237–272.
Sun, L., & van Heuven, V. J. (2007). Perceptual assimilation of English vowels by Chinese listeners: Can native-language interference be predicted? *Linguistics in the Netherlands, 24*(1), 150–161.
Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. E. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics, 33*(3), 263–290.
Turner, J., Netelenbos, N., Rosen, N., & Li, F. (2015). Stop consonant production by French-English bilingual children in Southern Alberta. *Proceedings of the 18th international congress of phonetic sciences, Glasgow*, Scotland.
Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research, 50*(6), 1510–1545.
Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America, 117*(1), 338–350.
Wang, X. (1997). *The acquisition of English vowels by Mandarin ESL learners: A study of production and perception*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. AAT MQ24263).
Wassink, A. B. (2006). A geometric representation of spectral and temporal vowel features: Quantification of vowel overlap in three linguistic varieties. *Journal of the Acoustical Society of America, 119*, 2334–2350.
Watson, I. (1991). Phonological processing in two languages. In E. Bialystock (Ed.), *Language processing in bilingual children* (pp. 25–48). Cambridge, United Kingdom: Cambridge University Press.
Yang, J. & Fox, R. A. (2013). Acoustic development of vowel production in American English children. *Proceedings of the 14th annual conference of the international speech communication association (Interspeech 2013)*, Lyon, France.
Yang, J., & Fox, R. A. (2014). Perception of English vowels by bilingual Chinese-English and corresponding monolingual listeners. *Language and Speech, 57*(2), 215–237.
Yang, J., Fox, R. A., & Jacewicz, E. (2015). Vowel development in an early Mandarin-English bilingual child: A longitudinal study. *Journal of Child Language, 42*(5), 1125–1145.
Yusa, N., Nasukawa, K., Koizumi, M., Kim, J., Kimura, N., & Emura, K. (2010). Unexpected effects of the second language on the first. In K. Dziubalska-Ko□aczyk, M. Wrembel, & M. Kul (Eds.), *New sounds 2010: Proceedings of the 6th international symposium on the acquisition of second language speech* (pp. 580–584). Poznan, Poland: Adam Mickiewicz University.
Zee, E. (2001). The phonetic value of the vowels, diphthongs, and triphthongs in Beijing Mandarin. *Proceedings of the 5th national conference on modern phonetics*, pp. 54–60.
Zhou, N., & Xu, L. (2008). Lexical tone recognition with spectrally mismatched envelopes. *Hearing Research, 246*, 36–43.