



ELSEVIER

Contents lists available at ScienceDirect

Journal of Phonetics

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

A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting

Grace E. Oh^{a,*}, Susan Guion-Anderson^a, Katsura Aoyama^b, James E. Flege^c,
Reiko Akahane-Yamada^d, Tsuneo Yamada^e

^a University of Oregon, USA

^b Texas Tech University Health Sciences Center, USA

^c Via della Torretta 12, 01017 Tuscania (VT), Italy

^d ATR Human Information Processing Research Laboratories, Japan

^e National Institute of Multimedia Education, Japan

ARTICLE INFO

Article history:

Received 22 March 2010

Received in revised form

3 December 2010

Accepted 14 January 2011

ABSTRACT

The effect of age of acquisition on first- and second-language vowel production was investigated. Eight English vowels were produced by Native Japanese (NJ) adults and children as well as by age-matched Native English (NE) adults and children. Productions were recorded shortly after the NJ participants' arrival in the USA and then one year later. In agreement with previous investigations [Aoyama, Flege, Guion, Akahane-Yamada, & Yamada (2004). *Journal of Phonetics*, 32, 233–250], children were able to learn faster, leading to higher accuracy than adults in a year's time. Based on the spectral quality and duration comparisons, NJ adults had more accurate production at Time 1 but showed no improvement over time. The NJ children's productions, however, showed significant differences from the NE children's for English "new" vowels /ɪ/, /ɛ/, /ɑ/, /ʌ/ and /ʊ/ at Time 1 but produced all eight vowels in a native-like manner at Time 2. An examination of NJ speakers' productions of Japanese /i/, /a/, /u/ over time revealed significant changes for the NJ Child Group only. Japanese /i/ and /a/ showed changes in production that can be related to second language (L2) learning. The results suggest that L2 vowel production is affected importantly by age of acquisition and that there is a dynamic interaction, whereby the first and second language vowels affect each other.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

The effect of age on second language (L2) acquisition in bilingual speech production has been demonstrated in many studies (e.g., Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Aoyama, Guion, Flege, Yamada, & Akahane-Yamada, 2008; Baker & Trofimovich, 2005; Baker, Trofimovich, Flege, Mack, & Halter, 2008; Flege et al., 2006; Flege, MacKay, & Meador, 1999; Guion, 2003; Kang & Guion, 2006). As an example, Aoyama et al.'s (2004) study examined Japanese children and adults learning American English for one year in the United States (US) and found that the Japanese children's but not adults' productions of English /l/ and /r/ improved significantly in a year's time. In another study, Kang and Guion (2006) reported that late Korean–English bilinguals ($M=21.4$ years old at time of arrival in the US) showed significantly different VOT, H1–H2 (i.e., voice quality of the following vowel), and f_0 values in producing

voiced and voiceless English consonants from English monolinguals, whereas early Korean–English bilinguals ($M=3.8$ years old at time of arrival in the US) produced both English consonants in a native manner. Furthermore, early bilinguals were able to distinguish English stops from Korean stops in production, establishing five distinct stop types, while late bilinguals only showed three stop types in their combined Korean–English systems.

Relatedly, early bilinguals have been shown to have a better ability to separate the L2 from the first language (L1) system in production with increases in amount of experience than late bilinguals. Baker and Trofimovich (2005) compared the L1–L2 interaction of early and late bilinguals with one and seven years of L2 exposure. Through a comparison of early and late bilinguals' English vowel production with that of monolinguals, the authors were able to show that early bilinguals were much better than late bilinguals at distinguishing their L1 and L2 systems after seven years of experience. In addition, Guion (2003) examined simultaneous, early, and late Quichua (L1)–Spanish (L2) bilinguals and found that simultaneous and early bilinguals were able to accurately acquire Spanish vowels whereas late bilinguals seemed to be using their Quichua vowels in Spanish production. More specifically,

* Corresponding author.

E-mail address:

goh1@uoregon.edu (G.E. Oh).

Guion (2003) found that the early bilinguals established a greater number of distinct vowel categories for each language than the late bilinguals. These studies suggest that early bilinguals are not only faster at learning a new language but also have a higher chance of creating L2 phonetic categories that are distinct from the L1 than late bilinguals, as hypothesized by the Speech Learning Model (e.g., Flege, 1995, 2002, 2003).

It is not necessarily the case, however, that all early learners produce an L2 with no influence from their L1. Some studies revealed that even early bilinguals may show a notable L1 influence on their production of the L2 (Asher & Garcia, 1969; Baker & Trofimovich, 2005; Baker et al., 2008; Flege et al., 2006). Such findings suggest that even a substantial amount of early exposure to the L2 may not totally override the effect of the L1 on the L2 system. However, other studies have reported some native-like aspects of L2 speech production among early learners (Flege et al., 1999; Guion, 2003; Kang & Guion, 2006; Tsukada et al., 2005).

In cases where early L2 learners have been found to differ from L2 native speakers, continuing interaction between the L1 and L2 sound systems is thought to be a primary obstacle in the acquisition of specific L2 sounds. Studies investigating the perceptual similarity between L2 and L1 segments have found differential effects on the production of L2 segments, depending on the degree of cross-language similarity. For example, Baker et al. (2008) investigated Korean adults' ($M=25$ years old at time of arrival) and children's ($M=10$ years old at time of arrival) production of four English vowels. The participants' productions were compared to those of English monolinguals in order to determine whether L1–L2 perceptual similarity (i.e., similarity judgment and goodness of fit rating) would predict L2 production accuracy. "New" English vowels, defined as vowels without a close counterpart in Korean (i.e., English /ʊ/), appeared to be harder for Korean children to learn to produce than "similar" vowels, defined as English vowels having a counterpart in Korean (i.e., English /u/). However, the children did produce new English vowels such as /ɪ/ and /ʊ/ more accurately than the adults did.

A proposed explanation for differences in acquiring new and similar sounds rests upon the idea that the degree of L2 production accuracy may be limited by perceptual factors (Flege, 1981, 1995; Ioup, 1995). The Perceptual Assimilation Model (PAM) by Best (1994, 1995) postulates that the discrimination of the difference between pairs of phonemically distinct L2 sounds is influenced by the perceptual similarity between specific L1 and L2 sounds. Furthermore, Best and Tyler (2007) have proposed that perceptual similarity can affect L2 speech learning.

In addition to the effects of the L1 on the L2, the L1 itself may be affected by L2 learning. Recently, de Leeuw, Schmid, and Mennen (2010) reported that the production of German (L1)–English (L2) bilingual adults ($M=27$ years old at time of arrival) who resided in Canada or the Netherlands for an average of 37 years were perceived to be non-native speakers of German (their L1) by 19 German monolinguals. Even when German bilingual speakers had no extensive L2 learning prior to immigration and the L2 was acquired after adolescence, 37 out of 57 German bilinguals were evaluated as having detectable global foreign accents in German. The authors argued that the extensive exposure to the L2 created changes in the phonetic elements of the L1, although the overall degree of perceived foreign accent in the L1 varied by the degree of participation in communicative settings in which the L1 was exclusively used.

Studies focusing on the production of individual segments have also found changes in the L1 related to L2 learning. Flege (1987) compared the /t/ production of French (L1)–English (L2) and English (L1)–French (L2) bilinguals. It was found that French–English bilinguals produced French /t/ with longer VOT than French monolinguals and English–French bilinguals produced English /t/ with

shorter VOT than English monolinguals, suggesting an effect of the L2 on L1 VOT. Also, Guion (2003) found that Quichua vowels produced by early Quichua–Spanish bilinguals were higher in the vowel space than those produced by Quichua monolinguals. The acquisition of Spanish /e/ and /o/ appeared to lower F1 values for Quichua /i/ and /u/. Guion proposed that the raising of Quichua vowels relative to the Spanish vowels served to maintain a sufficient perceptual distance between the L1 and L2 vowel systems in what may be termed category dissimilation (see below).

According to the Speech Learning Model (SLM) proposed by Flege (1995, 2002, 2003), category *assimilation* occurs when a new category fails to be established for an the L2 speech sound, despite audible differences between the L2 and closest L1 speech sound. By hypothesis, category formation will be blocked if instances of an L2 speech category continue to be identified as instances of an L1 category. That is, in cases where the L1 and L2 sounds have been "perceptually equated" a merged category will develop over time (see also Kingston, 2003). On the other hand, the SLM claims that phonetic category *dissimilation* may occur when a new category has been established for an L2 speech sound. The dissimilation is caused by a newly established L2 category and the nearest L1 speech category shifting away from one another in phonetic space to maintain phonetic contrast between elements found in the combined L1–L2 phonological space. The SLM proposes that the greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the greater the likelihood of category formation for the L2 sound (see Flege, 1995). Additionally, the SLM proposes that younger learners are more likely to develop new L2 categories and, so, are more likely to exhibit dissimilation.

Researchers working in the area of L2 speech acquisition typically recognize two types of category dissimilation. First, Flege, Schirru, and MacKay (2003) provided evidence of dissimilation of an L2 vowel from the closest L1 vowel. They compared early and late Italian–English bilinguals who differed according to self-reported frequency of Italian use (relatively low vs. high) and compared their productions of English /eɪ/ to those of monolingual English speakers, whose typical /eɪ/ productions show far more formant movement than is typical for the Italian /e/. Flege et al. (2003) found more formant movement in English vowels produced by the early-low bilinguals and less movement in late bilinguals compared to the vowels produced by English monolinguals. The authors posited that early-low bilinguals "overshot" formant movement in English /eɪ/ as the result of dissimilation of the newly acquired English /eɪ/ from the virtually steady-state Italian /e/, whereas the late bilinguals' "undershoot" of formant movement indicated assimilation of English /eɪ/ to Italian /e/.

The second type of dissimilation involves the shift of an L1 category. For example, Flege and Eefting (1988) investigated the VOT of Spanish–English bilinguals' and Spanish monolinguals' production for Spanish /p, t, k/ and found that early, but not late, bilinguals produced shorter VOT for their Spanish consonants compared to the monolinguals. The results were interpreted to mean that early bilinguals shortened the VOT of Spanish consonants to create a greater distinction from the English consonants. These studies indirectly supported the SLM's prediction that younger learners would be more likely to exhibit category dissimilation between their L1 and L2 than adult learners. In the present study, we sought to further investigate the effect of age of acquisition on the likelihood of new L2 category formation and to determine if such category formation would result in dissimilation between first and second language vowels. More specifically, the current study will focus on the production of English vowels by Japanese learners of English during their first years in an English-speaking setting. A brief review of previous studies in this area is provided in the following section.

1.1. Native Japanese speakers' production of English vowels

Generally, whether one is considering vowel duration or spectral quality, Japanese speakers show difficulty learning some English vowels (Tsukada, 1999, 2009; Lambacher, Martens, Kakehi, Marasinghe, & Molholt, 2005). For example, Tsukada (1999) reported that the English vowels produced by Japanese learners revealed extensive overlap between neighboring vowels, especially for English /ɑ/–/ʌ/ and /ʊ/–/u/. In other words, Japanese learners seem to assimilate English /ɑ/–/ʌ/ and /ʊ/–/u/ into the single categories of Japanese /a/ and /u/ in production, respectively. Japanese learners seem to use a vowel category that encompassed Japanese /a/, English /ɑ/ and English /ʌ/ and another that encompassed Japanese /u/, English /u/ and English /ʊ/. As a result, the use of the F1 and F2 dimensions for each English vowel was much more widely dispersed in English production by Japanese learners than in production by English monolinguals. Tsukada's study provided valuable insights into an interim state of Japanese accented English, but leaves uncertain the source of inter-subject variability. The participants were residing in Australia from 3 weeks to 16 years when tested. If a relation exists between amount of L2 experience and L2 vowel production accuracy, we can infer that important learning took place over time. If not, we would need to infer that success in producing L2 vowels depends on as yet-undefined "individual differences".

In a training study for English learning adults, Lambacher et al. (2005) investigated whether native Japanese speakers would improve in identifying and producing English low/mid-vowels, /æ/, /ɑ/, /ʌ/, /ɔ/ and /ɜ/, after 6 weeks of auditory training. The results of the pre-/post-identification tests revealed a significant effect of training, showing a notable improvement of the trained group (16% gain) compared to the untrained group (5% gain). The identification scores improved more than 20% for all but the /ɑ/, /ʌ/ and /æ/ vowels, which showed less improvement at the post-test (15%, 11% and 10% gains, respectively). Additionally, the production post-tests, which measured three formant frequencies and vowel duration, also revealed a positive effect of auditory training. Specifically, the participants were able to produce /ɜ/, /æ/ and /ɔ/ with more accuracy in the post-test. However, /ɑ/ and /ʌ/ were still significantly different in the post-test from those produced by English speakers (/ɑ/ produced with higher F2 and /ʌ/ with higher F1). As for duration, the significantly longer /ʌ/ produced by Japanese participants substantially shortened after identification training. The durational difference in the /ɑ/–/ʌ/ contrast also became native-like in the post-test. Lambacher et al. indicated that Japanese adults were able to gain accuracy in both perceiving and producing some L2 vowels but not others, such as /ɑ/ and /ʌ/, due to their perceived similarities to L1 sounds.

1.2. The current study

The study reported here focused on factors affecting L2 vowel acquisition, as well as L1–L2 vowel interaction. More specifically, in Experiment 1, the effect of one year of L2 exposure and age of acquisition on L2 vowel production was investigated. In Experiment 2, the mutual effects of the L2 and L1 on vowel production were further examined. English and Japanese vowel productions were elicited from native Japanese (NJ) adults and children approximately 5 months after their arrival in an English-speaking environment and then another sample was recorded roughly one year later. Also, English vowel productions from age-matched native English (NE) adults and children were elicited. Results of several aspects of these data have been reported previously

(Aoyama et al. (2004), for production and perception of English /r/ and /l/; Aoyama et al. (2008), for production and perception of English fricatives and perceived foreign accent), but no published report has examined vowel production.

In Experiment 1, we addressed the hypothesis that NJ children would show greater improvements over a 1-year period in their production of English vowels than the NJ adults across the one-year time span of the study. Although children are known to be better at learning new sounds in early stages of L2 acquisition (Baker et al., 2008), given Snow and Hoefnagel-Hoehle (1978) and García Mayo and García Lecumberri (2003)'s studies that non-native adults can be more successful at L2 learning than children until about 4–5 months of immersion exposure, we also hypothesized that the adults in the current study, who had been residing in the US for approximately 5 months, would initially be more proficient than children in English vowel production.

Furthermore, this study investigates the production of the English tense and lax vowels /i/, /eɪ/, /ɑ/, /u/ and /ɪ/, /ɛ/, /ʌ/ and /ʊ/, respectively. Previous studies on Japanese speakers cross-language mapping of these vowels to Japanese vowels found more consistent identification of tense than lax English vowels. Strange et al. (1998) found that English /i/, /eɪ/, /ɑ/ and /u/ were rated as being more similar to Japanese vowels, whereas /ɪ/, /ɛ/, /ʌ/ and /ʊ/ were rated as less similar to vowels in Japanese. Likewise, Frieda and Nozawa (2007) found that inexperienced Japanese listeners identified English /i/, /eɪ/, /ɑ/ and /u/ as the Japanese vowels /i:/, /eɪ/, /a:/ and /u:/ over 90% of the time whereas English /ɪ/, /ɛ/, /ʌ/ and /ʊ/ received less consistent identification as a Japanese vowels. Given the SLM hypothesis that the likelihood of category formation in L2 increases as a function of perceived cross-language phonetic dissimilarity, we will refer to the English lax vowels /ɪ/, /ɛ/, /ʌ/ and /ʊ/ as "new" and the English tense vowels, /i/, /eɪ/, /ɑ/ and /u/ as "similar". Furthermore, according to the SLM's (Flege, 1995, 2002, 2003) hypothesis that L2 sounds that are dissimilar to L1 sounds are likely to be learned, we predicted that the relatively more dissimilar (or "new", for short) vowels /ɪ/, /ɛ/, /ʌ/ and /ʊ/ would show more improvement compared to the other vowels in production at Time 2.

As previous research has yielded differences in terms of whether native-like production was or was not observed for early learners (Baker & Trofimovich, 2005; Baker et al., 2008; Flege et al., 1999; Tsukada et al., 2005), we sought to add to the body of knowledge and determine whether NJ children would attain native-like production of English vowels over the course of a year of residence in an L2 speaking country. In addition to examining the factors of age and duration of L2 exposure on L2 vowel acquisition in Experiment 1, we also investigated whether there were any interactions between the L1 and L2 systems in Experiment 2. In cases where changes in L2 (English) production over the one-year time span were found, the second experiment examined the production of L1 (Japanese) vowels in the neighboring vowel space. We hypothesized that L1 productions could change over time through interaction with the L2 categories. First, if the L2 category were assimilated to the L1 category, it may have taken on characteristics of both the L1 and L2 vowel. In this case, the L1 vowel may come to be produced more like a neighboring L2 vowel over the year of the study. Second, if an L2 vowel were established independently from similar L1 vowel categories, evidence for a dissimilation of the L1 and L2 vowels might be found. In this case, production of the L1 vowel might shift to help it maintain phonetic contrast from a newly formed L2 vowel that is adjacent to it in vowel space. Given the studies reviewed in the section above, we hypothesized that assimilation and dissimilation would take place around the "new" English lax vowels /ɪ/, /ɛ/, /ʌ/ and /ʊ/.

Table 1
Native Japanese and native English participants.

	Number (gender)	Age	Mean LOR	
			Time 1	Time 2
NJ adults	16 (8 female)	39.9 (3.8)	0.5 (0.2)	1.6 (0.3)
NE adults	16 (9 female)	40.3 (4.7)	–	–
NJ children	16 (7 female)	9.9 (2.4)	0.4 (0.2)	1.6 (0.3)
NE children	16 (6 female)	10.6 (2.1)	–	–

Chronological age and length of residence (LOR) in the United States at Time 1 and Time 2 in years are shown with standard deviation in parentheses.

2. Experiment 1: English vowel production

2.1. Method

2.1.1. Participants

The participants of the current study are the same as those who participated in the studies reported in Aoyama et al. (2004, 2008). Sixteen Native Japanese (NJ) adults and 16 NJ children living in Houston and Dallas, Texas participated, as well as 16 age-matched Native English (NE) speaking adults and children living in Birmingham, Alabama as control groups (see Table 1). The NJ and NE adults were the parents of the NJ and NE children. Two sets of recordings were made for all groups with a one-year interval. At the time of the first recording (Time 1), NJ participants had resided in the US approximately 5 months and at the time of the second recording (Time 2), the average length of residence (LOR) was 1.6 years for both NJ children and NJ adults.¹ The 16 NJ adults had studied English in Japan while only one NJ child had experience learning English before coming to the US (4 years of spoken English study). The NJ Adult Group reported an average of 12.1 and 9.8 years learning written and spoken English respectively. A self-report on English use suggested that, on a daily basis, the average time spent speaking English was 3.4 hours for the NJ Adult Group and 4.7 hours for the NJ Child Group. Time spent daily watching English TV and listening to English music was 2.2 hours for the NJ Adult Group and 2 hours for the NJ Child Group.

2.1.2. Stimuli

We recorded 16 words that occur frequently in English and thus were likely to be known by all of the participants. A total of eight vowel categories represented in sixteen words, two words exemplifying each vowel category, were produced by NJ adults, NJ children, NE adults and NE children: /i/ (*feet, eat*), /ɪ/ (*fish, six*), /ɛ/ (*egg, neck*), /eɪ/ (*cage, vase*), /ɑ/ (*dog, sock*), /ʌ/ (*bug, hug*), /ʊ/ (*book, foot*) and /u/ (*food, shoe*). These 16 words were part of a larger list that was recorded. The vowels were all in monosyllables and stressed. The monosyllabic words chosen were constrained by two main factors; they needed to be familiar to beginning learners and children and to be imageable for the elicitation procedure which used pictures of the words. Although the surrounding consonantal context varied due to these constraints, our analysis is not compromised as each vowel category is compared across time of testing or group, and thus, is compared within the same consonantal context.

¹ Because each group consisted of different numbers of male and female participants, ANOVA tests for group comparisons were separately conducted for the male and female groups to insure that the difference was not substantially driven by the effect of gender. The results for gender-matched groups returned the same effects as those found for the combined groups.

2.1.3. Procedure

The experiment was conducted in a quiet room in the home of the participant or at the University of Alabama at Birmingham. Pictures of the English words were presented in a random order to the participants on the screen of a laptop computer. The participants wore a head-mounted Shure microphone (Model SM 10A) and the speech was recorded on a Sony DAT tape recorder at a 22,050 Hz sampling rate with 16 bit quantization. The sixteen words were elicited three times each in random order from each participant. Japanese orthography in Hiragana was also provided with the picture to make sure that the participants clearly understood what the objects shown on the screen were. The experimenter presented a prerecorded production of the corresponding words at the first presentation only and the participants were asked to repeat the words. The experimenter did not provide the auditory cue on the second and third presentations unless the participants were unable to remember the words. Only the second and third tokens of each word, that is, the non-cued productions, were analyzed. The Japanese vowel productions (described in Section 4) were recorded in the same session as the English materials, with intervening tasks and a short break. The same procedures were followed when a second recording of the same materials were made roughly one year later.

2.1.4. Acoustic measurements

2.1.4.1. Formant frequencies. A total of 4096 tokens (16 words × 2 repetition × 2 times of testing × 4 groups × 16 participants) for the English vowel productions were analyzed using PCquirer acoustic analysis software developed by Scicon R&D. First and second formant frequencies of each vowel were measured at the temporal midpoint of each vowel, and first and second formant frequencies of the off-glided /eɪ/ were measured at the temporal 1/4, 1/2, and 3/4 points in order to examine vowel movement. FFT (512 points) and LPC (26 coefficients) spectra were calculated and plotted in a single window. The peaks from the formant-picking algorithm based on the LPC analysis were recorded except in the very few instances (2%) in which the LPC peaks did not align with the FFT spectral peaks on visual inspection. In these cases, the spectra were recalculated 5–10 ms later or earlier in the waveform to find an alignment of spectral peaks. Because the formant frequencies showed little differences across the two repetitions, mean values were used in all formant analyses.

2.1.4.2. Vowel duration. The duration of the 4096 English vowels produced by all groups at Time 1 and Time 2 were measured in milliseconds from spectrographic and time domain waveform displays. The vowels were measured from the onset of voicing in the vowel to the beginning of the consonant constriction (in words such as “feet”) or the end of voicing in the case of the one stimulus that was not closed by an obstruent consonant, that is, “shoe”. The onset and offset of clear energy in the second formant frequency on the sound spectrogram served as a reference, along with the waveform, to determine the onset and offset of the vowel. The mean duration averaged across two repetitions was submitted to analysis.

2.1.5. Statistical analyses

2.1.5.1. English vowels: formant frequency. In order to determine whether a difference existed between vowels produced at Time 1 and Time 2, MANOVAs examining English vowels produced by the two NJ Groups were conducted. The F1 and F2 frequencies obtained at the temporal midpoint were examined. The factors of Vowel (8) and Time (Time 1 vs. Time 2), as a repeated measure, were specified. Because the vowel /eɪ/ has relatively greater formant movement across its duration in comparison to the other

vowels, a separate MANOVA was carried out to examine potential changes in this formant movement that would not be captured by a midpoint measurement.

In the instances in which a significant Time \times Vowel interaction was obtained, 8 MANOVAs testing Time for each vowel were conducted. The alpha level was adjusted to 0.006 for 8 comparisons. Partial eta squared (η_p^2) values are provided for all analyses to provide information on the effect size. The univariate tests for F1 and F2 are reported for each significant MANOVA comparison.

In the case of a significant effect of Time on NJ children's English vowel production, NE children's production across Time was examined to determine whether the formant frequency changes found in NJ children's production were most likely due to developmental effects or to L2 learning. The dependent variables for the NE Child Group were F1 and F2 frequencies and the independent variables were Time (Time 1 vs. Time 2) and Vowel (8). It was theorized that changes in NJ children's vowels that were different from the developmental changes found for NE children's vowels would reflect the effect of L2 learning.

In the next set of statistical analyses, the native-likeness of the NJ Groups' English vowel production was assessed. To investigate which English vowel productions were native-like and which were not for the NJ Adult and Child Groups, comparisons with the NE Groups were made. As reported below, Time 1 and Time 2 differences were not found in the NJ Adult Group. Accordingly, only Time 1 vowels were compared for the NJ and NE Adult Groups. Because differences were found in the NJ Child Group across Time, the NE Child and NJ Child Group productions were compared at both Time 1 and Time 2, in order to determine whether the changes led to more native-like production. The dependent variables for all comparisons were F1 and F2 frequency and the independent variables were Vowel (8), conducted with repeated measures, and Group. A separate MANOVA was conducted for the English vowel /eɪ/. In the case of a significant interaction between Group and Vowel, 8 MANOVAs were conducted to test the effect of Group on each vowel. The alpha level was adjusted to 0.006 for 8 comparisons. The univariate tests for F1 and F2 are reported for each significant MANOVA comparison.

2.1.5.2. English vowels: duration. The mean duration of eight English vowels produced by the NJ Child and Adult Groups were separately analyzed using ANOVAs with a repeated measures design. The effects of Time (Time 1 vs. Time 2) and vowel (8) were investigated. In the case of a significant effect of Time or interaction with Time, comparisons with the NE Groups at Time 1 and Time 2 were made. In the case of no significant effect of Time, comparisons with the NE Groups were only made at Time 1. The group comparisons used the factors of Group (2) and Vowel (8). Significant interactions between Group and Vowel were explored with 8 ANOVAs testing the effect of Group on each Vowel. The alpha level was adjusted to 0.006 for 8 comparisons.

3. Results

3.1. The effect of time on NJ Child and Adult Groups

3.1.1. Formant frequency

Productions of 8 English vowels by the NJ adults and children were compared at Time 1 and Time 2. Our purpose was to determine whether the NJ Adult and Child Groups demonstrated differential learning over time. The vowel productions in terms of F1 and F2 frequencies are shown in Fig. 1(a) and (b) for the NJ Child Group across Time and Fig. 2(a) for the NJ Adult Group at Time 1. Due to the substantial overlap between the trajectory of English off-glided /eɪ/ and the neighboring high-front vowels, the

three temporal points of /eɪ/ were excluded from the figures for clarity's sake. In this and the following figures, the mean is represented by the placement of the vowel symbol and the ellipses enclose ± 2 standard deviations, rotated along the axis of the first principal component to reflect the correlation between the formants.

First, a MANOVA on the NJ Adult Group revealed no significant effect of Time [$F(2,14)=0.332, p > 0.05, \eta_p^2=0.182$], nor a Time by Vowel interaction [$F(14,208)=0.867, p > 0.05, \eta_p^2=0.062$]. However, a MANOVA on the NJ Child Group showed a significant effect of Time [$F(2,14)=8.945, p < 0.05, \eta_p^2=0.561$] and a Time by Vowel interaction [$F(14,208)=5.408, p < 0.05, \eta_p^2=0.267$]. Eight MANOVAs testing the effect of Time on each Vowel separately for the NJ Child Group showed significant effects for /i/ [$F(2,14)=9.042, p < 0.006, \eta_p^2=0.564$], /ɪ/ [$F(2,14)=11.270, p < 0.006, \eta_p^2=0.617$] and /a/ [$F(2,14)=19.138, p < 0.006, \eta_p^2=0.732$]. However, the effect of Time for the three temporal points of /eɪ/ was not significant [$F(2,14)=0.53, p > 0.006, \eta_p^2=0.090$]. Accordingly, the univariate tests for the F1 and F2 frequencies for each significant MANOVA comparison were conducted. The F1 frequency for /i/ significantly decreased over a year [$F(1,15)=13.834, p < 0.006, \eta_p^2=0.480$], whereas F2 frequency [$F(1,15)=0.180, p > 0.006, \eta_p^2=0.012$] stayed almost the same. /ɪ/ showed a significant increase in F1 [$F(1,15)=13.673, p < 0.006, \eta_p^2=0.477$] as well as in F2 frequency [$F(1,15)=14.048, p < 0.006, \eta_p^2=0.483$]. Also, F1 frequency for /a/ significantly increased [$F(1,15)=17.920, p < 0.006, \eta_p^2=0.544$] but F2 frequency did not change significantly [$F(1,15)=0.013, p > 0.006, \eta_p^2=0.015$]. It should be noted that, although the MANOVA test for /eɪ/ was not significant, F2 frequency in the univariate test significantly decreased [$F(1,15)=11.293, p < 0.006, \eta_p^2=0.430$].

These results indicate that the NJ Adult Group did not change their English vowel production after one year's residence in the US, whereas the NJ Child Group did (see Fig. 1). The NJ Child Group produced /i/ with a lower F1 frequency (i.e., higher in the vowel space), whereas they produced /ɪ/ with higher F1 frequency values (i.e., lower) and higher F2 frequency values (i.e., more front), /a/ with an increased F1 frequency (i.e., lower), and produced /eɪ/ with a trend toward a decreased F2 frequency (i.e., more back).

3.1.2. Duration

The mean duration of eight English vowels produced by NJ Adult and NJ Child Groups were compared across Time. The durational means by group and vowel at Time 1 are reported in Table 2. An ANOVA on the NJ Adult Group revealed no significant effect of Time [$F(1,15)=4.583, p > 0.05, \eta_p^2=0.184$], nor a Time by Vowel interaction [$F(7,9)=0.415, p > 0.05, \eta_p^2=0.244$]. Also, no significant change was observed in the NJ Child Group across Time [$F(1,15)=0.735, p > 0.05, \eta_p^2=0.047$] and the Time by Vowel interaction [$F(3,13)=0.59, p > 0.05, \eta_p^2=0.072$] was not significant. The lack of a Time effect or interaction suggests that, unlike spectral quality, neither the NJ Adult nor the NJ Child Group made changes in duration across Time in any vowel.

3.2. The effect of time on NE Children's production

As noted in the previous section, the NJ children differed in vowel production between Time 1 and Time 2. Accordingly, the NE children's vowel productions across Time were compared in this section (see Fig. 1(c) and (d)). The aim of this analysis was to determine whether the significant differences between Time 1 and Time 2 in the NJ children's production reflected phonetic learning, that is, a closer approximation to the production of the NE speakers, or instead resulted from changes in the size of the

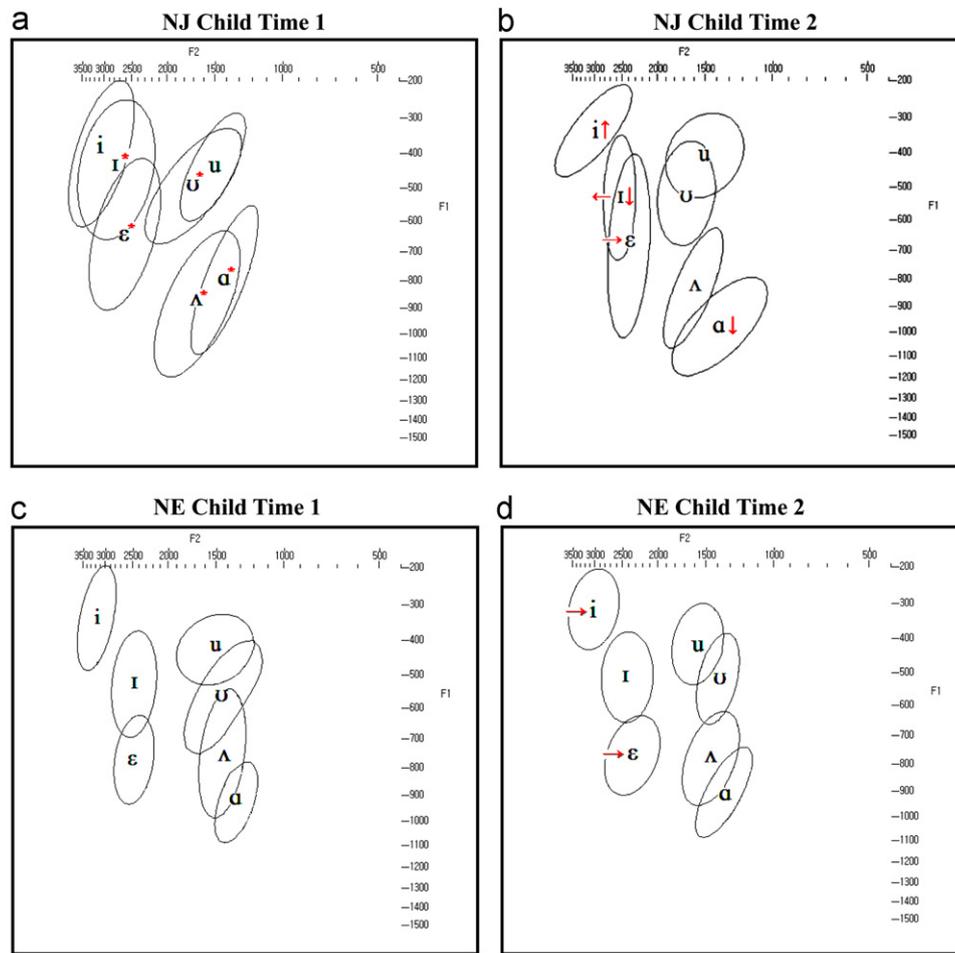


Fig. 1. F1 and F2 frequencies of seven English vowels produced by NJ Child (a) and NE Child (c) Groups at Time 1 and by NJ Child (b) and NE Child (d) Groups at Time 2 are shown. The NJ children's vowels that were significantly different from the NE children's (panel (a) vs. (c) and (b) vs. (d)) are indicated with an asterisk and the vowels that differed significantly between Time 1 and Time 2 (panel (a) vs. (b) and (c) vs. (d)) are marked with an arrow pointing in the direction of the movement.

children's vocal tracts over the course of one year, which could possibly affect vowel frequencies in the absence of learning.

Developmental changes in the NE children were assessed through a comparison of their Time 1 and Time 2 productions. A MANOVA on the NE Child Group showed no significant effect of Time [$F(2,14)=3.538$, $p > 0.05$, $\eta_p^2=0.336$], but showed a significant Time by Vowel interaction [$F(14,208)=3.039$, $p < 0.05$, $\eta_p^2=0.170$]. Eight MANOVAs testing the effect of Time on each Vowel separately showed significant effects for / ϵ / [$F(2,14)=9.189$, $p < 0.006$, $\eta_p^2=0.568$]. The univariate tests for / ϵ / returned a significant decrease in F2 frequency [$F(1, 15)=14.286$, $p < 0.006$, $\eta_p^2=0.488$]. Although not significantly different in the MANOVA tests, the relevant univariate tests also showed that the F2 frequency for / i / [$F(1, 15)=11.727$, $p < 0.006$, $\eta_p^2=0.439$] and for / e / at the 1/4 point [$F(1, 31)=9.632$, $p < 0.006$, $\eta_p^2=0.237$] and 3/4 point [$F(2, 30)=8.185$, $p < 0.006$, $\eta_p^2=0.353$] were significantly decreased at Time 2. Thus, the front vowels / i /, / e / and / ϵ / showed significantly decreased F2 frequencies at Time 2 in the NE Child Group. In Section 3.1, however, the NJ children's production only exhibited an F2 decrease for / ϵ /. Different from the NE Child Group, the NJ Child Group also exhibited changes in F1 for / i /, / I / and / a / and an increase in F2 for / I /. Many studies have shown significant decreases of F2 frequencies for front vowels with increase of age (Busby & Plant, 1995; Lee, Potamianos, & Narayanan, 1999; Vorperian & Kent, 2007). We reasoned, therefore, that the decreased F2 frequency found in the NE Child Group, as well as that of / ϵ / for the NJ Child Group,

at Time 2 was probably due to developmental effects. That is, the changes resulted from an increase in size of the children's oral cavities. The F1 changes in / i /, / I /, / a / in the NJ Child Group, however, were likely due to L2 learning.

3.3. Comparisons between NJ and NE English vowel production

3.3.1. Formant frequency

To examine the native-likeness of the English vowels produced by the NJ Adult and Child Groups, comparisons with the NE Groups were made. As the NJ Adult Group's productions at Time 1 and Time 2 were not significantly different, only Time 1 vowels were compared for the Adult Groups. Fig. 1 displays the F1 and F2 frequency distributions for the NJ and NE Child groups at Time 1 and Time 2 and Fig. 2 displays the F1 and F2 frequency distributions for the NJ and NE Adult groups at Time 1.

First, a MANOVA on the F1 and F2 frequencies of English vowels produced by the NJ Adult and NE Adult Groups at Time 1 revealed a significant effect of Group [$F(2, 29)=5.827$, $p < 0.05$, $\eta_p^2=0.287$], as well as an interaction of Group and Vowel [$F(14, 17)=13.461$, $p < 0.05$, $\eta_p^2=0.917$], which indicated that the effect of Group was significant for some vowels. Results from the MANOVA on each vowel found that some vowels were not different for the two groups: / i / [$F(2,29)=0.240$, $p > 0.006$, $\eta_p^2=0.016$], / ϵ / [$F(2,29)=0.959$, $p > 0.006$, $\eta_p^2=0.244$], / a / [$F(2,29)=5.793$,

$p > 0.006$, $\eta_p^2 = 0.286$], /u/ [$F(2,29) = 1.869$, $p > 0.006$, $\eta_p^2 = 0.114$] and all three points of /eɪ/, (1/4) [$F(2,29) = 0.815$, $p > 0.006$, $\eta_p^2 = 0.053$], (1/2) [$F(2,29) = 0.327$, $p > 0.006$, $\eta_p^2 = 0.022$], (3/4) [$F(2,29) = 0.175$,

$p > 0.006$, $\eta_p^2 = 0.012$]. A significant effect of Group was obtained for three vowels, however: /ɪ/ [$F(2,29) = 13.899$, $p < 0.006$, $\eta_p^2 = 0.489$], /ʌ/ [$F(2,29) = 11.349$, $p < 0.006$, $\eta_p^2 = 0.439$], and /ʊ/ [$F(2,29) = 26.564$, $p < 0.006$, $\eta_p^2 = 0.647$]. The univariate tests showed that the F2 frequency for /ɪ/ [$F(1,30) = 15.423$, $p < 0.006$, $\eta_p^2 = 0.340$] was significantly higher in NJ adults' production than in NE adults' production. Both F1 and F2 frequencies were significantly higher for /ʌ/ (F1 [$F(1,30) = 23.164$, $p < 0.006$, $\eta_p^2 = 0.436$] and F2 [$F(1,30) = 16.980$, $p < 0.006$, $\eta_p^2 = 0.361$]) and F1 was lower and F2 higher for /ʊ/ (F1 [$F(1,30) = 14.044$, $p < 0.006$, $\eta_p^2 = 0.319$] and F2 [$F(1,30) = 13.011$, $p < 0.006$, $\eta_p^2 = 0.303$]) in NJ adults' production.

These results indicate that the NJ Adult Group produced /ɪ/ with a higher F2 (i.e., further front in the vowel space), /ʌ/ with higher F1 and F2 frequencies (lower and further front in the vowel space) and /ʊ/ with lower F1 and higher F2 frequencies (higher and further front in the vowel space) than the NE Adult Group (see Fig. 2). Although not significant in the MANOVA, the production of /a/ by the NJ Adult Group revealed a significantly higher F2 frequency [$F(1,30) = 9.862$, $p < 0.006$, $\eta_p^2 = 0.247$] than the NE Adult Group in the univariate tests. Of the eight English vowels, four English vowels /ɪ/, /a/, /ʌ/ and /ʊ/, were produced with significant differences between the NJ and NE adults' production.

Next, NJ and NE children's productions were compared in separate analyses at Time 1 and Time 2. Fig. 1 presents the F1 and F2 frequency of seven English vowels produced by the NJ Child Group vs. NE Child Group at Time 1 and Time 2. A MANOVA on the NJ and NE Child Groups at Time 1 revealed a significant effect of Group [$F(2,29) = 8.827$, $p < 0.05$, $\eta_p^2 = 0.432$], and a Vowel by Group interaction [$F(14,17) = 6.275$, $p < 0.05$, $\eta_p^2 = 0.838$]. In the children's productions, five new vowels returned a significant effect of Group: /ɪ/ [$F(2,29) = 8.906$, $p < 0.006$, $\eta_p^2 = 0.381$], /ɛ/ [$F(2,29) = 8.435$, $p < 0.006$, $\eta_p^2 = 0.368$], /a/ [$F(2,29) = 21.480$, $p < 0.006$, $\eta_p^2 = 0.597$], /ʌ/ [$F(2,29) = 7.869$, $p < 0.006$, $\eta_p^2 = 0.352$], and /ʊ/ [$F(2,29) = 26.294$, $p < 0.006$, $\eta_p^2 = 0.645$]. Three other vowels were not significantly different across Group: /i/ [$F(2,29) = 4.966$, $p > 0.006$, $\eta_p^2 = 0.266$], /u/ [$F(2,29) = 0.082$, $p > 0.006$, $\eta_p^2 = 0.057$], and the three points of /eɪ/, (1/4) [$F(2,29) = 2.176$, $p > 0.006$, $\eta_p^2 = 0.130$], (1/2) [$F(2,29) = 0.556$, $p > 0.006$, $\eta_p^2 = 0.037$], (3/4) [$F(2,29) = 0.902$, $p > 0.006$, $\eta_p^2 = 0.059$].

The univariate tests at Time 1 showed that the F2 frequency in the NJ Child Group was significantly higher than that in the NE Child Group for /ɪ/ [$F(1,30) = 8.667$, $p = 0.006$, $\eta_p^2 = 0.224$], /ʌ/ [$F(1,30) = 14.956$, $p < 0.006$, $\eta_p^2 = 0.333$] and /ʊ/ [$F(1,30) = 10.550$, $p < 0.006$, $\eta_p^2 = 0.260$], while lower F1 frequency for /ɛ/ [$F(1,30) = 7.891$, $p < 0.006$, $\eta_p^2 = 0.208$] and /a/ [$F(1,30) = 11.255$, $p < 0.006$, $\eta_p^2 = 0.273$] and the higher F1 frequency for /ʌ/ [$F(1,30) = 8.862$, $p = 0.006$, $\eta_p^2 = 0.228$] were found in the NJ children's production. These results indicate that most of the lax vowels produced by NJ children at Time 1 were higher and further fronted in the vowel space than those produced by NE children except for the new English vowel /ʌ/, which was lower and more centralized than NE children's /ʌ/. A MANOVA on NJ and NE Child Groups at Time 2 showed no significant main effect of Group [$F(2,29) = 0.653$, $p > 0.05$, $\eta_p^2 = 0.025$] nor a Group by Vowel interaction [$F(14,17) = 1.583$, $p > 0.05$, $\eta_p^2 = 0.024$], which indicated that

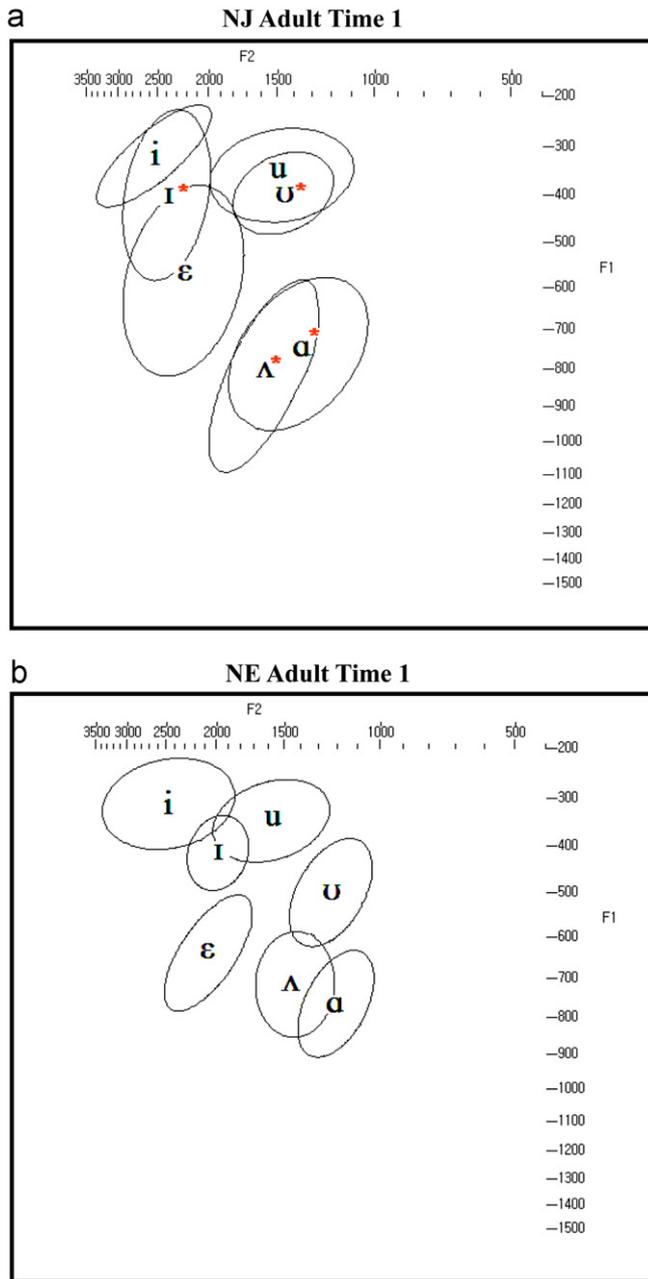


Fig. 2. F1 and F2 frequencies of seven English vowels produced by NJ Adult (a) and NE Adult (b) Groups at Time 1 are shown. The NJ adults' vowels that were significantly different from the NE adults' vowels (panel (a) vs. (b)) are indicated with an asterisk.

Table 2 Mean duration of eight English vowels in ms. produced by adults (top half) and children (bottom half) at Time 1 are shown with standard deviation in parentheses. Significantly different mean values between native English and Japanese Groups are marked in bold.

	i	ɪ	eɪ	ɛ	ɑ	ʌ	u	ʊ
NE adults	143(25)	129(32)	227(40)	177(30)	228(42)	197(33)	225(62)	122(28)
NJ adults	207(34)	136(61)	260(40)	188(40)	190(37)	204(45)	234(63)	141(29)
NE children	152(28)	137(25)	255(53)	192(36)	237(39)	216(56)	279(58)	128(27)
NJ children	173(56)	134(43)	243(65)	174(48)	171(59)	200(58)	223(62)	141(38)

NJ children's productions were not statistically different from NE children's productions at Time 2.

To summarize, the NJ and NE Adult Groups showed differences in producing /ɪ/, /ʌ/, /ɑ/ and /ʊ/ at Time 1, whereas the NJ Adult Group did not show any changes in English vowel production between Time 1 and Time 2. NJ children, on the other hand, produced /ɪ/, /ɛ/, /ʌ/, /ɑ/ and /ʊ/ differently from NE children at Time 1 but those differences were no longer statistically significant at Time 2. Overall, more vowels produced by NJ children than adults differed significantly from those of age-matched controls at Time 1 and thus were not "native-like" which suggests that NJ adults were able to outperform NJ children at the beginning stages of L2 learning but not after a year's time.

3.3.2. Duration

As neither the NJ Adult Group nor the Child Group exhibited changes in vowel duration across time, only the Time 1 vowel productions for the NJ and NE Groups were compared (see Table 2). An ANOVA on the mean duration of English vowels produced by the NJ and NE Adult Groups at Time 1 revealed no significant effect of Group [$F(1,30)=2.098$, $p > 0.05$, $\eta_p^2=0.069$], but a significant interaction of Vowel and Group [$F(7,24)=9.858$, $p < 0.05$, $\eta_p^2=0.742$], which indicates that some vowels were significantly different across Groups. Results from the ANOVA on each vowel showed that /i/ [$F(1,30)=35.746$, $p < 0.006$, $\eta_p^2=0.544$] and /ɑ/ [$F(1,30)=7.398$, $p < 0.006$, $\eta_p^2=0.198$] were significantly different, whereas /ɪ/ [$F(1,30)=0.176$, $p > 0.006$, $\eta_p^2=0.006$], /ɛ/ [$F(1,30)=0.787$, $p > 0.006$, $\eta_p^2=0.026$], /eɪ/ [$F(1,30)=5.566$, $p > 0.006$, $\eta_p^2=0.156$], /ʌ/ [$F(1,30)=0.247$, $p > 0.006$, $\eta_p^2=0.008$], /ʊ/ [$F(1,30)=3.668$, $p > 0.006$, $\eta_p^2=0.109$] and /u/ [$F(1,30)=0.180$, $p > 0.006$, $\eta_p^2=0.006$] were not significantly different between the two Adult Groups. As shown in Table 2, NJ adults' production for /i/ was longer and /ɑ/ was shorter than those produced by NE adults. As far as the tense–lax pairs are concerned (i.e., /i/–/ɪ/, /eɪ/–/ɛ/, /u/–/ʊ/ and /ɑ/–/ʌ/), NJ adults produced most of the tense vowels with greater duration than lax vowels except for the /ɑ/–/ʌ/ pair, which showed the reverse pattern.

An ANOVA on NJ and NE Child Groups at Time 1 was conducted as well. There was no significant effect of Group [$F(1,30)=1.333$, $p > 0.05$, $\eta_p^2=0.041$]. However, a Vowel by Group interaction [$F(7,24)=14.860$, $p < 0.05$, $\eta_p^2=0.813$] was significant. In the ANOVAs conducted on each vowel separately, /ɑ/ [$F(1,30)=13.564$, $p < 0.013$, $\eta_p^2=0.311$] was significantly different but the rest of the vowels were not significantly different across Group (/i/ [$F(1,30)=1.828$, $p > 0.013$, $\eta_p^2=0.047$] and /ɪ/ [$F(1,30)=0.057$, $p > 0.013$, $\eta_p^2=0.002$], /ɛ/ [$F(1,30)=1.502$, $p > 0.013$, $\eta_p^2=0.057$], /eɪ/ [$F(1,30)=0.399$, $p > 0.013$, $\eta_p^2=0.014$], /ʌ/ [$F(1,30)=0.790$, $p > 0.013$, $\eta_p^2=0.027$], /ʊ/ [$F(1,30)=0.279$, $p > 0.013$, $\eta_p^2=0.010$] and /u/ [$F(1,30)=4.842$, $p > 0.013$, $\eta_p^2=0.143$]). The NJ children's /ɑ/ was shorter than the NE children's, which led to /ɑ/ having a shorter duration than its counterpart /ʌ/, similar to the NJ Adult Group. Both the NJ Adult and Child Groups produced English tense and lax vowels with native-like duration except for the /ɑ/–/ʌ/ pair.

4. Experiment 2: Japanese vowel production

In this section, we explore whether learning English influenced how the NJ adults or children produced vowels in their L1. Specifically we examined whether the acquisition of English vowels located in neighboring areas of the native Japanese vowel space triggered either a merger of English and Japanese vowels or a dispersion between English and Japanese vowels. Japanese vowels were examined for an effect of Time and then the

English–Japanese vowel pairs were analyzed in order to investigate whether Japanese vowels changed in relation to the neighboring English vowels.

4.1. Method

4.1.1. Participants

The same 16 Native Japanese (NJ) adults and 16 NJ children who participated in Experiment 1 also participated in this study.

4.1.2. Stimuli

The Japanese words, /rakuda/ 'camel' and /kirin/ 'giraffe' were produced to elicit the Japanese vowels, /a/, /u/ and /i/. The two Japanese words were part of a larger list.

4.1.3. Procedure

The same procedure used for the English words was also used for the Japanese words. A picture depicting the target word, along with a written presentation of the word in Hiragana script, was displayed as a prompt. The Japanese words were presented in a random order to the participants and the words were elicited 3 times. The experimenter provided prerecorded productions of the corresponding word at the first presentation and the participants were asked to repeat the words. Only the second and third productions, which were recorded without the auditory prompts, were analyzed.

4.1.4. Formant frequencies

A total of 256 tokens (2 words \times 2 repetitions \times 2 testing times \times 32 participants) for the Japanese vowel productions were analyzed using Scicon PCquirer acoustic analysis software. No occurrence of devoiced vowels or missing vowels was observed. The same measurement procedure that was used for the English vowels was also used for the Japanese vowels; First and second formant frequencies were measured at the temporal midpoint of each vowel.

4.1.5. Statistical analyses

Japanese vowels were selected based on the English vowels that were significantly different across Time to investigate whether the acquisition of English vowels would affect the Japanese vowels. In order to examine whether there was any difference between Time 1 and Time 2 Japanese vowel production for the NJ Child or Adult Groups, MANOVAs were conducted. The dependent variables were F1 and F2 frequencies and the independent variables were Time and Vowel. The alpha level was adjusted to 0.017 for multiple comparisons. The univariate tests for F1 and F2 are reported for each significant MANOVA comparison.

The relationship of the Japanese and English vowels was also investigated in a subsequent analysis. Separate MANOVAs for the NJ Adult and Child Group on the productions of neighboring English (/i/, /a, ʌ/, /u/) and Japanese vowel pairs (/i/, /a/, /u/) were compared to determine whether they were produced distinctly from each other at Time 1 or Time 2. The alpha level was adjusted to 0.013 for 4 English–Japanese vowel pair comparisons.

4.2. Results

In Experiment 1, the NJ Child Group showed significant changes in the values of F1 frequency for English /i/, /ɪ/ and /ɑ/ in a year's time (see Fig. 1). A comparison with the NE Child Group indicated that these changes were not likely due to developmental effects. Although NJ children's /ʊ/ did not change significantly across Time, the results from Section 3.3 indicated that NJ children were able to produce /ʊ/ with more native-likeness at

Time 2. Thus, Japanese /i/, /a/ and /u/, which are located near English /i/, /ɪ/, /ɑ/ and /ʊ/ in the vowel space, were examined in order to observe any changes that might have occurred in Japanese vowel production as a consequence of the acquisition of English vowels. Fig. 3 presents the NJ Child Group's production of three Japanese vowels at Time 1 and Time 2.

First, to examine the production of the three Japanese vowels across time, a MANOVA was conducted for each NJ Group separately. The MANOVA on the NJ Adult Group showed no significant effect of Time [$F(2,14)=0.048$, $p > 0.05$, $\eta_p^2=0.007$] nor was there an interaction of Time and Vowel [$F(4,14)=1.708$, $p > 0.05$, $\eta_p^2=0.363$]. On the other hand, the MANOVA on the NJ Child Group showed a significant effect of Time [$F(2,14)=9.112$, $p < 0.05$, $\eta_p^2=0.566$] but not a significant Time by Vowel interaction [$F(4,14)=1.588$, $p > 0.05$, $\eta_p^2=0.346$]. In planned comparisons, the effect of Time was examined for each vowel produced by the NJ Child Group. The Time effect for /i/ [$F(2,14)=4.161$, $p < 0.017$, $\eta_p^2=0.373$] and /a/ [$F(2,14)=11.129$, $p < 0.017$,

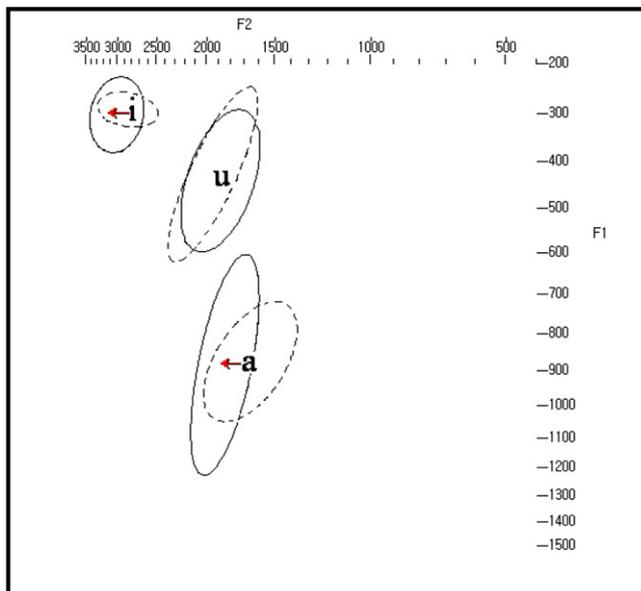


Fig. 3. F1 and F2 frequencies of three Japanese vowels produced by the NJ Child Group at Time 1 and Time 2 vowels are shown (Japanese vowels produced at Time 1 are marked with a dotted line). The significant changes in F2 frequencies are marked with an arrow pointing the direction of the movement.

$\eta_p^2=0.614$] was significant, but it was not for /u/ [$F(2,14)=2.608$, $p > 0.017$, $\eta_p^2=0.187$]. More specifically, univariate tests showed significant increases in the F2 frequency for the Japanese vowels /i/ [$F(1,15)=8.538$, $p < 0.017$, $\eta_p^2=0.363$] and /a/ [$F(1,15)=16.606$, $p < 0.017$, $\eta_p^2=0.525$]. As shown in Fig. 3, the Japanese vowels /i/ and /a/ came to be produced more forward in the vowel space for the NJ Child Group over the time span of the study.

Second, the relationship between the Japanese and neighboring English vowels was investigated for the two groups. The Japanese vowels /i/, /a/ and /u/ were compared to the English vowels /i/, /a, ʌ/ and /u/, respectively, at Time 1 and Time 2 (see Figs. 4 and 5). For the NJ Adult Group at Time 1, the comparisons revealed no difference between the English–Japanese vowel pairs /i/–/i/ [$F(2,14)=3.172$, $p > 0.013$, $\eta_p^2=0.276$], /ʌ/–/a/ [$F(2,14)=0.110$, $p > 0.013$, $\eta_p^2=0.025$] or /u/–/u/ [$F(2,14)=0.289$, $p > 0.013$, $\eta_p^2=0.038$]. However, there was a difference between the English–Japanese /a/–/a/ vowel pair [$F(2,14)=10.476$, $p < 0.013$, $\eta_p^2=0.599$]. F2 was significantly lower for English /a/ than Japanese /a/ [$F(1,15)=11.581$, $p < 0.013$, $\eta_p^2=0.436$]. At Time 2, none of the English–Japanese /i/–/i/ [$F(2,14)=0.875$, $p > 0.013$, $\eta_p^2=0.089$], /ʌ/–/a/ [$F(2,14)=0.081$, $p > 0.013$, $\eta_p^2=0.011$], /u/–/u/ [$F(2,14)=3.067$, $p > 0.013$, $\eta_p^2=0.248$] or /a/–/a/ vowel pairs [$F(2,14)=2.533$, $p > 0.013$, $\eta_p^2=0.213$] were significantly different. These results indicate that these English and Japanese vowel pairs were not produced distinctly, suggesting merged categories for Japanese and English and /i/–/i/, /u/–/u/ as well as /a/–/a/–/ʌ/ in terms of formant frequency (see Fig. 4).

However, for the NJ Child Group, some L1–L2 vowel pairs moved closer to each other and some vowels moved further away from one another in the vowel space at Time 2, as shown in Fig. 5. At Time 1, the English–Japanese vowel pair /i/–/i/ was significantly different [$F(2,14)=7.789$, $p < 0.013$, $\eta_p^2=0.527$] in both F1 [$F(1,15)=14.811$, $p < 0.013$, $\eta_p^2=0.497$] and F2 frequencies [$F(1,15)=9.556$, $p < 0.013$, $\eta_p^2=0.389$]. The English–Japanese vowel pair /a/–/a/ [$F(2,14)=23.442$, $p < 0.013$, $\eta_p^2=0.770$] was also significantly different in F1 [$F(1,15)=9.798$, $p < 0.013$, $\eta_p^2=0.395$] and F2 frequencies [$F(1,15)=48.454$, $p < 0.013$, $\eta_p^2=0.764$]. A significant difference was also found for the /u/–/u/ vowel pair [$F(2,14)=31.271$, $p < 0.013$, $\eta_p^2=0.817$], but only in F2 frequency [$F(1,15)=56.587$, $p < 0.013$, $\eta_p^2=0.791$]. Examining Fig. 5, one will notice that English /ʌ/ produced by the NJ Child Group at Time 1 overlapped with Japanese /a/. A MANOVA showed that the two vowels were not significantly different from each other at Time 1 [$F(2,14)=0.095$, $p > 0.013$, $\eta_p^2=0.013$].

At Time 2, English /i/ and Japanese /i/ were not significantly different [$F(2,14)=2.252$, $p > 0.013$, $\eta_p^2=0.243$]. The assimilation

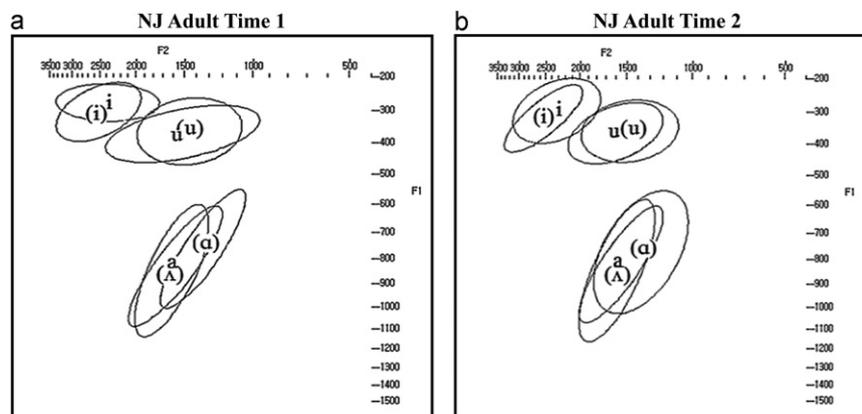


Fig. 4. F1 and F2 frequencies of three Japanese vowels and four similar English vowels produced by the NJ Adult Group at Time 1 (a) and at Time 2 (b) are shown (English vowels are marked with parentheses).

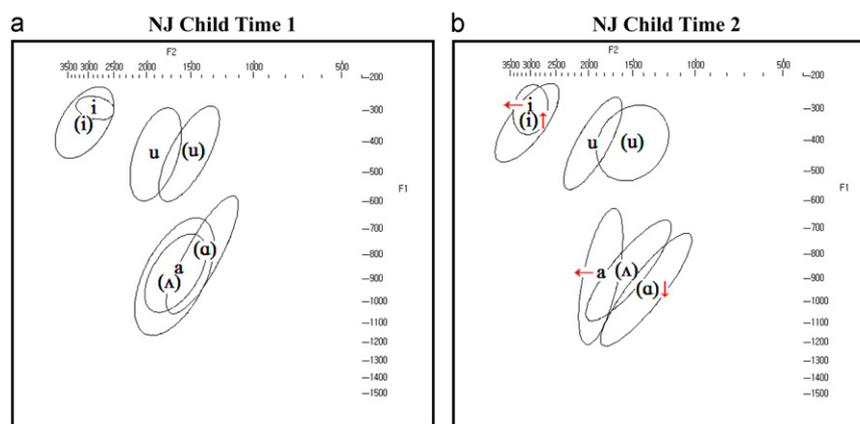


Fig. 5. F1 and F2 frequencies of three Japanese vowels and four similar English vowels produced by the NJ Child Group at Time 1 (a) and at Time 2 (b) are shown (English vowels are marked with parentheses).

occurred as a result of the lowering of F1 frequency for English /i/ and the raising of F2 frequency for Japanese /i/. However, the English–Japanese vowel pair /a/–/a/ was significantly different [$F(2,14)=99.743$, $p < 0.013$, $\eta_p^2=0.934$] in F2 frequency [$F(1,15)=114.811$, $p < 0.013$, $\eta_p^2=0.884$]. English /u/ and Japanese /u/ were also significantly different [$F(2,14)=24.655$, $p < 0.013$, $\eta_p^2=0.779$] in F2 frequency [$F(1,15)=44.798$, $p < 0.013$, $\eta_p^2=0.749$]. The English /ʌ/ and Japanese /a/, which were overlapping at Time 1, were significantly different [$F(2,14)=23.643$, $p < 0.013$, $\eta_p^2=0.772$] in F2 frequency [$F(1,15)=49.949$, $p < 0.013$, $\eta_p^2=0.769$] at Time 2. As shown in Fig. 5, the two vowels separated due to the increased F2 frequency for Japanese /a/.

5. General discussion

The results obtained here agreed with the findings of Aoyama et al. (2004)'s study on consonants in showing that children learned faster with higher accuracy than adults in a year's time. This was especially so with the “new” English vowels. More specifically, NJ children's production for English /i/, /ɪ/, /ɛ/ and /a/ changed significantly with regard to spectral quality between Time 1 and Time 2, whereas the NJ adults showed no significant change across Time. Considering the developmental changes found for NE children's vowels, specifically decreased F2 frequency for front vowels, we considered it likely that the NJ children's decreased F2 frequency for /ɛ/ was also a developmental effect. However, the changes found for /i/, /ɪ/ and /a/ were likely reflections of English vowel acquisition. The differential changes over a year's time for the NJ Child and Adult Groups' English vowel productions indicated that children are more flexible and rapid than adults in learning new vowels. Note that the variation for the NJ Child Group's vowels (see ellipse size in Fig. 1) is a much smaller at Time 2 compared to Time 1, indicating less variance in vowel production.

Following an examination of the Time effect, the NJ and NE Groups were compared to examine whether the changes observed in the NJ Child Group's vowel production resulted in more native-like English vowels at Time 2. We predicted that the NJ Child Group would show a later start than the NJ Adult Group but were more likely to produce new English vowels in a native-like manner at Time 2. The result agreed with the findings of Snow and Hoefnagel-Hoehle (1978) and Aoyama et al. (2004) in that the NJ Adult Group slightly outperformed the NJ Child Group at Time 1, differing from the NE Adult Group by four vowels /ɪ/, /ʌ/, /a/ and /ʊ/, compared to the five vowels for the NJ Child Group /ɪ/, /ʌ/, /a/, /ʊ/ and /ɛ/. At Time 2, the NE Adult Group's production of

these vowels had not changed from Time 1. Although NJ adults had the benefit of cross-linguistically similar vowels at an initial stage of L2 learning, they showed more difficulty acquiring English vowels that had no apparent counterpart in Japanese (see Baker et al., 2008). However, the NJ Child Group demonstrated learning, producing all eight English vowels with no statistical differences in F1 or F2 frequencies compared to the NE Child Group at Time 2. The result that NJ children, but not NJ adults, were able to produce the English vowels in a more native-like manner at Time 2 than at Time 1 indicated that NJ children were capable of creating new English vowel categories independent from the Japanese categories in a year's time.

As predicted on the basis of previous cross-language mapping studies (Frieda & Nozawa, 2007; Strange et al., 1998), the new English lax vowels were produced with less accuracy overall at Time 1 than the similar English tense vowels. Along with the lax vowels, the production of English tense vowel /a/ was significantly differently from the NE Groups at Time 1. This may be due to an assimilation of the English vowel to the Japanese vowel /a/, which has a higher F2 (Shibatani, 1990). Although it seems that the NJ Adult Group was partially able to distinguish English /a/ from Japanese /a/ at Time 1, the English vowel was not produced in a native-like manner.

Other difficulties with the central and back English vowels were also found. Notably, NJ children's production for English /ʌ/ had significantly higher F1 and F2 frequencies, which appeared to share the same F1 and F2 frequency with their Japanese /a/. The overlap between English /ʌ/ and Japanese /a/ was also predicted by the cross-language mapping result reported by Frieda and Nozawa (2007) and Strange et al. (1998), in which English /ʌ/ and /a/ were both mapped to Japanese /a/ or /a:/. Also, the NJ children's English /ʌ/ and /ʊ/ were produced with higher F2 values, which conform to the higher F2 formant values of the corresponding Japanese /a/ and /u/ reported by Tsukada (1999).

Along with formant frequencies, the duration of NJ adults and children's production of English vowels was investigated. According to Nishi, Strange, Akahane-Yamada, Kubo, and Trent-Brown's (2008) perceptual assimilation study, native English listeners assimilated both Japanese long and short vowels /i/, /e/, /a/, /o/ and /u/ to tense English vowels. The authors interpreted the result to indicate that English listeners ignored the temporal differences and only attended to the spectral similarity between Japanese and English vowels. However, since Japanese speakers are more likely to attend to temporal cues between Japanese and English due to the phonemic distinction in Japanese vowels (Ingram & Park, 1997), it was assumed that they would be better at distinguishing English tense and lax vowels in duration than in spectral qualities.

As expected, the duration results showed that most of the English short and tense vowels produced by NJ adults and children were already native-like in terms of durational differences at Time 1 except for the /ʌ/–/ɑ/ distinction, which is known to be difficult for Japanese speakers (see Lambacher et al., 2005). NJ adults produced English /i/ significantly longer than NE adults, while /ɑ/ was produced significantly shorter by both NJ adults and children. In the instance of the NJ Adult Group, it may be the case that NJ adults' /i/ was produced significantly longer than NE adults in order to exaggerate the /i/–/ɪ/ distinction which they failed to make with formant values. The English long tense vowel /ɑ/ was produced with an even shorter duration than its lax counterpart /ʌ/ in both NJ Groups, reflecting Japanese learners' difficulty distinguishing /ɑ/ and /ʌ/ with formant values as well as duration.

Finally, three Japanese vowels /i/, /a/ and /u/ were analyzed across the two times of testing in order to investigate the effect of newly acquired English vowels on native Japanese vowels. The NJ Adult Group showed no difference in Japanese vowel production across time. This was expected, as they showed no changes in L2 English vowel production across time. However, changes in the NJ Child Group were thought to be possible, due to the acquisition of the English /i/, /ɪ/, /ɑ/, /ʌ/ and /ʊ/ vowels.

The NJ Child Group showed a significant increase in F2 frequency for Japanese /i/ and /a/ at Time 2. As decreases in F2 frequency, which were detected in the NE Child Group's production across Time, have been reported to occur with increase in age (Busby & Plant, 1995; Lee et al., 1999; Vorperian & Kent, 2007), the increases in F2 found for the NJ Child Group are not likely to be developmental. From the changes observed in the NJ children's Japanese vowel production, we can infer that the significant decrease in F1 values for their English /i/ (see Fig. 5) was motivated by the category assimilation of English /i/ and Japanese /i/, which seemed to show a case of "merger" in production (Flege et al., 2003). The assimilation of English and Japanese /i/ also supports the claim that English /i/ and Japanese /i/ are the most perceptually similar category among all similar vowel pairs (Strange et al., 1998). Because English /i/ is produced with slightly lower F1 and higher F2 frequency than Japanese /i/ (Tsukada, 1999), it may be possible that /i/ in both languages moved to a middle-ground and was merged into a single category. This bidirectional movement between English and Japanese vowels in the NJ Child Group production suggests that younger learners' L1 system is more malleable, and therefore may exhibit greater effects of L2 on the L1 system than older learners.

There were also two types of category dissimilation observed in this study. First, Japanese children's production of English /ʌ/ and Japanese /a/ overlapped at Time 1, but the Japanese /a/ moved to a more central location at Time 2, as NJ children became more accurate in producing English /ʌ/ and /ɑ/. In other words, Japanese /a/ moved further front in the vowel space as result of acquiring new English low vowels. As it has been reported that Japanese central /a/ has higher F2 values than English /ɑ/ (Shibatani, 1990), it was likely that NJ children enhanced the higher F2 of Japanese /a/ to further emphasize the acoustic differences from their English counterparts, which is an example of L1 dissimilation from L2. Second, NJ children's production of English /ʌ/ also moved further away from Japanese /a/ by lowering the F2 value at Time 2. These findings represent the second type of dissimilation, where NJ children's productions of L2 vowels dissimilated from the neighboring L1 vowels (Mack, 1989; Piske, Flege, MacKay, & Meador, 2002). Overall, at Time 2, NJ children produced six different vowel types (i.e., English /ɑ/, /ʌ/, /u/, Japanese /a/, /u/ and English–Japanese /i/) for seven vowel types as a result of both assimilation and dissimilation, whereas the NJ Adult Group showed only assimilation, yielding three

vowel types (i.e., Japanese–English /i/, Japanese–English /a, ɑ, ʌ/, and Japanese–English /u/).

Along with specific changes in the formant frequency of individual English and Japanese vowels, the overall use of vowel space also appeared to change across Time. Tsukada (1999) reported the more dispersed use of vowel space compared to English monolinguals, possibly due to Japanese speakers' varying amount of LOR in the US. In this study, Fig. 1 showed that the overall variance for each vowel is smaller in the NE Child Group than in the NJ Child Group at both Times. That is, although the NJ Child Group's vowel quality refined overall as they gained more experience, each vowel occupied more acoustic space, and therefore more overlap between vowels was detected compared to the NE Child Group's production at Time 2. This may indicate that each vowel in NJ children's production was more variable than the NE children's production. In addition, not only the NJ Child but also the NE Child Group showed tighter distribution of vowels in the vowel space at Time 2 than at Time 1. However, Fig. 4 shows that the NJ Adult Group's variance for English /ɑ/ became even larger at Time 2, diminishing the difference between English /ɑ/ and Japanese /a/ that was observed at Time 1. This finding suggests that more native-like use of the vowel space for accurate production of the intended vowel target in the NJ children may be attributed to both developmental and learning effects.

Overall, our findings on the relationship of English and Japanese vowels are in agreement with Baker and Trofimovich (2005) in that the NJ children's L1–L2 vowel system revealed a bidirectional interaction while the NJ adults only showed a unidirectional L2 to L1 influence. We suggest that in order to fully understand the degree and direction of the L1–L2 interaction, the changes observed in learners' L1 should be described with reference to the changes occurring in their L2 system and vice versa.

Acknowledgments

This study was supported by NIH grant (DC00257) to J.E. Flege while at the University of Alabama at Birmingham. We are grateful to Linda Legault for help with data collection, to Jesse Blackburn-Morrow and Mark Post for help with data analysis, and to Vsevolod Kapatsinski and an anonymous reviewer for comments on the manuscript.

References

- Aoyama, K., Flege, J., Guion, S., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32, 233–250.
- Aoyama, K., Guion, S., Flege, J., 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.
- Asher, J. J., & Garcia, R. (1969). The optimal age to learn a second language. *The Modern Language Journal*, 53, 334–341.
- Baker, W., & Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language & Speech*, 48, 1–27.
- Baker, W., Trofimovich, P., Flege, J., Mack, M., & Halter, R. (2008). Child–adult differences in second-language phonological learning: The role of cross-language similarity. *Language & Speech*, 51, 316–341.
- Best, C. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In C. Goodman, & H. Nusbaum (Eds.), *The development of speech perception* (pp. 167–224). Cambridge: MIT Press.
- Best, C. (1995). A direct realist view of cross-language speech. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 171–204). Baltimore: York Press.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. J. Munro (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13–34). Amsterdam: John Benjamins.

- Busby, P. A., & Plant, G. L. (1995). Formant frequency values of vowels produced by preadolescent boys and girls. *Journal of Acoustical Society of America*, 97, 2603–2606.
- de Leeuw, E., Schmid, M., & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, 13(1), 33–40.
- Flege, J. (1981). The phonological basis of foreign accent. *TESOL Quarterly*, 15, 443–455.
- Flege, J. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47–65.
- Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233–277). Timonium, MD: York Press.
- Flege, J. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217–244). Trier: Wissenschaftlicher Verlag.
- Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer, & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production, differences and similarities* (pp. 319–355). Berlin: Mouton de Gruyter.
- Flege, J., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics*, 34, 153–175.
- Flege, J., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America*, 83, 729–740.
- Flege, J., 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.
- Frieda, E. M., & Nozawa, T. (2007). You are what you eat phonetically: The effect of linguistic experience on the perception of foreign vowels. In O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language learning* (pp. 79–96). Amsterdam: John Benjamins.
- García Mayo, M., & García Lecumberri, M. (Eds.) (2003). Age and the acquisition of English as a foreign language. Clevedon, England: Multilingual Matters.
- 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.
- Ingram, J., & Park, S. G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, 25, 343–370.
- Ioup, G. (1995). Evaluating the need for input enhancement in post-critical period language acquisition. In D. Singleton, & Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 95–123). Clevedon: Multilingual Matters.
- Kang, K.-H., & Guion, S. (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, 1672–1683.
- Kingston, J. (2003). Learning foreign vowels. *Language & Speech*, 46(2–3), 295–349.
- Lambacher, S., Martens, W., Kakehi, K., Marasinghe, C., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics*, 26, 227–247.
- Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America*, 105(3), 1455–1468.
- Mack, M. (1989). Consonant and vowel perception and production: Early English–French bilinguals and English monolinguals. *Perception & Psychophysics*, 46, 187–200.
- Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., & Trent-Brown, S. (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *Journal of the Acoustical Society of America*, 124(1), 576–588.
- Piske, T., Flege, J., MacKay, I., & Meador, D. (2002). The production of English vowels by fluent early and late Italian–English bilinguals. *Phonetica*, 59, 49–71.
- Shibatani, M. (1990). *The languages of Japan*. Cambridge, UK: Cambridge University Press.
- Snow, C. E., & Hoefnagel-Hoehle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, 49, 1114–1128.
- Strange, W., Akane-Yamada, R., Kubo, R., Trent, S. A., Nish, K., & Jenkins, J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, 26, 311–344.
- Tsukada, K. (1999). *An acoustic phonetic analysis of Japanese-accented English*. Unpublished doctoral dissertation, Macquarie University, Sydney, Australia.
- Tsukada, K. (2009). Durational characteristics of English vowels produced by Japanese and Thai second language learners. *Australian Journal of Linguistics*, 29(2), 287–299.
- Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native English adults and children. *Journal of Phonetics*, 33, 263–290.
- 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.