

Perception of American–English Vowels by Early and Late Spanish–English Bilinguals

Miriam Baigorri

Long Island University, Brooklyn, NY, USA

Luca Campanelli

The Graduate Center, CUNY, New York, NY, USA

Haskins Laboratories, New Haven, CT, USA

Erika S. Levy

Teachers College, Columbia University, New York, NY, USA

Abstract

Increasing numbers of Hispanic immigrants are entering the US and learning American–English (AE) as a second language (L2). Previous studies investigating the relationship between AE and Spanish vowels have revealed an advantage for early L2 learners for their accuracy of L2 vowel perception. Replicating and extending such previous research, this study examined the patterns with which early and late Spanish–English bilingual adults assimilated naturally-produced AE vowels to their native vowel inventory and the accuracy with which they discriminated the vowels. Twelve early Spanish–English bilingual, 12 late Spanish–English bilingual, and 10 monolingual listeners performed perceptual-assimilation and categorical-discrimination tasks involving AE /i,ɪ,ɛ,ʌ,æ,ɑ,o/. Early bilinguals demonstrated similar assimilation patterns to late bilinguals. Late bilinguals' discrimination was less accurate than early bilinguals' and AE monolinguals'. Certain contrasts, such as /æ-ɑ/, /ʌ-ɑ/, and /ʌ-æ/, were particularly difficult to discriminate for both bilingual groups. Consistent with previous research, findings suggest that early L2 learning heightens Spanish–English bilinguals' ability to perceive cross-language phonetic differences. However, even early bilinguals' native-vowel system continues to influence their L2 perception.

Keywords: Vowel perception, second language acquisition, bilingualism, American–English, Spanish

1 Introduction

The percentage of immigrants entering the US who are Hispanic has more than doubled (from 22.4% to more than 47.8%) since the 1990s (US Census Bureau, 2006), when several studies of speech perception in Spanish-English bilinguals were performed (e.g., Flege, 1991; Flege, Bohn, & Jang, 1997). These immigrants learn American-English (AE) as a second language (L2). Accurate perception of AE vowels is important for comprehension of the language, as vowels carry a large part of the speech signal (Kewley-Port, Burkle, & Lee, 2007). Because speech sound learning is most effective in early childhood (Flege & MacKay, 2004; Flege, MacKay, & Meador, 1999; Mack, 1989) and may impact social and professional performance (Agha, 2003; Flege, 1988), it is important to better understand the L2 perceptual skills of this large population of bilinguals.

The perceptual interaction between native language and non-native speech sounds is explained through models of cross-language speech perception. Among the most influential models are the Perceptual Assimilation Model (PAM) for naïve listeners (Best, 1995) and the Perceptual Assimilation Model for L2 learners (PAM-L2) (Best & Tyler, 2007). The PAM (Best, 1995) makes predictions about how naïve listeners will assimilate non-native speech sounds to the phonological categories of their native language. The PAM-L2 (Best & Tyler, 2007) is an extension of the PAM and distinguishes four assimilation patterns that occur in L2 learning. This model posits that when both L2 speech sounds in a contrast are perceived as equivalent to two native categories, discrimination will occur with ease. When both L2 speech sounds in a contrast are perceived as equivalent to the same native category, but one is perceived as being more deviant than the other, L2 learners will discriminate the L2 speech sounds well. With continued L2 exposure, the L2 learner may learn the deviant speech sound as a variant of the native language speech sound. When the two L2 speech sounds are perceived as equivalent to the same native category and are both equally good or poor instances of that category, the learner will, at least initially, have difficulty discriminating the L2 speech sounds in the contrast. Lastly, when both L2 speech sounds in a contrast are not perceived as a specific native category, but rather as between native categories, the accuracy in discrimination will depend on the distance in phonological space between the two speech sounds. That is, if each of the L2 speech sounds has similarities to native categories that are distant from one another in phonological space, discrimination of these speech sounds will be relatively easy.

Differences in the native language and L2 vowel inventories may render some L2 vowels more difficult to perceive than others (Best & Tyler, 2007). The AE inventory of 11 vowels (/i/, /ɪ/, /u/, /ʊ/, /e/, /o/, /ɛ/, /ʌ/, /ɔ/, /æ/, /ɑ/) (Strange, Weber, Levy, Shafiro, Hisagi, & Nishi, 2007) is larger than the Spanish inventory of five vowels: /i/, /u/, /e/, /o/, and /a/ (Bradlow, 1995). AE vowels include a corresponding tense and lax vowel (/i/- /ɪ/, /u/- /ʊ/) for each of the high front (/i/) and high back (/u/) vowels found in Spanish. AE has five mid vowels /e/, /o/, /ɛ/, /ʌ/, and /ɔ/, whereas Spanish has two mid vowels /e/ and /o/. Additionally, AE has two low vowels (/æ/, /ɑ/) and Spanish has only one low, central vowel (/a/). Acoustically, AE vowels /i/, /e/, /o/, and /u/ have higher second formant frequencies than their Spanish counterparts, indicating a more fronted lingual position in AE than in Spanish (Bradlow, 1995). It might be expected that learning to perceive vowels of a larger L2 inventory than a listener's native inventory, as would be the case for native

Spanish speakers learning AE, might be more difficult than learning to perceive vowels from a smaller L2 inventory (Bohn & Flege, 1990; Bradlow, 1995; Escudero, 2014; Iverson & Evans, 2007, 2009; McAllister, Flege, & Piske, 2002).

In addition to the similarities and differences in native language and L2 vowel inventories, factors found to affect accuracy of L2 speech perception include age of learning the L2 (Archila-Suerte, Zevin, Bunta, & Hernandez, 2011; Flege & MacKay, 2004; Flege, Schirru, & MacKay, 2003; Flege et al., 1999; Mack, 1989; Piske, Flege, MacKay, & Meador, 2002; Polka & Werker, 1994) and the length of residence (LOR) in the L2 country (Flege, 1991; Flege et al., 1997; Morrison, 2002). The strongest predictor of L2 speech perception is the age at which L2 is learned (Archila-Suerte et al., 2011; Flege & MacKay, 2004; Flege et al., 1999, 2003).

Studies examining early and late bilinguals' L2 vowel perception have shown a strong advantage for early L2 learning on vowel perception accuracy. For example, Flege et al. (1999) found that early Italian-Canadian English bilingual listeners performed more similarly to monolingual Canadian English listeners than to late Italian-Canadian English bilingual listeners in a categorial discrimination task containing Canadian-English and Italian vowels, suggesting that later L2 learning may result in perceptual abilities in the L2 that are not native-like. Additionally, Archila-Suerte et al. (2011) found that early L2 acquisition in Spanish-English bilingual listeners resulted in accurate within- and between-categorization of AE vowels in a similarity judgment task. Conversely, late English exposure resulted in accurate between-categorization of novel L2 vowels for only highly proficient Spanish-English bilingual listeners. Although vowel contrast / Λ - α / posed a challenge for all listeners, these findings revealed an advantage for early L2 learning.

Although the advantage of early L2 learning for perceptual accuracy is well-established, findings conflict regarding whether *early* bilinguals can perceive L2 vowels in a native-like manner. For example, Mack (1989) found early French-English bilingual listeners' perceptual patterns to be different from those of AE monolinguals. Additionally, Pallier, Bosch, and Sebastián-Gallés (1997) found that early exposure to Catalan was not sufficient for Spanish-Catalan bilingual listeners to discriminate Catalan vowel contrasts in a monolingual native-like manner. Furthermore, Sebastián-Gallés and Soto-Faraco's (1999) study of Spanish-Catalan bilinguals with exposure to Catalan since birth and those with exposure since the ages of 3 or 4 further demonstrated the effect of age of L2 acquisition on the identification of Catalan vowels. Vowel perception by the latter group of early bilinguals was not native-like, despite their exposure to Catalan starting at the young age of 3 or 4. In contrast, Flege et al. (1999) found native-like discrimination of Canadian-English vowels by early Italian-English bilingual listeners.

Such inconsistent findings across studies may reflect the participants' different language backgrounds, as well as the use of different experimental procedures and stimuli. For instance, Mack (1989) examined the identification and discrimination of vowels on a synthetic continuum. Flege et al. (1999), on the other hand, conducted a categorial discrimination test with natural tokens of each vowel category.

The LOR in the L2 country has also been examined as a factor impacting the accuracy of L2 vowel perception by late Spanish-English bilingual listeners. Morrison (2002), for example, found that Spanish listeners with six months of exposure to an L2 (Canadian-English) identified Canadian-English vowels more accurately than did Spanish lis-

teners with one month of exposure to the L2. Similarly, Flege (1991) investigated the identification of AE vowels /i/, /ɪ/, /ɛ/, and /æ/ by late Spanish-English bilinguals who differed in LOR (4 months vs. 10 years). Listeners were instructed to identify these AE vowels according to their Spanish equivalents by circling one of the five letters ("a", "e", "i", "o", "u") or to use the label "none" if they judged the AE vowel to not be a Spanish vowel. Late bilingual listeners with a longer LOR in the US used the label "none" more often than those with a shorter LOR in the US, suggesting that their greater L2 exposure helped separate the listeners' native and L2 vowel categories.

To our knowledge, no previous study has examined patterns with which Spanish-English bilingual adults discriminate naturally-produced AE vowels in relation to how they assimilate the vowels into their native phonological inventory. Previous studies on vowel perception by Spanish-English bilinguals have focused on *AE vowel discrimination* by *early* and *late Spanish-English bilinguals* (Archila-Suerte et al., 2011) and compared *perceptual assimilation* patterns and English vowel *identification* of *late Spanish-English bilinguals* to *native AE monolinguals* (Flege, 1991; Flege et al., 1997; Morrison, 2002). Other studies have compared early and late *Italian-English* bilinguals' vowel perception (Flege & MacKay, 2004; Flege et al., 1999, 2003). The results of these studies indicate that individuals who began learning their L2 (English) in childhood perceived English vowels more like native English listeners than individuals who began learning English in late adolescence or early adulthood. However, whether L2 vowel perception in early bilinguals, and in particular, in early Spanish-English bilinguals, is distinguishable from that of native listeners requires further investigation (Flege et al., 1999; Mack, 1989).

Additionally, the relationship between AE and Spanish vowels requires more extensive investigation. The relative ease or difficulty in L2 vowel perception depends on the relationship between L2 vowels and native vowels (Best & Tyler, 2007). Most studies investigating Spanish-English bilinguals' L2 vowel perception have focused on the identification of AE, Canadian-English, and Scottish-English synthetic vowels in an /i/ - /ɪ/ continuum (Escudero & Boersma, 2004; Flege et al., 1997; Morrison, 2002). AE /ʌ, æ, ɑ/ require further investigation, given the perceptual challenges these vowels present to non-native listeners (Archila-Suerte et al., 2011).

Moreover, perception of naturally-produced AE vowels by native Spanish listeners has received limited attention (Archila-Suerte et al., 2011; Flege, 1991), with more research examining perception of synthetic Canadian-English, Scottish-English, and British-English vowels (e.g., Escudero Boersma, 2004; Escudero & Chládková, 2010; Morrison, 2002, 2008). Speakers' dialects have been shown to influence listeners' L2 perception. For example, Escudero and Chládková (2010) investigated the perception of Standard Southern British English (SSBE) and AE vowels by monolingual Peruvian Spanish listeners. Listeners assimilated AE and SSBE vowels to different Spanish vowel categories, suggesting a dialect-specific effect on L2 vowel perception. Thus, especially with the large population of native speakers of Spanish in the United States, further research could shed light on this population's perception of AE vowels and suggest appropriate targets for perceptual training.

1.1 The current study

The current study examined the perception of AE vowels in natural speech in the presence of noise by early and late Spanish-English bilinguals, replicating and extending previous

examinations of the relationship between English and Spanish vowels (Archila-Suerte et al., 2011; Flege, 1991). Early and late Spanish-English bilinguals' discrimination of naturally-produced AE vowels was examined, as in Archila-Suerte et al. (2011); however, in the present study, the listeners' discrimination patterns were also examined in relation to their perceptual assimilation of the AE vowels into their native phonological inventory, within the framework of PAM-L2 (Best & Tyler, 2007). Moreover, while Flege (1991) examined late Spanish-English bilinguals' identification of naturally-produced AE vowels /i/, /ɪ/, /ɛ/, and /æ/, the present study investigated both early and late Spanish-English bilinguals' perceptual assimilation of a larger AE vowels set /i/, /ɪ/, /ɛ/, /ʌ/, /æ/, /ɑ/, and /o/. Additionally, the current study provided further data on assimilation patterns in L2 learning, potentially elucidating the "none" category in Flege's (1991) study.

The first experiment investigated whether perceptual assimilation of AE vowels would vary as a function of the Spanish-English bilingual listeners' language background (early vs. late bilingual) and the particular vowel. We hypothesized that perceptual assimilation patterns would differ between early and late bilingual listeners and that late bilingual listeners would accept AE vowels as better instances of their Spanish categories as suggested by previous studies on cross-language speech perception (Levy, 2009a). Certain AE vowels, such as /ʌ/ and /æ/, were expected to be more difficult to categorize than others for both groups of listeners (Levy, Leone, Garcia, & Baigorri, 2010).

In the second experiment, we asked whether discrimination accuracy of AE vowels would vary as a function of the listeners' language background (native AE monolingual vs. early vs. late bilingual) and the particular vowel. We hypothesized that early Spanish-English bilingual listeners would discriminate AE vowels more accurately than late Spanish-English bilingual adult listeners, as suggested by previous research (Flege & MacKay, 2004; Flege et al., 1999, 2003). However, it was unclear whether the early bilinguals would demonstrate native-like discrimination accuracy. We hypothesized that L2 vowels that were assimilated to distinct native vowels would yield high discrimination accuracy. Conversely, we expected low discrimination accuracy for L2 vowels that were assimilated to a single native vowel category.

2 Methods

2.1 Participants

Three groups of listeners participated in the experiment: 10 native monolingual AE adults; 12 "early" Spanish-English bilingual adults; and 12 "late" Spanish-English bilingual adults. Both bilingual groups performed a perceptual assimilation and a categorial discrimination task, whereas the monolingual group performed only the categorial discrimination task. Participants were between the ages of 18 and 48, with the mean age of 28 years (see Table 1 for participant characteristics). Listeners reviewed and signed Institutional Review Board consent and completed a language background questionnaire. All passed a bilateral hearing screening at 20 dB at 500, 1000, 2000, and 4000 Hz and had no reported history of a speech or language disorder.

The native monolingual AE group of listeners was born and raised in an AE-speaking environment in the US and had minimal Spanish experience. The "early" bilingual listeners were born in a Spanish speaking country and immigrated to the US prior to 11 years of

age, at which time they learned AE. These listeners were raised in a monolingual Spanish household and had no English exposure in their native country. "Late" bilinguals were born in a Spanish speaking country and raised in a monolingual Spanish household. They immigrated to the US no earlier than 13 years of age. They reported no AE instruction or interaction with AE speakers with any regularity prior to this age. All of the Spanish–English bilinguals were from Latin American countries (Colombia, Dominican Republic, Ecuador, Mexico, and Peru). They differed in their country of origin, as, unlike across English dialects, vowel production is thought to not differ considerably across Spanish dialects and a study by Morrison and Escudero (2007) found no significant difference in formant values between Peruvian and Castilian dialects. It should be noted, however, that even minimal differences in listeners' native dialects can yield different perceptual patterns, as demonstrated by Bohemian Czech versus Moravian Czech listeners' different assimilation patterns for Dutch vowels (Chládková & Podlipský, 2011), which is consistent with dialect-related perceptual differences found by Escudero and Boersma (2004) and Morrison (2008).

It is important to note that definitions for "early" versus "late" bilinguals vary across studies. The age cutoff selected in the current study was used in previous studies that defined early bilinguals as individuals who acquired English prior to puberty and late bilinguals as individuals who acquired English by the end of or beyond puberty (Johnson & Newport, 1989; Long, 1990; Shi, 2010).

2.2 Stimuli perceptual assimilation

The AE vowels /i/, /ɪ/, /ɛ/, /ʌ/, /æ/, /ɑ/, and /o/ were selected for the perceptual assimilation task to provide opportunities for observing various assimilation type patterns by listeners. The control AE vowel was /o/, expected to be assimilated to Spanish /o/ and easily discriminated from other AE vowels (Flege, 1991).

Natural speech was used in this experiment, with the following rationale: studies on the perception of AE vowels by late Spanish–English bilingual listeners have been performed primarily with synthetic speech stimuli (Flege et al., 1997; Morrison, 2002), with only a handful of studies using natural speech (Archila-Suerte et al., 2011; Flege, 1991; Flege, Munro, & Fox, 1994; Fox, Flege, & Munro, 1995). Yet natural speech may be perceived differently from synthetic speech. For example, Kangas and Allen (1990) found that adult listeners repeated words in their native language (AE) more accurately when stimuli were natural speech tokens than when they were synthesized tokens, suggesting that natural speech may better represent the actual speaking conditions.

In the present study, three female native monolingual AE talkers from the New York regional area were recorded producing the nonsense word /gəbVpə/ in /gəbVpə/ context in the carrier phrase "five /gəbVpə/ this time." Consonants /b/ and /p/ preceded and followed (respectively) the target vowel because these consonants do not involve tongue articulation, thus minimizing any coarticulatory influence on the target vowels (Levy, 2009a; Strange et al., 2007). Nonsense words were selected rather than real words to decrease any lexical effects (Neuman & Hochberg, 1983). A carrier phrase was used rather than vowels or words in isolation because phrases are more representative of everyday speech than are words or vowels in isolation. The talkers had minimal exposure to other languages and no history of speech or language disorders. They passed a bilateral hearing screening at 20dB at 500, 1000, 2000, and 4000 Hz.

Table 1. Participant characteristics (monolingual = MO; early bilingual = EB; late bilingual = LB; and n/a = not applicable).

Participant	Group	Age	Gender	Age of arrival to the US
101	MO	25	F	n/a
102	MO	24	F	n/a
103	MO	25	F	n/a
104	MO	27	F	n/a
105	MO	24	F	n/a
106	MO	24	F	n/a
107	MO	23	F	n/a
108	MO	33	F	n/a
109	MO	25	F	n/a
110	MO	32	F	n/a
201	EB	30	F	6
202	EB	25	F	3
203	EB	18	F	8
204	EB	23	F	8
205	EB	20	F	10
206	EB	19	F	9
207	EB	48	F	6
208	EB	20	F	2
209	EB	24	F	10
210	EB	25	F	2
211	EB	24	F	3
212	EB	34	F	4
301	LB	40	F	17
302	LB	38	F	22
303	LB	28	F	14
304	LB	33	F	21
305	LB	27	F	19
306	LB	23	F	13
307	LB	41	F	26
308	LB	44	F	14
309	LB	39	F	15
310	LB	23	M	14
311	LB	24	F	15
312	LB	39	F	16

Talkers were recorded in a sound-treated booth in the Speech Production and Perception Laboratory at Teachers College, Columbia University. Output was recorded through a Shure (SM58) microphone placed 15 cm from the talker's mouth and passed through a Shure (Prologue 200M) mixer to a Turtle Beach Rivera sound card of a Dell Pentium desktop computer using Soundforge™ 8.0 software, with a sample rate of 22,050 Hz, 16-bit resolution, on a mono channel. The experimenter was in the adjoining room and provided the talker with directions using an intercom. After being familiarized with the written protocol, talkers read 4 lists, producing 10 utterances (Five gəbVpə this time), on each list, which contained stimuli with 9 AE vowel targets (e.g., "Five gabeepa this time" for /i/). They were instructed to read each utterance as if talking to a good friend. Utterances on lists were randomized and the first utterance and the last utterance contained the same target vowel. In order to control for list-final intonation effect, the final utterance was discarded (Strange et al., 2007). The experimenter listened to the recording input via Sennheiser HD 280 pro headphones. If an utterance contained irregular rate, prosody, vocal quality, pronunciation, or noise, the talker was instructed to repeat the utterance. For each vowel, the second and third recording were used as stimuli that were determined by the primary investigator to contain noise or other distraction. Multiple tokens of the utterances were used to obtain information on categorial perception rather than physical discrimination (Gottfried, 1984). Stimuli were entered into the Paradigm software program (Tagliaferri, 2011) for presentation to the listeners.

2.2.1 Categorial discrimination. The stimuli for the categorial discrimination experiment were the same as those presented in the perceptual assimilation task. Noise was added in order to address ceiling effects that could occur in silent environments, as well as for greater ecological validity, that is, to better reflect the noisy communicative contexts in which speech sound discrimination typically takes place. Studies on Spanish-English bilinguals have focused on vowel perception in quiet conditions. Noisy environments can decrease perceptual accuracy of speech, especially of L2 words and sentences (Adachi, Akahane-Yamada, & Ueda, 2006; Broersma & Scharenborg, 2010; Garcia Lecumberri, Cooke, & Cutler, 2010; Ueda, Akahane-Yamada, & Komaki, 2002; von Hapsburg, Champlin, & Shetty, 2004). However, no noise was added to the stimuli for the perceptual assimilation task, as this categorial task was not subject to ceiling effects and was designed to provide as accurate as possible a representation of listeners' vowel assimilation patterns.

Speech-shaped noise yielding -2 dB signal-to-noise ratio (SNR) was thus added to stimuli using the Praat v. 5.2.22 program, based on the findings of a pilot study suggesting a floor effect at -4 dB SNR and its use in previous studies (Bradlow & Alexander, 2007; Rogers, Lister, Febo, Besing, & Abrams, 2006). The following vowel contrasts were tested: /i-ɪ/, /ɪ-ε/, /ʌ-ɑ/, /ʌ-æ/, /æ-ɑ/ and control contrast /ɑ-i/. Experimental contrasts contained vowels that approximated each other in L2 acoustic vowel space and whose segments were in close proximity to native vowels, to provide opportunities for examining a range of L2 learning patterns.

2.3 Procedure

2.3.1 Perceptual assimilation. Stimuli were presented via Sennheiser HD 280 pro headphones. Listeners were presented with the carrier phrase with a nonsense word: "Five /gəbVpə/ this time." They were instructed to "listen to the second vowel sound of the

word (e.g., gabeepa) and determine which Spanish sound is the best example of that sound by choosing one of the following Spanish nonsense words: bapo, bepo, bipo, bopo, bupo." Response options were displayed on the computer monitor and represented the Spanish vowels /a/, /e/, /i/, /o/, and /u/. Listeners then heard the stimulus again and were asked to rate the vowel on a scale from 1–9, with (1) indicating "least Spanish-like" and (9) indicating "most Spanish-like." They were instructed to use the entire spectrum of the scale. A total of 7,128 responses were collected from 24 listeners (297 from each early and late bilingual listener).

2.3.2 Categorical discrimination. For each trial, stimuli were produced by the three different talkers described above, with the order of talkers randomized. Three stimuli were presented in AXB trials via Sennheiser HD 280 pro headphones. For each contrast, half of the "same" vowel as the second (middle) stimulus were presented first, and half were presented third. Participants were instructed to click on "1" if the vowel in the second phrase was the same as the one in the first phrase and click on "3" if the vowel in the second phrase was the same as the one in the third phrase. A total of 7,616 responses were collected from 34 listeners (224 from each monolingual and early and late bilingual listener).

2.4 Analysis

Descriptive and inferential statistical analyses were performed on all data by means of Stata software, version 13 (StataCorp, 2013). To determine whether perceptual assimilation patterns and discrimination accuracy differed as a function of language background and particular vowel in this study, results were analyzed using a mixed effects logistic regression, which has been reported to give relatively reliable and robust results for categorical outcome variables (e.g., the forced-choice variables used in this study) and to be more appropriate than an analysis of variance approach on transformed data (e.g., arcsine square root transformation; Ferguson, 2012; Jaeger, 2008). These models contain both fixed and random effects. In this study, fixed effects included vowel or vowel contrast (vowel for the assimilation data and vowel contrast for the discrimination data), group, and their interaction. Listeners were considered random effects as they are thought of as a random selection of a much larger population. For all analyses, random slopes for the within-subject predictor were examined but never retained, either because of convergence failure or because they did not improve the model fit (as assessed by the likelihood-ratio test) and did not change the fixed effects component of the models significantly. Main effects and interactions were assessed using the likelihood-ratio test, and Bonferroni corrected pairwise comparisons were performed to follow up the significant effects.

A cross-language assimilation overlap method (Levy, 2009b) was used in order to determine whether errors in categorical discrimination were related to the listeners' perceptual assimilation patterns. This method quantifies the extent to which perceptual assimilation for one vowel in a contrast overlaps with the perceptual assimilation of the other vowel in the contrast. An overlap score was calculated by determining the percentage of responses that two L2 speech sounds in a contrast were assimilated to the same native category. Vowel contrasts were ranked according to their overlap scores, which were then compared to the listeners' discrimination accuracy to test the prediction by the PAM-L2 (Best & Tyler, 2007) that perceptual assimilation patterns (i.e., greater overlap in assimilation) predict discrimination accuracy (i.e., poorer discrimination). For example, if two AE vowels were

assimilated to a single native category yielding a high assimilation overlap score and high discrimination errors, while two other vowels assimilated to different categories (yielding a low overlap score) and were discriminated accurately, such patterns would support the PAM-L2.

Table 2. Perceptual assimilation patterns of American-English (AE) vowels to Spanish vowels by early bilingual listeners. Percentages chosen for the Spanish vowel responses selected are presented for each AE vowel stimulus.

AE vowel stimulus	Spanish vowel response				
	/a/	/e/	/i/	/o/	/u/
/i/	0.0	4.1	95.4	0.4	0.2
/ɪ/	0.6	48.5	50.0	0.2	0.7
/ɛ/	4.6	93.5	0.7	0.2	0.9
/æ/	91.7	7.0	0.2	0.9	0.2
/ɑ/	82.6	0.4	0.0	10.9	6.1
/ʌ/	29.4	4.3	0.2	24.4	41.7
/o/	0.9	0.0	0.0	96.0	3.1

Table 3. Perceptual assimilation patterns of American-English (AE) vowels to Spanish vowels by late bilingual listeners. Percentages chosen for the Spanish vowel responses selected are presented for each AE vowel stimulus.

AE vowel stimulus	Spanish vowel response				
	/a/	/e/	/i/	/o/	/u/
/i/	0.2	3.3	96.3	0.0	0.2
/ɪ/	0.2	48.9	50.0	0.9	0.0
/ɛ/	11.3	86.7	0.6	1.5	0.0
/æ/	82.0	13.5	0.6	3.5	0.4
/ɑ/	75.0	1.3	0.2	22.8	0.7
/ʌ/	53.1	6.7	0.7	38.5	0.9
/o/	0.3	1.5	0.0	97.8	0.3

3 Results

3.1 Perceptual assimilation: Language background and particular vowel effects

Our first research question concerned language background effects on perceptual assimilation of AE vowels. The early bilingual listeners demonstrated similar assimilation patterns to late bilingual listeners for all vowels; however, the two groups' goodness ratings for the vowels differed. Tables 2 and 3 show the Spanish vowel (perceptual assimilation) responses selected by early bilingual and late bilingual listeners for the AE vowel stimuli presented. Table 4 displays the modal Spanish vowels chosen by the early and late bilingual listener groups for all AE vowel stimuli presented. The left-hand column lists the AE vowel stimuli, followed by the second column, which lists the overall modal Spanish responses chosen. The column labeled "Mode percentage chosen" indicates the overall percentage of trials for which

that particular Spanish response was chosen by early and late bilingual listeners. The "Median rating" indicates the median of goodness ratings on a scale from 1 (least Spanish-like) to 9 (most Spanish-like) of all of the trials on which bilingual listeners selected the modal response category.

Overall, median goodness ratings for all vowels increased for listeners with a later age of L2 acquisition, suggesting that late bilingual listeners perceived AE vowels as more like their native vowels than early bilingual listeners. Furthermore, the majority of AE vowels (/ɪ/, /æ/, /ɑ/, /ʌ/, /o/) received high median goodness ratings (6, 6, 6, 6, 7, respectively) by late bilingual listeners suggesting that late bilingual listeners accepted AE vowels as good instances of their Spanish category and were possibly less attuned to the differences between these AE and Spanish vowels than early bilingual listeners (median goodness ratings: 3, 4, 4, 3, 6, respectively). Mixed effects regression indicated a significant difference in goodness ratings for early and late bilingual listeners, $z = 1.97$, $p = .049$.

Table 4. Perceptual assimilation of American-English (AE) vowels by early and late Spanish bilingual listeners: Percentages chosen for each modal response (most frequent category chosen) and median goodness ratings (scale from 1–9, with (1) indicating "least Spanish-like" and (9) indicating "most Spanish-like") are presented for each vowel.

AE stimulus	Spanish modal choice	Mode percentage chosen	Median rating
Early bilingual listeners			
/i/	/i/	95	8
/ɪ/	/i/	50	3
/ɛ/	/e/	94	6
/æ/	/a/	92	4
/ɑ/	/a/	83	4
/ʌ/	/u/	42	3
/o/	/o/	96	6
Late bilingual listeners			
/i/	/i/	96	8
/ɪ/	/i/	50	6
/ɛ/	/e/	87	6
/æ/	/a/	82	6
/ɑ/	/a/	75	6
/ʌ/	/a/	53	6
/o/	/o/	98	7

Regarding particular vowel effects, early and late bilingual listeners perceptually assimilated AE vowels to Spanish vowels that are acoustically similar to the AE vowels, as seen in the table 4. For example, early and late bilingual listeners assimilated AE front vowel /ɪ/ to both Spanish front /e/ (early bilinguals: 49%, late bilinguals: 49%) and Spanish front /i/ (early bilinguals: 50%, late bilinguals: 50%) and AE front vowel /ɛ/ to Spanish front /e/ (early bilinguals: 94%, late bilinguals: 87%). For the two groups of listeners, modes were similar to each other for AE vowels /ɪ/ (50%, 50%) and /ɛ/ (94%, 87%), suggesting that perceptual assimilation of these vowels did not vary as a function of language

background. Early and late bilingual listeners assimilated AE vowels /æ/ (early bilinguals: 92%, late bilinguals: 82%) and /ɑ/ (early bilinguals: 83%, late bilinguals: 75%) primarily to Spanish low vowel /a/. Spanish modes for AE vowels /æ/ and /ɑ/ increased as a function of language background, suggesting that earlier age of L2 learning was associated with more stability in vowel representation for these vowels. Early and late bilingual listeners assimilated AE central vowel /ʌ/ to Spanish central and back vowels. For example, early bilingual listeners most often assimilated AE /ʌ/ to Spanish /u/ (42%), Spanish /a/ (30%) and Spanish /o/ (24%), and late bilingual listeners assimilated AE /ʌ/ to Spanish /a/ (53%) and Spanish /o/ (39%). The Spanish modal choice for AE /ʌ/ differed for early and late bilingual listeners (/u/, /a/, respectively). Additionally, AE vowel /ʌ/ received a relatively low modal percentage score (42%, 53%) by early and late bilingual listeners respectively, indicating difficulty or inconsistency in categorizing this vowel.

When early and late bilingual listeners' responses were compared for each AE vowel stimulus presented, mixed effects logistic regression indicated a statistically significant difference in responses for AE vowel /ɛ/, /ʌ/, and /æ/ stimuli. (The complete tables of results are provided as Supplementary Material.) For example, a significant or approaching significant difference between early and late bilingual listeners' responses to AE vowel /ɛ/ was found for their Spanish vowel response /a/, $z = 2.37$, $p = .018$, and for /e/, $z = -2.12$, $p = .034$. Likewise, a significant difference between early and late bilingual listeners' responses to AE /ʌ/ was found for their Spanish vowel response /a/, $z = 2.43$, $p = .015$, and /u/, $z = -3.66$, $p < 0.001$. Lastly, a significant difference or a trend between early and late bilingual listeners' responses to AE vowel /æ/ was found for their Spanish vowel response /a/, $z = -2.6$, $p = .009$, /e/, $z = 1.95$, $p = .051$, and /o/, $z = 1.92$, $p = .054$.

3.2 Categorical discrimination: language background and particular vowel contrast effects

Our second research question asked about language background and specific vowel effects on categorical discrimination. (The complete tables of results can be found in Supplementary Material.) Figure 1 displays the percentage correct for each vowel contrast by each listener group. The AE vowel contrast is along the X-axis and the percentage correct with error bars representing standard error of the mean in percentage is along the Y-axis. Overall, discrimination accuracy was highest for monolingual AE listeners (mean accuracy = 97%), followed by early bilingual (mean accuracy = 80%) and late bilingual listeners (mean accuracy = 66%). Mixed effects logistic regression confirmed a main effect of group, $\chi^2(2) = 45.3$, $p < 0.001$: monolingual listeners performed significantly more accurately than early and late bilingual listeners, $z = -6.24$, $p < 0.001$, and $z = -9.21$, $p < 0.001$, respectively, and early bilingual listeners performed more accurately than late bilingual listeners, $z = -3.45$, $p = .002$. A significant main effect of vowel was found, $\chi^2(4) = 147$, $p < 0.001$. Results of post-hoc pairwise comparisons indicated significant differences between vowel contrasts except for /i-ɛ/ versus /i-i/, $p = 1$, and /ʌ-ɑ/ versus /ʌ-æ/, $p = .93$.

As seen in the graph in Figure 1, discrimination accuracy varied as a function of the particular vowel contrast for early and late bilingual listeners (for the interaction effect, $\chi^2(8) = 21.98$, $p = .005$). Monolingual AE listeners performed close to ceiling for all vowel contrasts: /i-i/ (99%), /ʌ-æ/ (98%), /i-ɛ/ (96%), /ʌ-ɑ/ (96%), and /æ-ɑ/ (94%). A descriptive analysis revealed that for early bilingual listeners, higher discrimination accuracy

was evident for vowel contrasts /ɪ-ɛ/ (91%) and /ɪ-i/ (86%) than for vowels contrasts /ʌ-æ/ (77%), /ʌ-ɑ/ (73%), and /æ-ɑ/ (72%). Similarly, late bilingual listeners showed higher discrimination accuracy for /ɪ-ɛ/ (78%) and /ɪ-i/ (74%) and showed lower accuracy in discriminating vowel contrasts /æ-ɑ/ (60%), /ʌ-ɑ/ (59%), and /ʌ-æ/ (57%). Results of post-hoc pairwise comparisons indicated significant differences for all vowel contrasts between all listener groups except for vowel contrast /ɪ-ɛ/ between early bilingual and monolingual listeners, $p = .119$, and vowel contrast /æ-ɑ/ between early and late bilingual listeners, $p = .086$.

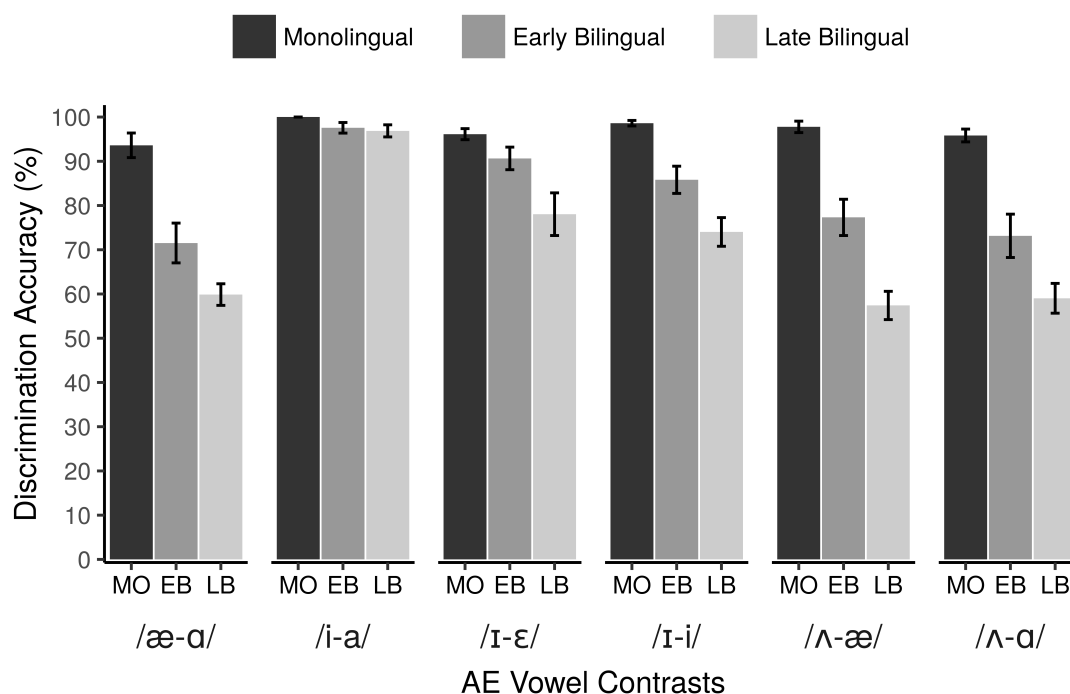


Figure 1. Mean discrimination accuracy of American-English (AE) vowel contrasts by monolingual AE listeners (MO), and early (EB) and late bilingual (LB) listeners. Percentages correct and standard error of the mean are given.

To determine the relationship between perceptual assimilation and discrimination, a cross-language assimilation overlap score (Levy, 2009b) was calculated for each vowel contrast in each bilingual group and compared with the contrast's discrimination error score, as shown in Table 5. The left-hand column lists the vowel contrast and is followed by the cross-language assimilation overlap score and discrimination percentage error, arranged by overlap score in ascending order. Visual inspection of the descriptive results revealed variable support for the PAM-L2 (Best & Tyler, 2007). For example, for the control contrast /i-ɑ/, the AE vowels were perceptually assimilated to different categories (Spanish /i/ and /a/), yielding a low percentage of cross-language assimilation overlap, 0.92%, 0.73%, respectively, by early and late bilingual listeners. Thus, this contrast was discriminated with fewer errors, 2.4% and 3.1%, respectively, as would be predicted by the PAM-L2. Similarly, both segments /æ/ and /ɑ/ were perceptually assimilated to Spanish /a/ by early and late

bilingual listeners, yielding a higher cross-language assimilation overlap score, 79.8% and 75.9%, respectively, for vowel contrast /æ-ɑ/ when compared to other vowel contrasts. As expected, the contrast was poorly discriminated by early and late bilingual listeners, which resulted in higher discrimination errors, 28.5% and 40.1%, respectively, when compared to other vowel contrasts. However, the contrasts' similarly high cross-language assimilation overlap scores yielded quite variable discrimination error scores, which would not be predicted by the model. Moreover, for early bilinguals, cross-language assimilation-overlap scores appeared higher for vowel contrasts /ɪ-ɛ/ and /ɪ-i/, 48.3% and 51.3%, respectively, than /ʌ-æ/ and /ʌ-ɑ/, 32.4% and 42%, respectively. However, discrimination errors were lower for vowel contrasts /ɪ-ɛ/ and /ɪ-i/, 9.4%, 14.2%, respectively, than for vowel contrasts /ʌ-æ/ and /ʌ-ɑ/, 22.7% and 26.9%, respectively. Late bilinguals followed variable patterns, as well, although overall, there was a general increase in discrimination errors with increases in overlap.

A Spearman rank order correlation indicated a strong correlation between overlap scores and discrimination errors, $\rho = .543$, $p = .266$. However, this correlation did not reach statistical significance, likely due to the small sample size. When the control contrast was excluded, the findings were similar. Similarly, for late bilingual listeners, as the cross-language assimilation-overlap increased, so did discrimination errors, $\rho = .829$, $p = .042$. This suggests that for late bilingual listeners, perceptual assimilation patterns were highly related to discrimination performance, although results should be interpreted with caution due to the small sample size.

Table 5. Discrimination accuracy and cross-language assimilation overlap by (early and late bilingual) listener group and vowel contrast. A cross-language assimilation overlap score and a categorial discrimination percentage error score for each bilingual group are presented.

Vowel contrast	N	Cross-language assimilation overlap (%)	Categorial discrimination percentage errors (%)
Early bilingual listeners			
/i-ɑ/	12	0.9	2.4
/ʌ-æ/	12	32.4	22.7
/ʌ-ɑ/	12	42.0	26.9
/ɪ-ɛ/	12	48.3	9.4
/ɪ-i/	12	51.3	14.2
/æ-ɑ/	12	79.8	28.5
Late bilingual listeners			
/i-ɑ/	12	0.7	3.1
/ɪ-ɛ/	12	49.1	22.0
/ɪ-i/	12	52.8	26.0
/ʌ-æ/	12	60.8	42.6
/æ-ɑ/	12	75.9	40.1
/ʌ-ɑ/	12	76.1	41.0

4 Discussion

This study aimed to examine differences between AE vowel perception in early and late Spanish-English bilinguals and to determine whether early bilinguals could perceive AE vowels in a native-like manner. Results revealed that early L2 learning was associated with greater ability to perceive cross-language phonetic differences than later L2 learning. However, even the early bilinguals did not demonstrate native-like perception of AE vowels. Our findings were, for the most part, consistent with the PAM-L2 (Best & Tyler, 2007) in that vowels that were assimilated to separate categories by early and late bilinguals in the current study were often discriminated more accurately than vowels that were assimilated to the same native category. These results are also in line with previous findings from cross-language speech perception studies (Archila-Suerte et al., 2011; Flege, 1991; Flege et al., 1999; Flege & MacKay, 2004; Levy 2009a; Levy & Strange, 2008; Morrison 2002; Pallier et al., 1997).

4.1 Perceptual assimilation: Language background and particular vowels

The present study provided new information on Spanish-English bilinguals' perceptual assimilation patterns. Early and late bilingual listeners consistently assimilated AE front vowel / ϵ / to a single native category (Spanish / e /), for example, suggesting that AE / ϵ / was perceived as a good exemplar of Spanish / e /). Conversely, early and late bilingual listeners at times assimilated AE vowels to more than one native category. For example, AE central vowel / Λ / was most often assimilated to Spanish back vowels (/u/ and /o/) and to a Spanish central vowel /a/. This AE vowel may have been assimilated to Spanish /a/ due to the vowels' proximity in vowel space (AE vowel / Λ /: F1 820 Hz, F2 1522 Hz; Spanish /a/: F1 638 Hz, F2 1353 Hz) (Bradlow, 1995).

Surprisingly, AE vowel / Λ / was assimilated to Spanish /u/ in "bupo." However, we speculate that this finding may be less attributable to perceptual patterns than to native language orthographic interference during labeling of the stimuli. Specifically, the English orthographic representation of "u" often corresponds with the AE speech sound / Λ /; thus, the listener, when presented with / Λ / may have associated the vowel with the English letter "u" and selected the "bupo" response. (A limitation of the familiarization was that listeners were not instructed to pronounce stimuli, potentially permitting such orthographic interference.) Of interest, early bilinguals selected the /u/ response far more often than late bilinguals, potentially reflecting stronger English interference, that is, greater difficulty treating the response options as Spanish throughout the task, than did late bilinguals. The affected response choices may also have contributed to the unclear relationship between perceptual assimilation and discrimination.

Lastly, early and late bilingual listeners assimilated several AE vowels (AE back vowel / α / and AE front vowel / \ae /) to a single native category (central Spanish vowel /a/). Although AE / \ae / and Spanish / e / are front vowels, listeners may have assimilated AE / \ae / to Spanish /a/ because Spanish / e / was perceived as a good exemplar of AE vowel / ϵ /).

Overall, the perceptual assimilation patterns found in the present study are similar to those found in Morrison's (2008) perceptual assimilation study of Western Canadian-English /i/ and /I/ synthetic vowels by monolingual Mexican-Spanish and monolingual Peninsular-Spanish listeners, Flege's (1991) identification study of AE /i/, /I/, / ϵ /, and

/æ/ vowels by late Spanish-English bilingual listeners, as well as Escudero and Chládková's (2010) perceptual assimilation study of synthetic SSBE and AE /i/, /ɪ/, /u/, /ʊ/, /ɛ/, /ʌ/, /ɔ/, /æ/, and /ɑ/ by Peruvian Spanish monolingual listeners, despite the differences in the listeners' language backgrounds among studies. As will be discussed further regarding discrimination patterns, although assimilation and identification of AE /i/ and /ɪ/ have received considerable attention in the literature (Escudero & Boersma, 2004; Escudero & Chládková, 2010; Flege, 1991; Flege et al., 1997; Morrison, 2002, 2008), assimilation of AE /æ/, /ɑ/, and /ʌ/ vowels also appears to pose challenges for native speakers of Spanish, especially for the late learners, given that single category assimilation patterns likely yield poor discrimination (Best & Tyler, 2007).

Although similar perceptual assimilation patterns were found between early and late bilingual listeners in the present study, late bilinguals accepted AE vowels as better instances of their Spanish categories, as indicated by their high median goodness ratings. This is consistent with Levy's (2009a) finding that overall median goodness ratings for Parisian French vowels were higher for AE listeners with minimal French experience than for AE listeners with extensive French experience, suggesting that listeners perceived L2 vowels as less like their native vowels with increased French language experience. It should be noted, though, that in the present study, perception by early versus late learners was examined, rather than that of only late learners with moderate versus extensive (including immersion) language experience. Therefore, consistent with the PAM-L2, we found evidence that learning to recognize phonetic differences between native and L2 vowels comes with extensive exposure to the L2.

4.2 Categorical discrimination: language background and particular vowel contrast effects

In this study, Spanish-English bilinguals' early L2 learning was associated with increased ability to discriminate cross-language phonetic differences, consistent with conclusions from the cross-language speech perception literature (e.g., Flege & MacKay, 2004; Flege et al., 1999) and hypothesized at the outset of the study. Our finding that monolingual AE listeners discriminated AE vowels most accurately, followed by early bilingual and late bilingual listeners is in line with, for example, Flege and MacKay's (2004) and Flege et al.'s (1999) findings that early Italian-English bilinguals discriminated Canadian-English vowels more accurately than late Italian-English bilinguals. The significant language background and experience effects are also consistent with Levy's (2009b) and Levy and Strange's (2008) finding that late bilingual listeners who had studied French as an L2 discriminated most French vowels more accurately than non-French-speaking AE listeners. Furthermore, AE listeners with extensive French immersion experience discriminated Parisian French vowel contrasts more accurately than AE listeners with only formal French experience (Levy, 2009b). However, in the present study, although discrimination accuracy of L2 vowels improved with early age of L2 learning (mean age = 5 years), early bilingual listeners' vowel discrimination was not native-like, suggesting that early bilingual listeners' native vowel system may continue to influence their L2 vowel perception. These findings are consistent with Pallier et al. (1997), who found that early Spanish-Catalan bilingual listeners had difficulty discriminating Catalan vowel contrasts despite their early exposure to Catalan, but differ from Flege et al. (1999) who did find native-like discrimination by Italian-English

bilingual listeners as their performance did not significantly differ from Canadian-English monolingual listeners.

The inconsistency of the current findings with those of Flege et al.'s (1999) study may be due to variables such as the different ages of participants. The early bilinguals examined in the current study had learned their L2 prior to 12 years of age, and were a mean age of 26 years at the time of testing. The participants examined in Flege et al.'s (1999) study, in contrast, were first exposed to their L2 at 7 years of age and were a mean age of 48 at the time of testing and were therefore more likely to have used their L2 for a longer period of time, which may have impacted results. Thus, further examination of length of L2 use is needed to ascertain its effect on L2 vowel perception by early bilingual listeners.

Individual differences among early bilingual listeners in the present study may have influenced their perceptual accuracy. For example, the early bilinguals' daily use of Spanish varied considerably (25% vs. 75%). High use of a native language may affect the perception of phonetic properties of the native language despite listeners' early age of L2 learning (Flege & MacKay, 2004). Additionally, in the present study, LOR varied among early bilingual listeners (10 vs. 42 years), also potentially impacting the accuracy of L2 vowel perception (Flege, 1991; Flege et al., 1997; Morrison, 2002). Further studies with more participants are needed to control for such factors confounded with age of L2 learning.

Discrimination accuracy varied as a function of the particular vowel contrast in this study, suggesting that the relationship between Spanish and AE vowel inventories causes some vowels to be more difficult to discriminate than others. Higher discrimination accuracy was revealed when AE vowels were assimilated to distinct native vowels. For example, the control contrast /a-i/, which contains vowels distant in vowel space, resulted in few discrimination errors.

Other assimilation patterns were also evident. For example, although early and late bilingual listeners assimilated each AE vowel in the contrast /ɪ-ε/ to Spanish /e/, they assimilated AE /ε/ to the Spanish /e/ category more often than they assimilated AE /ɪ/ to that category. Therefore, AE vowel /ε/ may have been perceived as a "better instance" of Spanish /e/ than AE vowel /ɪ/, yielding moderate to very good discrimination accuracy according to the PAM-L2 (Best & Tyler, 2007).

Additionally, the vowel contrast /ɪ-i/ was discriminated with higher accuracy than vowel contrasts /ʌ-æ/, /ʌ-ɑ/, and /æ-ɑ/ by all groups of listeners. Despite the proximity of these vowels in vowel space, each AE vowel was assimilated to two different native vowels (Spanish /e/ and /i/), which resulted in fewer discrimination errors. This finding conflicts with previous studies that indicate poor discrimination accuracy for the AE /ɪ-i/ vowel contrast (Escudero & Chládková, 2010; Morrison, 2002, 2008). However, it should be noted that previous studies included different English dialects (Scottish, British, and Canadian English synthetic vowels) and participants (Spanish monolingual listeners). Perceptual learning patterns are also expected to change throughout the process of L2 learning. Thus, the reasons for the inconsistent findings cannot be ascertained, but are of interest for future exploration.

Conversely, AE vowel contrasts that included AE central vowel /ʌ/, such as /ʌ-æ/ and /ʌ-ɑ/ were discriminated with poorer accuracy. Findings from Archila-Suerte et al.'s (2011) study also indicated poor discrimination of /ʌ-ɑ/ by bilingual listeners. In the present study, listeners assimilated the vowels in each of these contrasts to a single native vowel (Spanish

/a/). Vowel contrast /æ-ɑ/, containing two low vowels, was also discriminated poorly. These findings support the hypothesis that L2 vowels will be discriminated less accurately if both L2 vowels are assimilated to instances of a single native language vowel (Best & Tyler, 2007). Because of the perceptual challenges the contrasts pose, these contrasts might be of particular importance in the design of perceptual training programs for native Spanish listeners learning English.

By examining the relationship between native and L2 vowels through the framework of the PAM-L2 (Best & Tyler, 2007), predictions can be made regarding the listeners' patterns of L2 perceptual learning. Previous cross-language speech perception studies have tested predictions proposed by the PAM with naïve learners, that is, those new to learning (or listening to) the unfamiliar language. For example, Fabra and Romero (2012) found that late Catalan learners of English varying in English proficiency demonstrated poor discrimination when AE vowels were assimilated in a single category assimilation type pattern and good discrimination when vowels were assimilated in a category goodness assimilation type pattern.

In the present study, the predictions were examined for listeners with far more extensive language experience than in Fabra and Romero's (2012) Catalan study. The PAM-L2 (Best & Tyler, 2007) was supported for the most part in our finding of a strong correlation between perceptual assimilation patterns and discrimination errors. The absence of statistical significance could be a reflection of the small sample size and/or of certain vowel pairs' unpredicted discriminability—and this varied with L2 experience. For example, while the /ɪ-ε/ contrast showed similar overlap scores in early and late bilinguals, late bilinguals revealed 22% discrimination errors, whereas early bilinguals had only 9.4% discrimination errors for this contrast, suggesting far greater ease of discrimination (but little assimilation difference) with early L2 experience. A different pattern was evident for contrasts involving the AE central vowel /ʌ/, with early bilinguals showing less assimilation overlap of this vowel with other vowels, and, as predicted, fewer discrimination errors involving this speech sound. However, unlike /ɪ-ε/, discrimination was consistently relatively poor for /ʌ/ contrasts, even for listeners with early AE experience, who revealed 25% discrimination errors. Thus, when individual contrasts' perceptual patterns are examined, the learnability of particular speech sounds appears to differ in the domains of assimilation versus discrimination, resulting in the relationship between assimilation and discrimination and the predictability of discrimination accuracy differing as a function of age of acquisition of L2 vowels.

4.3 Limitations and future directions

By examining several non-native contrasts with various assimilation patterns, it was possible to test the predictions of the PAM-L2 regarding the changes that occur in L2 vowel learning as listeners gain experience. Especially with the increasing numbers of Hispanic immigrants who are entering the US (US Census Bureau, 2006) and learning AE as a L2, this information is valuable in that inaccurate speech-sound perception may negatively impact communication in social, academic and professional settings (Agha, 2003; Flege, 1988). Findings on the perceptual assimilation and discrimination of AE vowels by Spanish-English bilinguals may thus assist professionals working with this population. For example, knowing the AE contrasts that are perceived least accurately by native Spanish listeners (e.g., /ʌ-æ/, /ʌ-ɑ/, and /æ-ɑ/) may be of use in the development of training programs that tar-

get perceptual accuracy of these challenging contrasts as early as possible in the listeners' learning of the L2.

It is important to note the methodological limitations of this study. For example, the stimuli were uttered by only three native AE talkers who were all from the New York regional area for consistency of dialect. Previous studies have shown that the dialect of the speaker's language may have an effect on the way vowels are perceived by non-native listeners (Escudero & Chládková, 2010). Additionally, the vowels were presented in one (bilabial) consonantal context, although alveolar context, for example, can yield different perceptual assimilation and discrimination patterns from bilabial context in L2 learners (Levy, 2009b; Levy & Strange, 2007). Such factors limit the generalizability of the results. Lastly, Spanish listeners' perception of other non-native sounding AE vowels, such as /ə/, /ɜ/, /ɝ/, should be investigated in future studies.

A further limitation of the study was the complexity of the listeners' language histories. Listeners in this study came from diverse language backgrounds with a range of continued use of the native language and LOR in the L2 country. These factors have been found to affect accuracy of L2 speech perception (Flege, 1991; Flege & MacKay, 2004; Flege et al., 1997; Levy, 2009b; Levy & Strange, 2008; Morrison, 2002, 2008). Additionally, listeners from various Latin American countries with different dialects were grouped together, which may have obscured perceptual differences between listeners (Chládková & Podlipský, 2011; Escudero & Boersma, 2004; Morrison, 2008). An analysis grouping the listeners by continued use of the native language, dialect, and LOR, may reveal perceptual differences in the listeners. Future studies with more participants could allow subgroup comparisons without loss of statistical power. Additionally, future studies with a greater age gap between listener groups, as well as correlational studies, could further clarify commonalities and differences between early and late L2 learners.

Additionally, examining the relationship between perception and production of L2 vowels (Flege et al., 1997, 1999; Jia, Strange, Wu, Collado, & Guan, 2006; Levy & Law, 2009) will further shed light on repercussions for communication in the L2 environment. Finally, an extension of this study to children's AE vowel perception would help document changes in native-Spanish speaking children's L2 perception and serve as a foundation to investigate perceptual challenges experienced by native Spanish-speaking children with communication disorders who are faced with a new phonological inventory.

Acknowledgements

The authors wish to thank the Speech Production and Perception Laboratory at Teachers College, Columbia University, with special thanks to Dorothy Leone, Gemma Moya-Gale, and Sih-Chiao Hsu. The authors also express their appreciation to Carol Hammer, Hansun Waring, Megan McAuliffe, and Robert Remez for their suggestions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Adachi, T., Akahane-Yamada, R., & Ueda, K. (2006). Intelligibility of English phonemes in noise for native and non-native listeners. *Acoustical Science and Technology*, 27(5), 285–289.
- Agha, A. (2003). The social life of cultural value. *Language and Communication*, 23(3–4), 231–273.
- Archila-Suerte, P., Zevin, J., Bunta, F., & Hernandez, A. E. (2011). Age of acquisition and proficiency in a second language independently influence the perception of non-native speech. *Bilingualism: Language and Cognition*, 15(1), 190–201.
- 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 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.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam, The Netherlands: John Benjamins.
- Bohn, O. S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11(3), 303–328.
- Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America*, 97(3), 1916–1924.
- Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *Journal of the Acoustical Society of America*, 121(4), 2339–2349.
- Broersma, M., & Scharenberg, O. (2010). Native and non-native listeners' perception of English consonants in different types of noise. *Speech Communication*, 52(11–12), 980–995.
- Chládková, K., & Podlipský, V. J. (2011). Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners. *Journal of the Acoustical Society of America*, 130(4), EL186–EL192.
- Escudero, P. (2014). The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. *Journal of the Acoustical Society of America*, 135(3), 1577–1584.
- Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, 26(4), 551–585.
- Escudero, P., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *Journal of the Acoustical Society of America*, 128(5), EL254–EL 260.
- Fabra, L. R., & Romero, J. (2012). Native Catalan learners' perception and production of English vowels. *Journal of Phonetics*, 40(3), 491–508.
- Ferguson, S. H. (2012). Talker differences in clear and conversational speech: Vowel intelligibility for older adults with hearing loss. *Journal of Speech, Language, and Hearing*

- Research*, 55(3), 779–790.
- Flege, J. E. (1988). The production and perception of foreign language speech sound. In H. Winitz (Ed.), *Human communication and its disorders, A review - 1988* (pp. 224–401). Norwood, NJ: Ablex.
- Flege, J. E. (1991). The interlingual identification of Spanish and English vowels: Orthographic evidence. *Quarterly Journal of Experimental Psychology*, 43(3), 701–731.
- Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26(1), 1–34.
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437–470.
- Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*, 106(5), 2973–2987.
- Flege, J. E., Munro, M. J., & Fox, R. A. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, 95(6), 3623–3641.
- Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40(4), 467–491.
- Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America*, 97(4), 2540–2551.
- Garcia Lecumberri, M. L., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, 52(11–12), 864–886.
- Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics*, 12(2), 91–114.
- Iverson, P., & Evans, B. G. (2007). Learning English vowels with different first-language vowel systems: perception of formant targets, formant movement, and duration. *Journal of the Acoustical Society of America*, 122(5), 2842–2854.
- Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *Journal of the Acoustical Society of America*, 126(2), 866–877.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446.
- Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *Journal of the Acoustical Society of America*, 119(2), 1118–1130.
- Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60–99.
- Kangas, K. A., & Allen, G. D. (1990). Intelligibility of synthetic speech for normal-hearing and hearing-impaired listeners. *Journal of Speech and Hearing Disorders*, 55(4), 751–755.

- Kewley-Port, D., Burkle, T. Z., & Lee, J. H. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *Journal of the Acoustical Society of America*, *122*(4), 2365–2375.
- Levy, E. S. (2009a). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *Journal of the Acoustical Society of America*, *125*(2), 1138–1152.
- Levy, E. S. (2009b). On the assimilation–discrimination relationship in American English adults’ French vowel learning. *Journal of the Acoustical Society of America*, *126*(5), 2670–2682.
- Levy, E. S., & Law II, F. (2009). Perception–production relationship in French vowel learning in adulthood. *Journal of the Acoustical Society of America*, *125*(4), 2772–2773.
- Levy, E. S., & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, *36*(1), 141–157.
- Levy, E. S., Leone, D., Garcia, P., & Baigorri, M. (2010, November). *American English adults’ and children’s perception of Spanish accented conversational and clear speech*. Presented at the American Speech-Language-Hearing Association convention, Philadelphia, PA.
- Long, M. H. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, *12*(3), 251–285.
- Mack, M. (1989). Consonant and vowel perception and production: Early English–French bilinguals and English monolinguals. *Perception and Psychophysics*, *46*(2), 186–200.
- McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English, and Estonian. *Journal of Phonetics*, *30*(2), 229–258.
- Morrison, G. S. (2002). Perception of English /i/ and /I/ by Japanese and Spanish listeners: longitudinal results. In G. S. Morrison & L. Zsoldos (Eds.), *Proceedings of North West Linguistics Conference 2002* (pp. 29–48). Burnaby, BC: Simon Fraser University Linguistics Graduate Student Association.
- Morrison, G. S. (2008). Perception of synthetic vowels by monolingual Canadian-English, Mexican-Spanish, and Peninsular-Spanish listeners. *Canadian Acoustics*, *36*(4), 17–23.
- Morrison, G. S., & Escudero, P. (2007). A cross-dialect comparison of Peninsula- and Peruvian-Spanish vowels. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1505–1508). Saarbrücken, Germany: University of Saarbrücken.
- Neuman, A., & Hochberg, I. (1983). Children’s perception of speech in reverberation. *Journal of the Acoustical Society of America*, *73*(6), 2145–2149.
- Pallier, C., Bosch, L., & Sebastian-Galles, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, *64*(3), B9–B17.
- Piske, T., Flege, J. E., MacKay, I. R. A., & Meador, D. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica*, *59*(1), 49–71.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*,

- 20(2), 421–435.
- Rogers, C., Lister, J., Febo, D., Besing, J., & Abrams, H. (2006). Effects of bilingualism, noise and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, 27(3), 465–485.
- Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, 72(2), 111–123.
- Shi, L.-F. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of Speech, Language, and Hearing Research*, 53(4), 821–835.
- StataCorp. (2013). *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
- Strange, W., Weber, A., Levy, E., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *Journal of the Acoustical Society of America*, 122(2), 1111–1129.
- Tagliaferri, B. (2011). Paradigm (Version 1.0.2) [Computer software]. Retrieved from <http://www.paradigmexperiments.com>
- U.S. Census Bureau (2006). *U.S. Hispanic Population: 2006*. Retrieved from https://www.census.gov/content/dam/Census/library/working-papers/2007/demo/CPS_Hispanic_Population_2006.pdf
- Ueda, K., Akahane-Yamada, R., & Komaki, R. (2002). Identification of English /r/ and /l/ in white noise by native and non-native listeners. *Acoustical Science and Technology*, 23(6), 336–338.
- Von Hapsburg, D., Champlin, C. A., & Shetty, S. R. (2004). Reception thresholds for sentences in bilingual (Spanish/English) and monolingual (English) listeners. *Journal of the American Academy of Audiology*, 15(1), 88–98.

Supplementary Material

Perception of American–English vowels by early and late Spanish–English bilinguals

M. Baigorri, L. Campanelli, & E. S. Levy

Table S1. Experiment 1 (Perceptual Assimilation). Summary of mixed effects logistic regression analyses for perceptual assimilation patterns in EB (early bilingual) and LB (late bilingual) listeners.

AE Vowel	Spanish Response Vowel	EB (%)	LB (%)	Odds Ratio	Std. Err.	<i>z</i>	<i>p</i>	95% CI
/i/	/a/	0	0.2	—				
	/e/	4.1	3.3	1.286	0.978	0.33	.741	[0.29; 5.707]
	/i/	95.4	96.3	0.82	0.545	−0.3	.766	[0.223; 3.02]
	/o/	0.4	0	—				
	/u/	0.2	0.2	1	1.416	0	1	[0.062; 16.03]
/ɪ/	/a/	0.6	0.2	0.332	0.384	−0.95	.34	[0.034; 3.203]
	/e/	48.5	48.9	0.625	0.691	−0.42	.671	[0.072; 5.459]
	/i/	50	50	1.857	2.058	0.56	.577	[0.212; 16.301]
	/o/	0.2	0.9	5.019	6.443	1.26	.209	[0.405; 62.135]
	/u/	0.7	0	—				
/ɛ/	/a/	4.6	11.3	3.393	1.748	2.37	.018	[1.236; 9.312]
	/e/	93.5	86.7	0.38	0.173	−2.12	.034	[0.156; 0.929]
	/i/	0.7	0.6	0.749	0.574	−0.38	.706	[0.167; 3.361]
	/o/	0.2	1.5	8.167	9.076	1.89	.059	[0.925; 72.11]
	/u/	0.9	0	—				
/æ/	/a/	91.7	82	0.39	0.141	−2.6	.009	[0.192; 0.794]
	/e/	7	13.5	2.197	0.888	1.95	.051	[0.995; 4.851]
	/i/	0.2	0.6	2.989	3.864	0.85	.397	[0.237; 37.654]
	/o/	0.9	3.5	3.839	2.683	1.92	.054	[0.976; 15.107]
	/u/	0.2	0.4	2.004	2.457	0.57	.571	[0.181; 22.163]
/ɑ/	/a/	82.6	75	0.679	0.482	−0.54	.586	[0.169; 2.733]
	/e/	0.4	1.3	4.407	5.537	1.18	.238	[0.375; 51.721]
	/i/	0	0.2	—				
	/o/	10.9	22.8	1.834	1.274	0.87	.383	[0.47; 7.156]
	/u/	6.1	0.7	0.174	0.223	−1.37	.172	[0.014; 2.141]
/ʌ/	/a/	29.4	53.1	9.729	9.122	2.43	.015	[1.549; 61.119]
	/e/	4.3	6.7	1.675	1.171	0.74	.461	[0.426; 6.59]
	/i/	0.2	0.7	4.022	4.504	1.24	.214	[0.448; 36.105]
	/o/	24.4	38.5	2.545	1.837	1.29	.196	[0.618; 10.477]
	/u/	41.7	0.9	0.001	0.001	−3.66	<.001	[0; 0.035]
/o/	/a/	0.9	0.3	0.332	0.384	−0.95	.34	[0.034; 3.2]
	/e/	0	1.5	—				
	/i/	0	0	—				
	/o/	96	97.8	0.972	1.304	−0.02	.983	[0.07; 13.485]
	/u/	3.1	0.3	0.415	1.19	−0.31	.759	[0.002; 114.374]

Table S2. Experiment 2 (Categorical Discrimination). Summary of mixed effects logistic regression analysis for the effects of Group and Vowel contrast on response accuracy.

Fixed Effect	Estimate(SE)	z	p	95% CI
/r-i/	1.056(0.534)	1.98	0.048	[0.01; 2.103]
/æ-a/	-0.541(0.359)	-1.51	0.131	[-1.244; 0.162]
/ʌ-æ/	0.574(0.458)	1.25	0.21	[-0.324; 1.473]
/ʌ-a/	-0.082(0.39)	-0.21	0.834	[-0.845; 0.682]
group-EB	-0.892(0.433)	-2.06	0.04	[-1.742; -0.043]
group-LB	-2.098(0.415)	-5.06	<.001	[-2.911; -1.285]
/r-i/ × EB	-1.55(0.579)	-2.68	0.007	[-2.684; -0.415]
/r-i/ × LB	-1.28(0.559)	-2.29	0.022	[-2.374; -0.185]
/æ-a/ × EB	-0.898(0.414)	-2.17	0.03	[-1.71; -0.087]
/æ-a/ × LB	-0.344(0.391)	-0.88	0.379	[-1.111; 0.423]
/ʌ-æ/ × EB	-1.681(0.504)	-3.33	0.001	[-2.67; -0.692]
/ʌ-æ/ × LB	-1.564(0.484)	-3.23	0.001	[-2.513; -0.615]
/ʌ-a/ × EB	-1.269(0.441)	-2.87	0.004	[-2.134; -0.403]
/ʌ-a/ × LB	-0.839(0.42)	-2	0.046	[-1.662; -0.016]
Intercept	3.395(0.353)	9.61	<.001	[2.703; 4.088]
<i>Random Effect</i>				
Intercept	0.393(0.123)			[0.212; 0.727]

Note. EB = early bilingual; LB = late bilingual. Monolingual AE listeners are the reference group.

Table S3. Experiment 2 (Categorical Discrimination). Pairwise comparisons for the effect of Group on discrimination accuracy. P values are corrected for multiple comparisons using the Bonferroni method.

Comparison	Estimate(SE)	z	p	95% CI
EB vs MO	-1.985(0.318)	-6.24	<.001	[-2.609; -1.361]
LB vs MO	-2.913(0.316)	-9.21	<.001	[-3.533; -2.293]
LB vs EB	-0.928(0.269)	-3.45	.002	[-1.455; -0.401]

Note. EB = early bilingual; LB = late bilingual; MO = monolingual.

Table S4. Experiment 2 (Categorical Discrimination). Pairwise comparisons for the effect of Vowel contrast on discrimination accuracy. P-values are corrected for multiple comparisons using the Bonferroni method.

Comparison	Estimate(SE)	z	p	95% CI
/ɪ-ɛ/ vs /ɪ-i/	-0.058(0.185)	-0.31	1	[-0.577; 0.462]
/æ-ɑ/ vs /ɪ-i/	-1.037(0.169)	-6.13	<.001	[-1.512; -0.563]
/ʌ-æ/ vs /ɪ-i/	-0.628(0.189)	-3.32	.009	[-1.16; -0.097]
/ʌ-ɑ/ vs /ɪ-i/	-0.883(0.175)	-5.05	<.001	[-1.375; -0.392]
/æ-ɑ/ vs /ɪ-ɛ/	-0.98(0.14)	-7.01	<.001	[-1.372; -0.587]
/ʌ-æ/ vs /ɪ-ɛ/	-0.571(0.164)	-3.49	.005	[-1.03; -0.112]
/ʌ-ɑ/ vs /ɪ-ɛ/	-0.826(0.147)	-5.62	<.001	[-1.238; -0.414]
/ʌ-æ/ vs /æ-ɑ/	0.409(0.145)	2.82	.048	[0.002; 0.816]
/ʌ-ɑ/ vs /æ-ɑ/	0.154(0.126)	1.22	<.001	[-0.199; 0.507]
/ʌ-ɑ/ vs /ʌ-æ/	-0.255(0.152)	-1.68	.933	[-0.682; 0.171]

Table S5. Experiment 2 (Categorical Discrimination). Pairwise comparisons for the Group \times Vowel contrast interaction on discrimination accuracy. P values are corrected for multiple comparisons using the Bonferroni method.

Comparison	Estimate(SE)	z	p	95% CI
<i>/i-i/</i>				
EB vs MO	-2.442(0.549)	-4.45	<.001	[-3.517; -1.367]
LB vs MO	-3.377(0.541)	-6.24	<.001	[-4.438; -2.317]
LB vs EB	-0.936(0.315)	-2.97	.009	[-1.553; -0.318]
<i>/i-ε/</i>				
EB vs MO	-0.892(0.433)	-2.06	.119	[-1.742; -0.043]
LB vs MO	-2.098(0.415)	-5.06	<.001	[-2.911; -1.285]
LB vs EB	-1.205(0.334)	-3.61	.001	[-1.861; -0.55]
<i>/æ-ɑ/</i>				
EB vs MO	-1.791(0.369)	-4.86	<.001	[-2.513; -1.068]
LB vs MO	-2.442(0.365)	-6.69	<.001	[-3.157; -1.727]
LB vs EB	-0.651(0.298)	-2.19	.086	[-1.235; -0.068]
<i>/Λ-æ/</i>				
EB vs MO	-2.573(0.469)	-5.49	<.001	[-3.492; -1.654]
LB vs MO	-3.661(0.464)	-7.89	<.001	[-4.571; -2.752]
LB vs EB	-1.088(0.301)	-3.62	.001	[-1.678; -0.499]
<i>/Λ-ɑ/</i>				
EB vs MO	-2.161(0.4)	-5.41	<.001	[-2.944; -1.378]
LB vs MO	-2.937(0.396)	-7.42	<.001	[-3.713; -2.161]
LB vs EB	-0.776(0.298)	-2.6	.028	[-1.361; -0.191]

Note. EB = early bilingual; LB = late bilingual; MO = monolingual.