PERCEPTION OF NATURAL VOWELS BY MONOLINGUAL CANADIAN-ENGLISH, MEXICAN-SPANISH, AND PENINSULAR-SPANISH LISTENERS

Geoffrey Stewart Morrison
Forensic Voice Comparison Laboratory, School of Electrical Engineering & Telecommunications,
University of New South Wales, UNSW Sydney, NSW 2052, Australia
e-mail: geoff-morrison@forensic-voice-comparison.net

ABSTRACT

On the basis of a previously-reported synthetic-vowel perception experiment, it was hypothesized that the location of the perceptual boundary between Spanish /i/ and /e/ differed for monolingual Peninsular-Spanish and Mexican-Spanish listeners (north-central Spain and Mexico City), and that this would affect the perception of the Canadian-English /i/–/ɪ/ contrast (western Canada): Peninsular-Spanish listeners were predicted to identify almost all tokens of Canadian-English /i/ as Spanish /i/ and almost all tokens of Canadian-English /ɪ/ as Spanish /e/ (two-category assimilation); whereas Mexican-Spanish listeners were predicted to identify almost all tokens of Canadian-English /i/ as Spanish /i/, but identify some tokens of Canadian-English /ɪ/ as Spanish /i/ and some as Spanish /e/. Monolingual Peninsular-Spanish and Mexican-Spanish listeners' perception of natural tokens of English /i/, /ɪ/, /e/, and /ɛ/ produced by monolingual Canadian-English speakers was tested. Both the Peninsular-Spanish and the Mexican-Spanish listeners had results consistent with the perceptual pattern predicted for the Peninsular-Spanish listeners. The results call into question the assumption that first-language-Spanish learners of English have difficulty learning the English /i/–/ɪ/ contrast because they initially assimilate most tokens of both English vowel categories to a single Spanish vowel category, Spanish /i/.

RÉSUMÉ

En nous fondant sur les résultats d'une expérience antérieure consacrée à la perception de voyelles synthétiques, nous avons émis l'hypothèse qu'en espagnol, la frontière perceptive entre /i/ et /e/ différait chez les auditeurs monolingues en fonction de leur origine géographique (nord-est de l'Espagne vs ville de Mexico) et que cela était susceptible d'affecter la perception que les auditeurs hispanophones ont du contraste entre les voyelles anglo-canadiennes /i/ et /ɪ/ (de l'ouest du Canada). Plus précisément, nous postulons 1. que les auditeurs hispanophones originaires d'Espagne identifieront, respectivement, la quasi-totalité des occurrences des voyelles /i/ et /ɪ/ produites par des locuteurs anglo-canadiens comme des occurrences des voyelles espagnoles /i/ et /e/ (l'assimilation en deux catégories); 2. que les auditeurs mexicains identifieront également la quasi-totalité des occurrences de la voyelle anglaise /i/ comme celles d'un /i/ espagnol, mais que certaines occurrences de la voyelle anglaise /ɪ/ seront identifiées comme des /e/ espagnols alors que d'autres seront associées à un /i/ espagnol. Nous avons testé la perception que des auditeurs hispanophones monolingues avaient des voyelles anglaises /i/, /ɪ/, /e/, et /ɛ/ telles que produites par des locuteurs anglo-canadiens monolingues. Il s'avère que tant les auditeurs natifs d'Espagne que ceux originaires du Mexique présentent des résultats conformes à l'hypothèse 1 visant les auditeurs originaires d'Espagne. Ces résultats mettent en question la supposition selon laquelle les apprenants ayant l'espagnol comme première langue éprouvent des difficultés, lors de leur apprentissage de l'anglais, à appréhender le contraste entre /i/ et /ɪ/, car ils ont, en premier lieu, assimilé une majorité des occurrences appartenant à deux classes vocaliques distinctes en anglais à une classe vocalique unique en espagnol, le /i/.

1. INTRODUCTION

The present paper is the second in a series of two papers of which Morrison (2008b) "Perception of *synthetic* vowels by monolingual Canadian-English, Mexican-Spanish, and Peninsular-Spanish listeners" is the first.

First-language Spanish speakers have often been reported to have difficulty learning the English /i/–/ɪ/ contrast (e.g., Bohn, 1995; Flege, Bohn, & Jang, 1997; Escudero & Boersma, 2004; Morrison, 2008a, 2009), and it has been hypothesized that this is because they assimilate most tokens of both English vowel phonemes (English /i/ and /ɪ/) to a single Spanish phoneme (Spanish /i/). Results of studies on Peninsular-Spanish and American-Spanish speakers listening to English from the south east of England (Álvarez González, 1980, ch. 5; Escudero, 2005, §1.2.2), and studies of American-Spanish speakers listening to English from the United States (Flege, 1991; Møller Glasbrenner, 2005) have found that:

1. First-language Spanish second-language English (L1-Spanish L2-English) listeners misidentify L1-English speakers' productions of English /i/ as English /ɪ/ and vice versa.
2. Monolingual-Spanish listeners assimilate the majority of tokens of English /i/ to the Spanish /i/ category.
3. Monolingual-Spanish listeners assimilate the majority of tokens of English /ɪ/ to the Spanish /i/ category.

However, these studies also report that Spanish listeners assimilate some tokens of English /ɪ/ to Spanish /e/, and identify some tokens of English /ɪ/ as English /ɛ/.

It is well known that English vowel-phoneme realizations can vary substantially across dialects (Wells, 1982) and it is not therefore unexpected that Spanish listeners' perception of the English /i/–/ɪ/ contrast varies according to the dialect of English spoken. For example, /i/ and /ɪ/ in Scottish English have a larger first-formant (F1) separation and less difference in duration than their counterparts in English from the south east of England (Escudero & Boersma, 2004). Escudero (2005, §1.2.2) found that Peruvian-Spanish listeners assimilated tokens of Scottish English /i/ and /ɪ/ via a two-category assimilation to the Spanish /i/ and /e/ categories respectively, but assimilated tokens of southeastern-England English /i/ and /ɪ/ via a single-category or category-goodness-difference assimilation to the Spanish /i/ category (see Best's, 1995, Perceptual Assimilation Model for these terms).

What is less immediately apparent is whether there are

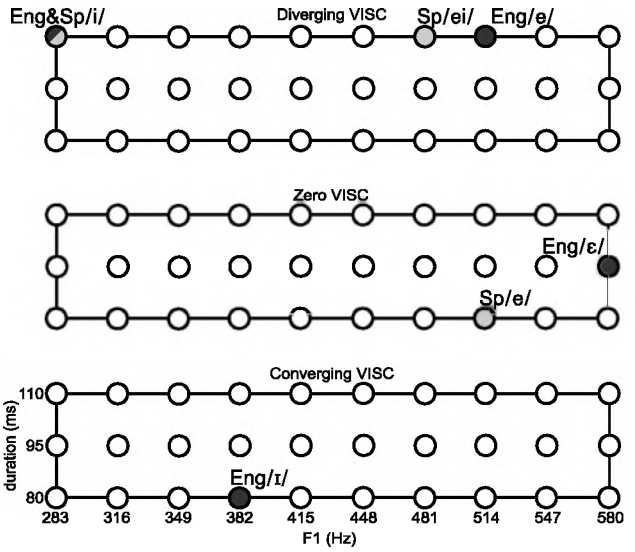

Figure 1. Properties of synthetic stimuli. Reproduced from Morrison (2008b).

differences in vowel-phoneme realizations across Spanish dialects which may lead to Spanish listeners of different dialects perceiving the English /i/–/ɪ/ contrast differently. It has not been uncommon in L2 speech-perception research for L1 Spanish listener groups to be made up of speakers of a mixture of different Spanish dialects (e.g., Flege, Bohn, & Jang, 1997; Escudero & Boersma, 2004; Morrison, 2008a, 2009) implying at least a tacit assumption that the listeners' Spanish dialect is not particularly relevant to their perception of the English /i/–/ɪ/ contrast.

Godínez (1978) tentatively suggested that there were differences in vowel formant values between Peninsular, Mexican, and Argentinean Spanish, but the number of participants in the study was too small to draw any stronger conclusion. Comparing Peninsular- and Peruvian-Spanish vowels in isolated vowels produced at the end of a carrier sentence, Morrison & Escudero (2007) failed to find significant differences in formant values with the exception of a mean 11% difference (as measured in hertz) in the second-formant (F2) value for /o/; however, comparing vowels in nonce words including various consonant contexts at the end of a carrier sentence, Chládková, Escudero, & Boersma (2011) found significant differences for F1 in /a/ (6.3%) and F2 in /e/ and /o/ (4.1% and 4.8%), and more widespread differences in certain consonant contexts.

Escudero & Williams (2012) found that L1 Spanish L2 Dutch listeners' first dialect influenced their perception of Dutch vowels. Peninsular-Spanish listeners were better than Peruvian-Spanish listeners at discriminating the Dutch /a/–/ɑ/ and /i/–/ɪ/ contrasts, and had higher correct-

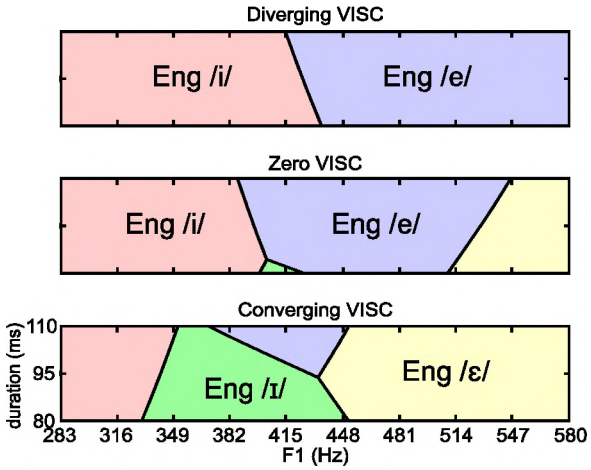

Figure 2. Territorial map of modal response areas from logistic-regression model of Canadian-English listeners' identification of synthetic stimuli. Reproduced from Morrison (2008b).

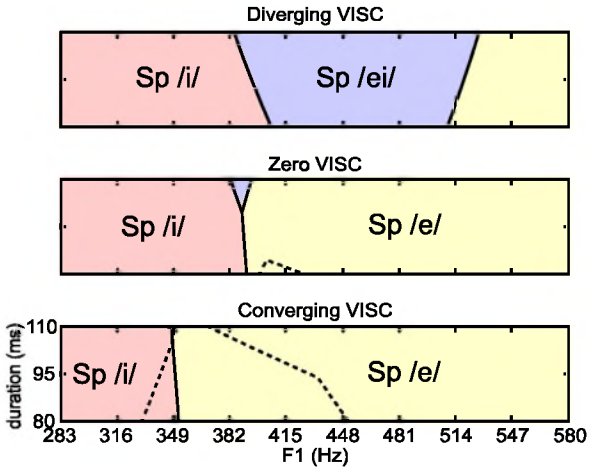

Figure 3. Territorial map of modal response areas from logistic-regression model of Peninsular-Spanish listeners' identification of synthetic stimuli. Reproduced from Morrison (2008b).

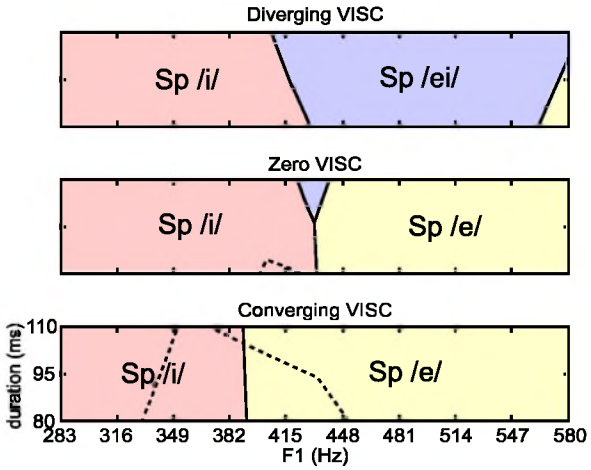

Figure 4. Territorial map of modal response areas from logistic-regression model of Mexican-Spanish listeners' identification of synthetic stimuli. Reproduced from Morrison (2008b).

classification rates for identification of Dutch /ɔ/, /ɑ/, /a/, /ɛ/, and /ʏ/. Also, as discussed below, Morrison (2008b) found evidence suggesting that there is a difference in monolingual Peninsular-Spanish and Mexican-Spanish listeners' vowel perception which could affect their perception of the English /i/–/ɪ/ contrast.

Morrison (2008b) tested the perception of a set of synthetic vowel tokens by monolingual Canadian-English, Peninsular-Spanish, and Mexican-Spanish listeners (19, 17, and 20 listeners respectively from western Canada, north-central Spain, and Mexico City). The synthetic vowels were imbedded in the word /bVpa/ in English and Spanish carrier sentences (the final non-stressed Spanish /a/ was acceptable as a schwa for the English listeners). Fig. 1 shows the duration and spectral values of the synthetic vowels in the stimulus set. Initial F2 covaried with initial F1 (F2 values were 2090, 2050, 2010, 1970, 1930, 1890, 1850, 1810, 1770, 1730 Hz). F1 and F2 either diverged [−99 Hz, +120 Hz], stayed flat [±0 Hz, ±0 Hz], or converged [+99 Hz, −120 Hz] over the timecourse of the vowel (diverging, zero, or converging vowel inherent spectral change, VISC), with the trajectory being a straight line in a log-hertz by milliseconds space. The durations given in Fig. 1 include 25 ms of consonant transitions. The filled circles in Fig. 1 represent the stimuli selected in pilot tests as the best exemplars of English /i/, /ɪ/, /e/, and /ɛ/, and Spanish /i/, /ei/, and /e/. The 90 stimuli were played in random order, each stimulus presented between 2 and 6 times to each listener (stimuli were selected using an adaptive procedure, Morrison, 2006a, for a total of 360 trials per listener). On each trial the listener identified the stimulus as one of English /i/, /ɪ/, /e/, or /ɛ/ if the listener was a monolingual English speaker, or one of Spanish /i/, /ei/, or /e/ if the listener was a monolingual Spanish speaker. Logistic regression models were fitted to the listeners' responses and territorial maps showing the modal response areas for each vowel category were made on the basis of these models, see Figs. 2–4. In Figs. 3 and 4 the modal response area for English /ɪ/ from Fig. 2 is superimposed on the Spanish /i/ and /e/ modal response areas (dashed lines).

For Peninsular-Spanish listeners the boundary between their modal response areas for Spanish /i/ and /e/ fell close to the boundary between the Canadian-English listeners' modal response areas for English /i/ and /ɪ/, but for Mexican-Spanish listeners the boundary between their modal response areas for Spanish /i/ and /e/ fell in the middle of the Canadian-English listeners' modal response

area for English /ɪ/. This leads to the hypothesis that Peninsular-Spanish listeners will assimilate most tokens of Canadian-English /i/ to Spanish /i/ and most tokens of Canadian-English /ɪ/ to Spanish /e/ – a two category assimilation. If this is the case then Peninsular-Spanish learners of Canadian English could transfer their Spanish /i/–/e/ boundary and have little difficulty perceiving and learning the English /i/–/ɪ/ contrast (there are anecdotal reports that this is the case). In contrast, Mexican-Spanish listeners are hypothesized to assimilate most tokens of Canadian-English /i/ to Spanish /i/, some tokens of Canadian-English /ɪ/ to Spanish /i/, and some tokens of Canadian-English /ɪ/ to Spanish /e/ – likely a mixture of category-goodness-difference assimilation and two-category assimilation. If this is the case then Mexican-Spanish learners of Canadian English could be expected to have substantial difficulty learning the English /i/–/ɪ/ contrast.

The present study tests whether the difference between monolingual Peninsular-Spanish and Mexican-Spanish listeners' Spanish /i/–/e/ boundary for synthetic vowels found in Morrison (2008b) is also manifested as a difference between their perception of natural tokens of Canadian-English /i/ and /ɪ/. Perception of an expanded set of Canadian-English vowels /i/, /ɪ/, /e/, and /ɛ/ is tested. Monolingual Canadian-English listeners' perception of the Canadian-English vowels are tested as a control. Also as controls, Peninsular-Spanish listeners' perception of tokens of Peninsular-Spanish /i/, /ei/, and /e/ is tested, and Mexican-Spanish listeners' perception of tokens of Mexican-Spanish /i/, /ei/, and /e/ is tested. To further explore whether there is a difference in Peninsular- and Mexican-Spanish listeners' Spanish /i/–/e/ boundary, Mexican-Spanish listeners were also tested on tokens of Peninsular-Spanish /i/, /ei/, and /e/. Canadian-English listeners' were also tested on tokens of Peninsular-Spanish /i/, /ei/, and /e/.

2. METHODOLOGY

2.1 Acoustic stimuli collection

Acoustic stimuli consisted of recordings of Spanish /i/, /ei/, and /e/ and English /i/, /ɪ/, /e/, and /ɛ/ vowels produced by monolingual Spanish and monolingual English speakers respectively.

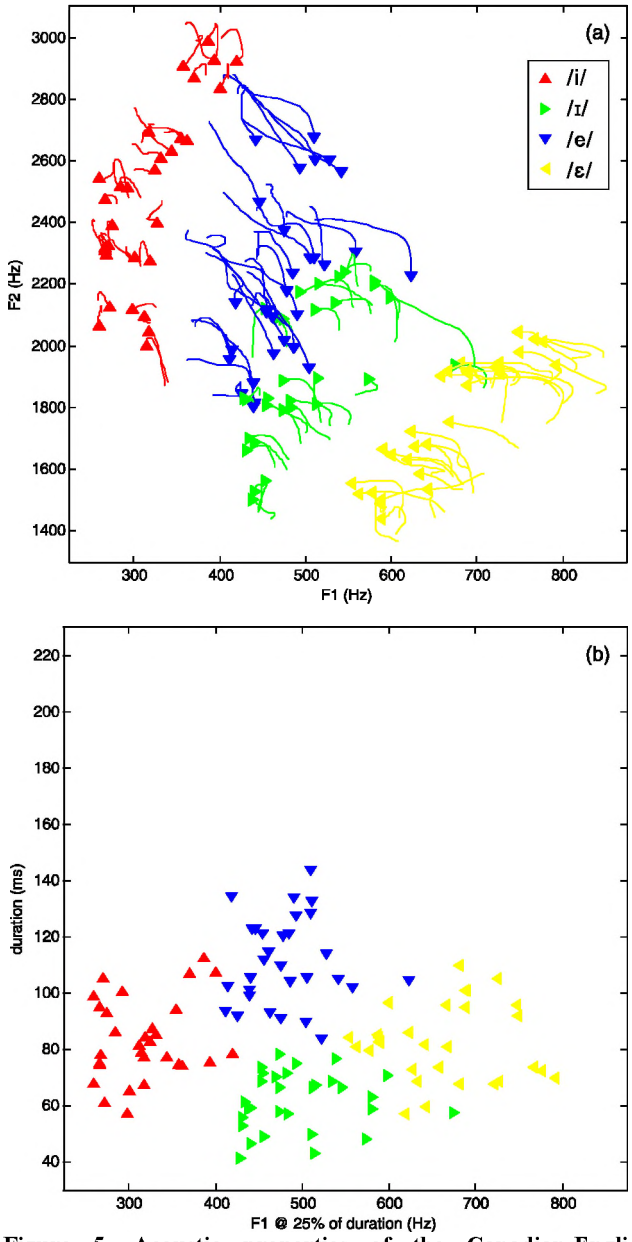

Figure 5. Acoustic properties of the Canadian-English speakers' vowel tokens used as stimuli in the present study. In the comet plots in (a), F1 and F2 at 25% of the duration of the vowel are represented by the symbol, and the trajectory from 25% to 75% of the duration of the vowel is represented by the tail. In (b) the symbol represents F1 at 25% of the duration of the vowel (*x*-axis) and the duration of the vowel (*y*-axis).

2.1.1 Speakers

Nineteen monolingual English speakers (8 male, 11 female) were recruited in Edmonton, Alberta, Canada. They came from western Canada (Alberta and Saskatchewan), and ranged in age from 18 to 54. None reported knowledge of any language other than English.

Seventeen monolingual Spanish speakers (8 male, 9 female) were recruited in Vitoria-Gasteiz, Autonomous

Region of the Basque Country, Spain (Vitoria-Gasteiz is traditionally a monolingual Spanish speaking part of the Basque Country). They came from several regions in north-central Spain (The Basque Country, Navarre, Burgos, León, and Madrid), and ranged in age from 25 to 53. None reported knowledge of any language other than Spanish beyond the level *some* from a choice of *a-little*, *some*, *well*, and *near-native*, and reported being unable to hold a conversation in any language other than Spanish.

Thirty five monolingual Spanish speakers (17 male, 18 female) were recruited in Mexico City. They came from Mexico City and the surrounding area and ranged in age from 18 to 31. None reported knowledge of any language other than Spanish beyond the level *some*, and reported being unable to hold a conversation in any language other than Spanish.

Potential participants who reported hearing or speech impediments were not included in the study.

2.1.2 Prompts

For the Canadian-English speakers, prompts consisted of written sentences "The next word is ____", and the prompt words were *BEEPA*, *BIPPA*, *BAYPA*, and *BEPPA* corresponding to /bipə/, /bɪpə/, /bepə/, and /bɛpə/. For the Peninsular-English speakers, prompts consisted of written sentences "La próxima palabra es ____" ("The next word is ____"), and the prompt words were *BIPA*, *BEPA*, and *BEIPA* corresponding to /bipa/, /bepa/, and /beipa/. The essentially-identical consonant contexts result in possible but non-existent words in the tested dialects of both languages. For the Mexican-Spanish speakers, prompt words were the same as for the Peninsular-Spanish speakers, but the written sentences were "En ____ tienes ____" ("In ____ you have ____"). The prompt word occurred in both lacunae but only the second reading of the prompt word was used in the present study. The Mexican-Spanish speakers also responded to prompt words including other vowel phonemes and other consonant contexts.

2.1.3 Procedure

Prompts were presented and responses recorded using custom-written software (a revised version of the software, *Acoustic recording software for speech production experiments*, is available from the author's website: http://geoff-morrison.net/). The monolingual English speakers completed the spelling-to-sound-correspondence

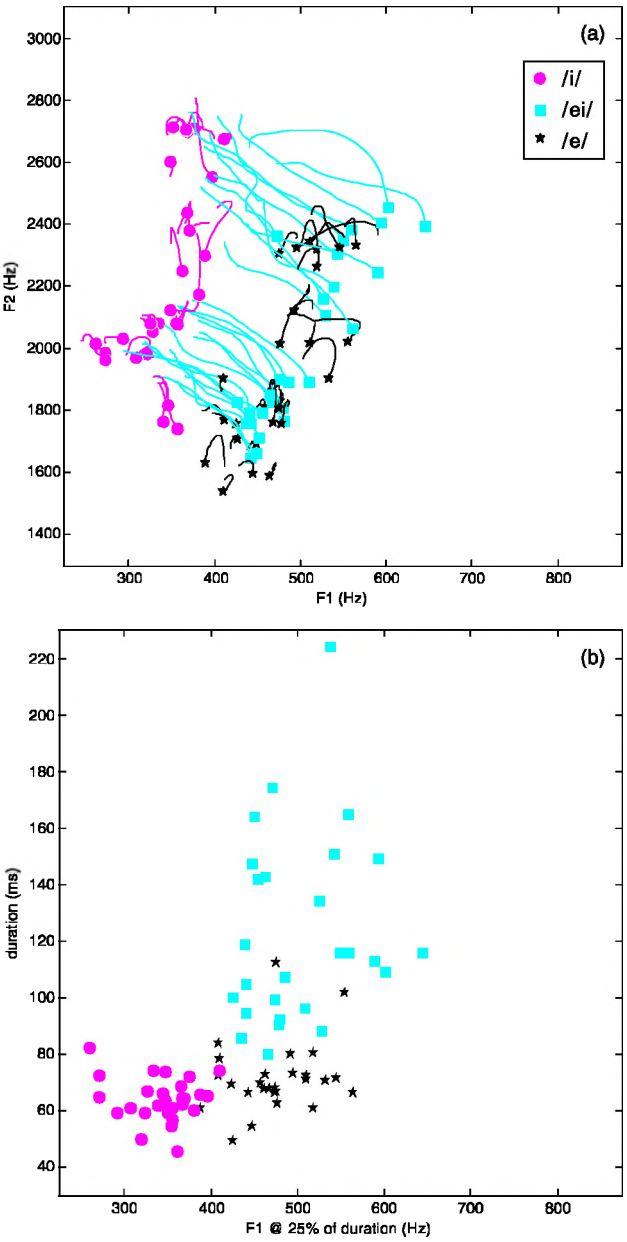

Figure 6. Acoustic properties of the Peninsular-Spanish speakers' vowel tokens used as stimuli in the present study.

training first. Speakers saw the prompt sentences and practiced reading them out loud until the researcher was confident that they could read the sentences smoothly without stumbling over the prompt words. Each prompt sentence was presented multiple times in randomized blocks (ten times for the Canadian-English and Peninsular-Spanish speakers, and three times for the Mexican-Spanish speakers, who also produced responses to a number of other prompts). The speaker heard a beep, saw a prompt sentence on a computer screen, and read the sentence out loud. The researcher monitored the recordings, and rejected recordings with problems such as stuttering, extraneous noise, and

clipping. Prompts corresponding to rejected recordings were repeated in randomized order at the end of each block. Prior to the experiment, listeners saw the instructions on the screen and heard them read out.

Recordings were made in sound booths at a sampling frequency of 44.1 kHz using a Sennheiser HMD 280 PRO headset and a Roland ED UA-30 USB Audio Interface with a Rolls MP13 preamplifier for the Canadian-English and Peninsular-Spanish speakers, and a Sennheiser HSP 2 head-mounted microphone with P48 XLP adapter and an Edirol UA-25 USB Audio Interface for the Mexican-Spanish speakers.

2.2 Acoustic stimuli preparation

Ten Canadian-English speakers (5 males and 5 females) were randomly selected. The speakers selected ranged in age from 19 to 28. Three recordings of each stimulus word were randomly selected from the recordings produced by each of these speakers. For each recording, the /bVpa/ was extracted from the sentence and normalized to 99% peak amplitude.

The above procedure was repeated for 9 randomly-selected Peninsular-Spanish speakers (5 males and 4 females aged from 34 to 50, data from an additional female speaker could not be used because of technical problems), and 10 randomly-selected Mexican-Spanish speakers (5 males and 5 females aged from 18 to 30).

F1 and F2 trajectories, and the durations of the natural stimuli are shown in Figs. 5–7. To allow for comparison, the synthetic stimuli from Morrison (2008b) are plotted in Fig. 8 on the same axes and using the same scale as in Figs. 5–7. The spread of natural stimuli along the diagonal of positive correlation between F1 and F2 can be explained by vocal-tract length differences between the speakers. Different phonemes are spread along the diagonal of negative correlation between F1 and F2.

The Mexican-Spanish speakers' /e/ tokens generally have higher F1 and lower F2 compared to those of the Peninsular-Spanish speakers. This difference is consistent with the difference in the location of the Spanish /i/–/e/ boundary found for the perception of synthetic stimuli in Morrison (2008b). Also consistent with the perception of synthetic stimuli in Morrison (2008b), the Peninsular-Spanish speakers' /e/ tokens have similar acoustic properties to the Canadian-English speakers' /ɪ/ tokens (although the English /ɪ/ tokens are generally shorter than the Spanish /e/ tokens, and generally have falling F1 and rising F2, which is not the case for the Spanish /e/ tokens).

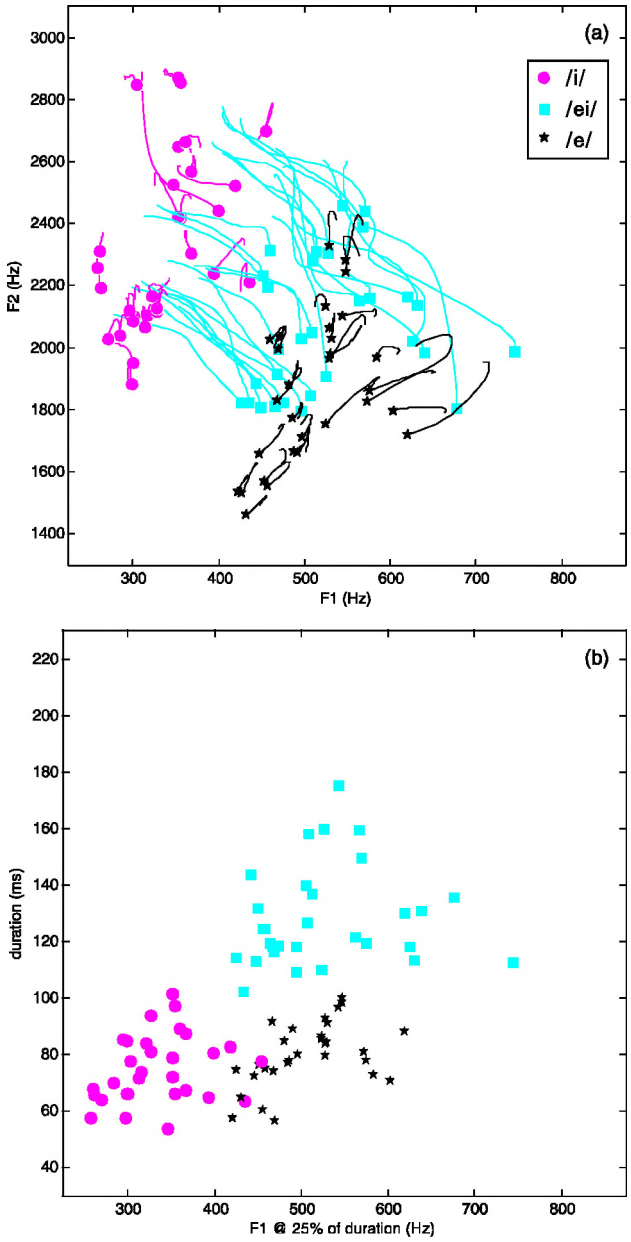

Figure 7. Acoustic properties of the Mexican-Spanish speakers' vowel tokens used as stimuli in the present study.

2.3 Listeners

Eleven monolingual Canadian-English listeners (1 male, 10 female) were recruited in Edmonton, Alberta. They came from western Canada (Alberta and Saskatchewan), and ranged in age from 18 to 27. None reported knowledge of any language other than English.

Eighteen monolingual Peninsular-Spanish listeners (11 male, 7 female) were recruited in Madrid, Spain. They all came from the Comunidad de Madrid, and ranged in age

from 18 to 27. None reported knowledge of any language other than Spanish beyond the level *a-little* from a choice of *a-little*, *some*, *well*, and *near-native*, and reported being unable to hold a conversation in any language other than Spanish.

Twenty monolingual Mexican-Spanish listeners (12 male, 8 female) were recruited in Mexico City. They all came from Mexico City and surrounding area, and ranged in age from 18 to 42. None reported knowledge of any language other than Spanish beyond the level *a-little* from a choice of *a-little*, *some*, *well*, and *near-native*, and reported being unable to hold a conversation in any language other than Spanish.

Potential participants who reported hearing or speech impediments were not included in the study.

2.4 Procedure

Listeners were tested one at a time. Testing of the Peninsular-Spanish listeners took place in a quiet conference room, and testing of the other two groups of listeners took place in sound booths. For the Spanish listeners, stimuli were presented using an Edirol UA-25 USB Audio Interface and AKG K701 headphones, and for the English listeners, stimuli were presented using a Roland Edirol UA-30 USB Audio Interface and a Sennheiser HMD 280 PRO headset.

Custom-written software was used to present stimuli and record responses In each trial, listeners heard a stimulus word and responded by clicking on the response button which corresponded to their identification of the word. A replay button allowed the stimulus to be heard up to two more times. A new stimulus was presented 750 ms after a response was given. In the Spanish experiment the response buttons were labelled *BIPA*, *BEIPA*, and *BEPA* representing /bipa/, /beipa/, and /bepa/ respectively, and in the English experiment the response buttons were labelled *BEEPA*, *BIPPA*, *BAYPA*, and *BEPPA* representing /bipə/, /bɪpə/, /bepə/, and /bɛpə/ respectively. Stimuli were presented in random order blocked by speaker. Each stimulus was presented once, except that, to allow for adaptation to each new voice, a single stimulus from the block was randomly selected and presented before the block proper. The response to this extra stimulus was not recorded. If a listener accidentally pressed a button other than their intended response button a "mistake" button was available after the presentation of the next stimulus. If the "mistake" button was pressed the current stimulus and the previous stimulus

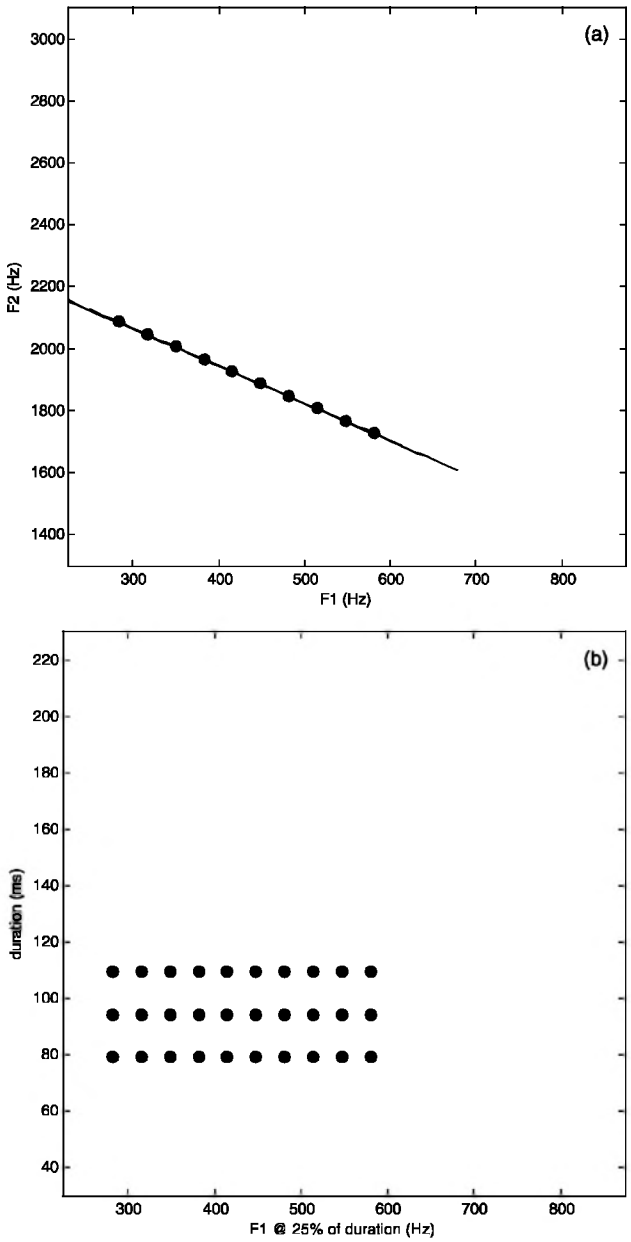

Figure 8. Acoustic properties of the synthetic vowel tokens used as stimuli in Morrison (2008b).

on which the mistake had been made were re-cued for presentation in random order at the end of the block, and a fresh stimulus was presented for identification.

Prior to the experiment, listeners saw the instructions on the screen and heard them read out. They also completed a practice experiment in which they identified stimuli spoken by a speaker of their first language whose stimuli were not included in the experiment proper (the practice speaker was a Canadian-English speaker for the monolingual English listeners and a Colombian-Spanish speaker for the monolingual Spanish listeners). Prior to the instructions and practice, monolingual-English listeners also underwent the

same spelling-to-sound correspondence training as described in §2.1.3 above.

3. RESULTS AND DISCUSSION

Tables I–III show confusion matrices for the control tests, each group of listeners identifying vowel tokens from their own first language and dialect. For all groups, the group's correct-identification rates were above 97% for all vowel categories except for Canadian-English listeners' identification of Canadian-English /ɪ/ tokens, which were identified as English /ɛ/ at a rate of 9%. A possible explanation for this could be related to the vowels having been produced in a sentence context but presented to the listeners in a word context. Another possible explanation for this could be related to a putative diachronic vowel shift in Canadian English in which (in traditional terms) /ɪ/ and /ɛ/ are lowering (Boberg, 2005; Clarke, Elms, & Youssef, 1995; Esling & Warkentyne, 1993; Hagiwara, 2006; Morrison, 2006b, §3.1) and in which perception lags behind production (Preston, 2007). One Canadian-English /ɪ/ token was an outlier and clustered with the /ɛ/ tokens, see Fig. 5. It is possible that this token was mis-spoken; however, removal of responses to this stimulus only reduced the percentage of /ɪ/ tokens identified as /ɛ/ to 7.2%.

TABLE I. Confusion matrix of monolingual Canadian-English listeners' identification of English vowel tokens produced by monolingual Canadian-English speakers. The number in each cell represent the proportion of tokens from the category given for the row identified as the category given for the column.

Produced	Classified			
	Eng /i/	Eng /ɪ/	Eng /e/	Eng /ɛ/
Eng /i/	**0.979**	0.021		
Eng /ɪ/		**0.885**	0.024	0.091
Eng /e/	0.003	0.009	**0.985**	0.003
Eng /ɛ/	0.012	0.009	0.003	**0.976**

TABLE II. Confusion matrix of monolingual Peninsular-Spanish listeners' identification of Spanish vowel tokens produced by monolingual Peninsular-Spanish speakers.

Produced	Classified		
	Sp /i/	Sp /ei/	Sp /e/
Sp /i/	**0.975**	0.023	0.002
Sp /ei/	0.004	**0.994**	0.002
Sp /e/	0.004		**0.996**

TABLE III. Confusion matrix of monolingual Mexican-Spanish listeners' identification of Spanish vowel tokens produced by monolingual Mexican-Spanish speakers.

Produced	Classified		
	Sp /i/	Sp /ei/	Sp /e/
Sp /i/	**0.988**	0.002	0.010
Sp /ei/	0.002	**0.998**	
Sp /e/	0.010	0.007	**0.983**

TABLE IV. Confusion matrix of monolingual Peninsular-Spanish listeners' identification of English vowel tokens produced by monolingual Canadian-English speakers.

Produced	Classified		
	Sp /i/	Sp /ei/	Sp /e/
Eng /i/	**0.985**	0.007	0.007
Eng /ɪ/	0.030	0.011	**0.959**
Eng /e/	0.037	**0.831**	0.131
Eng /ɛ/	0.013	0.030	**0.957**

TABLE V. Confusion matrix of monolingual Mexican-Spanish listeners' identification of English vowel tokens produced by monolingual Canadian-English speakers.

Produced	Classified		
	Sp /i/	Sp /ei/	Sp /e/
Eng /i/	**0.980**	0.015	0.005
Eng /ɪ/	0.030	0.022	**0.948**
Eng /e/	0.030	**0.837**	0.133
Eng /ɛ/	0.023	0.023	**0.953**

TABLE VI. Confusion matrix of monolingual Mexican-Spanish listeners' identification of Spanish vowel tokens produced by monolingual Peninsular-Spanish speakers.

Produced	Classified		
	Sp /i/	Sp /ei/	Sp /e/
Sp /i/	**0.974**	0.017	0.009
Sp /ei/	0.006	**0.991**	0.004
Sp /e/	0.006	0.009	**0.985**

Table IV shows the confusion matrix for the Peninsular-Spanish listeners' identification of the Canadian-English speakers' English vowel tokens. Consistent with the prediction from Morrison (2008b), English /i/ tokens were identified as Spanish /i/ at a rate of 99% and English /ɪ/ tokens were identified as Spanish /e/ at a rate of 96%.

Table V shows the confusion matrix for the Mexican-Spanish listeners' identification of the Canadian-English

speakers' English vowel tokens. English /i/ tokens were identified as Spanish /i/ at a rate of 98%, but, contrary to the prediction from Morrison (2008b), English /ɪ/ tokens were identified as Spanish /e/ at a rate of 95% and as Spanish /i/ at a rate of only 3%. Table VI shows the confusion matrix for the Mexican-Spanish listeners' identification of the Peninsular-Spanish speakers' Spanish vowel tokens. The correct-identification rate was 98%. Any difference which may exist in the location of the Spanish /i/–/e/ boundary for Mexican-Spanish versus Peninsular-Spanish listeners did not lead to a substantial difference in their perception of either the Peninsular-Spanish of the Canadian-English vowel tokens.

Table VII shows the confusion matrix for the Canadian-English listeners' identification of the Peninsular-Spanish speakers' Spanish vowel tokens. Spanish /i/ was identified as English /i/ at a rate of 94%, but identifications of Spanish /e/ were spread relatively evenly across English /e/ and /ɪ/ (44% and 38%) with a minority of /ɛ/ responses (17%). These results are consistent with the results found for Peninsular-Spanish listeners' perception of Canadian-English vowels: they almost always identified English /ɪ/ as Spanish /e/. This is consistent with Peninsular-Spanish /e/ being more spectrally similar to Canadian-English /ɪ/ than to Canadian-English /ɛ/ (see Figs. 5 and 6). Since Peninsular-Spanish /e/ has relatively little formant movement, and Canadian-English /ɪ/ and /e/ have relatively large magnitudes of formant movement but in opposite directions to each other, duration may be the primary factor determining whether the Canadian-English listeners gave English /ɪ/ or /e/ responses (see Fig. 2).

TABLE VII. Confusion matrix of monolingual Canadian-English listeners' identification of Spanish vowel tokens produced by monolingual Peninsular-Spanish speakers.

Produced	Classified			
	Eng /i/	Eng /ɪ/	Eng /e/	Eng /ɛ/
Sp /i/	**0.936**	0.040	0.024	
Sp /ei/	0.003		**0.997**	
Sp /e/	0.010	0.380	**0.441**	0.168

4. CONCLUSION

A synthetic-vowel perception experiment (Morrison, 2008b) found evidence to suggest that the location of the perceptual boundary between Spanish /i/ and /e/ differed for monolingual Peninsular-Spanish listeners (north-central Spain) and monolingual Mexican-Spanish listeners (Mexico City), and that this would affect their perception of the Canadian-English /i/–/ɪ/ contrast (western Canada): Peninsular-Spanish listeners were predicted to identify almost all tokens of Canadian-English /i/ as Spanish /i/ and almost all tokens of Canadian-English /ɪ/ as Spanish /e/ (two-category assimilation); whereas Mexican-Spanish listeners were predicted to identify almost all tokens of Canadian-English /i/ as Spanish /i/, but identify some tokens of Canadian-English /ɪ/ as Spanish /i/ and some as Spanish /e/.

The present study tested monolingual Peninsular-Spanish and monolingual Mexican-Spanish listeners' perception of natural tokens of English /i/, /ɪ/, /e/, and /ɛ/ produced by monolingual Canadian-English speakers. Consistent with the predictions from Morrison (2008b), Peninsular-Spanish listeners identified almost all tokens of Canadian-English /i/ as Spanish /i/ and almost all tokens of Canadian-English /ɪ/ as Spanish /e/; however, inconsistent with the prediction, the Mexican-Spanish listeners also identified almost all tokens of Canadian-English /i/ as Spanish /i/ and almost all tokens of Canadian-English /ɪ/ as Spanish /e/. If there is any difference between Peninsular-Spanish and Mexican-Spanish with respect to the location of the Spanish /i/–/e/ perceptual boundary, it was not found to have any substantial differential effect on monolingual Peninsular-Spanish versus Mexican-Spanish listeners' perception of natural tokens of Canadian-English /i/ and /ɪ/.

Given the caveats that the present study tested a single dialect of English and only two dialects of Spanish, and tested a single consonant context, the results call into question the assumption that L1-Spanish learners of English have difficulty learning the English /i/–/ɪ/ contrast because they assimilate most tokens of both English vowel categories to a single Spanish vowel category, Spanish /i/. The results indicate that for at least this consonant context both monolingual Peninsular-Spanish and Mexican-Spanish listeners assimilate tokens of Canadian-English /i/ and /ɪ/ to Spanish /i/ and /e/ via a two-category assimilation. On perceptual grounds, Peninsular-Spanish and Mexican-Spanish learners of Canadian-English would therefore not be expected to have difficulty learning the English /i/–/ɪ/ contrast. Given that there is evidence indicating that Mexican-Spanish learners of English do have difficulty learning the Canadian-English /i/–/ɪ/ contrast (Morrison, 2002, 2008a, 2009), one must therefore consider whether there are non-perceptual explanations for this difficulty. The

most likely non-perceptual explanations would seem to be (mis)education, students are often taught that English has a long "i" and a short "i" (Flege et al., 1997; Wang and Munro, 1999), and orthography, "i" in Spanish orthography corresponds to Spanish /i/ whereas "i" in English orthography most often corresponds to English /ɪ/ and never to English /i/ (Escudero & Wanrooij, 2010; Morrison, 2009).

ACKNOWLEDMENTS

This research was supported in part by Doctoral and Postdoctoral Fellowships from the Social Sciences and Humanities Research Council of Canada. Data collection took place between 2005 and 2010 when the author was affiliated with, sequentially, the Department of Linguistics, University of Alberta; the Department of Cognitive & Neural Systems, Boston University; the School of Language Studies, Australian National University; and the School of Electrical Engineering & Telecommunications, University of New South Wales. The author wishes to thank the speakers and listeners who volunteered to take part in the research and all those who assisted with recruitment and facilities at, sequentially, the Centre for Comparative Psycholinguistics, University of Alberta; the Phonetics Laboratory, University of the Basque Country; the Phonic Studies Laboratory, College of Mexico; the Biometric Recognition Group, Autonomous University of Madrid; and the Acoustics Laboratory, National Autonomous University of Mexico. Thanks to Vincent Arnaud for the French translation of the abstract.

REFERENCES

Álvarez González, J. A. (**1980**). *Vocalismo español y vocalismo inglés* (Spanish and English vowels), (PhD diss., Universidad Complutense de Madrid).

Best, C. T. (**1995**). "A direct realist view of cross-language speech perception," in *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange (York Press, Timonium, MD), pp. 171–204.

Bohn, O.-S. (**1995**). "Cross-language speech perception in adults: First language transfer doesn't tell it all," in *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange (York Press, Timonium, MD), pp. 279–304.

Boberg, C. (**2005**). "The Canadian shift in Montreal," Lang. Variation and Change, **17**, 133–154. doi:10.1017/S0954394505050064

Chládková, K., Escudero, P., & Boersma, P. (**2011**). "Context-specific acoustic differences between Peruvian and Iberian Spanish vowels," J. Acoust. Soc. Amer., **130**, 416–428. doi:10.1121/1.3592242

Clarke, S., Elms, F., & Youssef, A. (**1995**). "The third dialect of English: Some Canadian evidence," Lang. Variation and Change, **7**, 209–228. doi:10.1017/S0954394500000995

Escudero, P. (**2005**). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*, (PhD diss., University of Utrecht / LOT, Utrecht, The Netherlands).

Escudero, P., & Boersma, P. (**2004**). "Bridging the gap between L2 speech perception research and phonological theory," Stud. Second Lang. Acq. **26**, 551–585. doi:10.1017/S0272263104040021

Escudero, P., & Wanrooij, K. (**2010**). "The effect of L1 orthography on non-native vowel perception," Lang. Speech **53**, 343–365. doi:10.1177/0023830910371447

Escudero, P., & Williams, D. (**2012**). "Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch," J. Acoust. Soc. Amer., **131**, E406–E412. doi:10.1121/1.3701708

Esling, J. H., & Warkentyne, H. J. (**1993**). "Retracting of /æ/ in Vancouver English," in *Varieties of English Around the World, Vol. G11, Focus on Canada*, edited by S. Clarke (John Benjamins, Philadelphia, PA), pp. 229–246.

Flege, J. E. (**1991**). "The interlingual identification of Spanish and English vowels: Orthographic evidence," Quart. J. Exp. Psych. **43**, 701–731. doi:10.1080/14640749108400993

Flege, J. E., Bohn, O.-S., & Jang, S. (**1997**). "Effects of experience on non-native speakers' production and perception of English vowels," J. Phonetics **25**, 437-470. doi:10.1006/jpho.1997.0052

Godínez, M. Jr. (**1978**). "A comparative study of some romance vowels," UCLA Working Papers in Phonetics **41**, 3–19.

Hagiwara, R. (**2006**). "Vowel production in Winnipeg," Can. J. Ling. **51**, 127–142.

Møller Glasbrenner, M. (**2005**). *Vowel identification by monolingual and bilingual listeners: Use of spectral change and duration cues* (MSc diss., University of South Florida).

Morrison, G. S. (**2002**). "Perception of English /i/ and /ɪ/ by Japanese and Spanish listeners: Longitudinal results," *Proceedings of the North West Linguistics Conference 2002* edited by G. S. Morrison, & L. Zsoldos (Simon Fraser University Linguistics Graduate Student Association, Burnaby, BC), pp. 29–48.

Morrison, G. S. (**2006a**). "An adaptive sampling procedure for speech perception experiments," in *Proceedings of the Ninth International Conference on Spoken Language Processing: Interspeech 2006 — ICSLP, Pittsburgh* (ISCA, Bonn, Germany), pp. 857–860.

Morrison, G. S. (**2006b**). *L1 & L2 production and perception of English and Spanish vowels: A statistical modelling approach.* (PhD diss., University of Alberta).

Morrison, G. S. (**2008a**). "L1-Spanish speakers' acquisition of the English /i/–/ɪ/ contrast: Duration-based perception is not the initial developmental stage," Lang. Speech **51**, 285–315. doi:10.1177/0023830908099067

Morrison, G. S. (**2008b**). "Perception of synthetic vowels by monolingual Canadian-English, Mexican-Spanish, and Peninsular-Spanish listeners," Can. Acoust. **36**(4), 17–23. [Typesetting errata published in **37**(1), 34]

Morrison, G. S. (**2009**). "L1-Spanish speakers' acquisition of the English /i/–/ɪ/ contrast II: Perception of vowel inherent spectral change," Lang. Speech **52**, 437–462. doi:10.1177/0023830909336583

Morrison, G. S., & Escudero, P. (**2007**). "A cross-dialect comparison of Peninsular- and Peruvian-Spanish vowels" in *Proceedings of the 16th International Congress of Phonetic Sciences: Saarbrücken 2007* edited by J. Trouvain & W. J. Barry (Universität des Saarlandes, Saarbrücken, Germany), pp. 1505–1508.

Preston, D. R. (**2007**). "Why can't you understand your own language?" in *LACUS Forum XXXIII* edited by P. Reich, W. J. Sullivan, A. R. Lommel, & T. Griffen (Linguistic Association of Canada and the United States, Houston, TX), pp. 5–18.

Wang, X., & Munro, M. J. (**1999**). "The perception of English tense-lax vowel pairs by native Mandarin speakers: The effect of training on attention to temporal and spectral cues," in *Proceedings of the 14th International Congress of Phonetic Sciences: San Francisco 1999* edited by J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey (University of California Berkeley), pp.125–128.

Wells, J. C. (**1982**). *Accents of English* (Cambridge University Press, Cambridge, UK). Vol. 1 pp. 1–277, vol. 2 pp. 279–465, vol. 3 pp. 466–673.