

# PERCEPTUAL ASYMMETRIES AND LEXICAL EFFECTS IN L2 VOWEL DISCRIMINATION

Juli Cebrian

Universitat Autònoma de Barcelona  
juli.cebrian@uab.cat

## ABSTRACT

This paper examines the relationship between perceived similarity between first language (L1) and second language (L2) vowels and discrimination of L2 vowel pairs in light of L2 speech models' predictions on L2 vowel discrimination. The study also explores the role of word status (real vs. nonsense word) in L2 vowel discrimination. Forty-four Spanish learners of English performed a perceptual assimilation task and two L2 vowel discrimination tasks. The discrimination tasks presented the members of each pair in the two possible orders (e.g., /i/-/i/, /i/-/i/), and stimuli were embedded in real and nonsense words.

Results showed that discrimination accuracy was not always explained by crosslinguistic mapping relations as proposed by L2 speech models such as the Perceptual Assimilation Model-L2. Further, directional asymmetries emerged, mirroring previous findings in infant and adult perception. Finally, discrimination accuracy was greater in real words than in non-sense words, underscoring the influence of lexical categories in L2 perception.

**Keywords:** L2 speech perception, vowel discrimination, cross-linguistic similarity.

## 1. INTRODUCTION

Adult L2 learners' difficulty to perceive and produce target language phones accurately is explained by a number of learner factors and linguistic factors, including the amount of L2 experience and the influence of the L1 [6, 13]. Models of L2 speech perception attempt to characterize the way in which target phones are "assimilated" to (i.e., identified with) L1 sound categories and their consequences [1, 9]. According to the Native Language Magnet (NLM) model [9], native and non-native phones are perceived in terms of prototypical exemplars of L1 phones that act as perceptual magnets. These magnets attract perceptually similar non-native phones, affecting the formation of accurate L2 categories. According to the Perceptual Assimilation Model-L2 (PAM-L2) [1], likelihood of accurate discrimination of L2 phones is determined by the degree to which pairs of target phones are

assimilated to one or more L1 categories. Thus, two target phones that are assimilated to the two different L1 phones (two category assimilation) will be more accurately discriminated than two target phones that are assimilated to a single L1 category. In the latter case, the two target phones may be perceived as equally good or bad versions of the L1 category (single category assimilation) or one target phone may be perceived as a better match for the L1 category than the other target phone (category goodness assimilation). L2 discrimination is predicted to be better with the latter than with the former [1]. A number of studies have provided evidence in support of these predictions, e.g., [18] for vowel discrimination, among others.

Previous studies on vowel discrimination have also reported the existence of directional perceptual asymmetries: listeners appear to be better at discriminating between two members of a vowel contrast when the more peripheral vowel is presented second (e.g., the sequence /e/-/i/) than when the order is the reverse (/i/-/e/). According to the Natural Referent Vowel (NRV) framework [14], vowels that are more peripheral in the F1/F2 space act as natural referents or perceptual anchors in the development of vowel perception. This type of perceptual asymmetries has been found in infant perception of native (up to 12 months old) and non-native phones, and in adult perception of non-native contrasts [14]. Further, asymmetries disappear when listeners can access auditory memory, i.e., when the inter-stimulus interval (ISI) in the discrimination task stimuli is 500 ms or shorter [e.g., 10]. See [14] for a review of studies.

Finally, another factor that may affect L2 vowel perception is the role of lexical categories. Some studies have found that adult L2 learners are better at discriminating challenging L2 phones in real words than in nonwords, e.g., [11]. Similarly, Solé found that auditory priming effects triggered by confusable L2 words were more prevalent with nonwords than with real words in a lexical decision task [16]. These and other findings indicate that lexical representations play a role in the perception of segmental L2 contrasts [7, 19].

The current study had three main goals: a) to examine the relationship between categorization of target L2 phones and L2 discrimination accuracy; b)

to explore the prevalence directional perceptual asymmetries in L2 vowel discrimination; and c) to evaluate if L2 perception is affected by the lexical status of the stimuli. To that effect, a group of Spanish learners of English performed a perceptual similarity task (Experiment 1) and two discrimination tasks involving real words and nonwords (Experiment 2). These tasks were performed as part of a larger study and took place on different days. The methodology and results of each experiment are described in the next two sections.

## 2. EXPERIMENT 1. CROSS-LINGUISTIC PERCEIVED SIMILARITY

### 2.1 Participants

The participants were 44 Spanish L2 English speakers in their first year of an English Studies degree at a Spanish university (mean age = 19.4). Most had started learning English as children at school, and few had spent time in an English-speaking country beyond short summer visits, according to a language background and use questionnaire. The L1 in the perceptual assimilation task (PAT) was Spanish as all participants spoke Spanish even if many were Spanish-Catalan bilinguals. In fact, no differences were observed in the results of the PAT related to the language dominance indicated on the questionnaire, and the PAT results replicated earlier findings, as discussed below.

### 2.2 Perceptual assimilation task

In a perceptual assimilation task [8, 17, among others], listeners are presented with non-native stimuli and have to identify them in terms of native categories and provide goodness of fit ratings. The stimuli for this task consisted of nine Southern Standard British English (SSBE) vowels (/i: ɪ ε ɜ: æ ʌ ɑ: aɪ eɪ/) produced by three male native SSBE speakers in bVt sequences. On each trial, listeners responded by selecting one of eight options representing the Spanish vowels /i e a o u/ and the diphthongs /ai ei oi/,<sup>1</sup> and provided a goodness of fit rating on a 7-point Likert scale. The total number of trials was 108 (9 vowels x 3 talkers x 2 tokens x 2 repetitions). The task was conducted using Praat [2].

### 2.3 Results

The percentage of times that each English vowel was identified as a given Spanish vowel was calculated, together with the corresponding average goodness of fit rating. Table 1 presents a confusion matrix showing the assimilation percentages (i.e.,

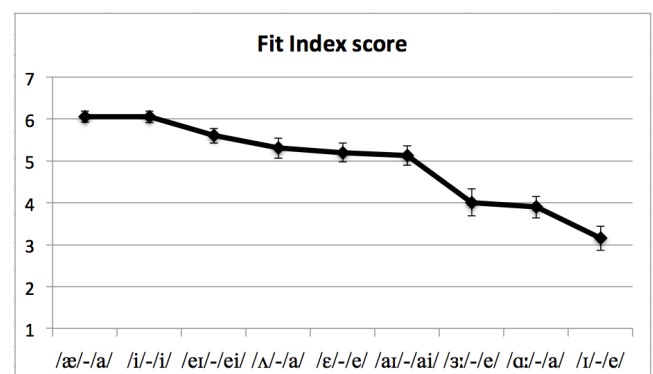
identification as an L1 vowel) and the goodness ratings (GR). The modal responses are highlighted in bold. As can be observed, a few SSBE vowels were perceived as very close to their Spanish counterparts, such as English /æ, ε, i:/ to Spanish /a, e, i/, respectively, followed by /eɪ, ʌ, aɪ/ to Spanish /ei, a, ai/, with assimilation rates above 80% and GR greater than 5. SSBE /ɑ:, ɜ:/ were mostly perceived as Spanish /a, e/ with lower GR (77-73%, GR: 4.9-4.2), while /ɪ/ obtained the lowest assimilation rates (65% as Spanish /e/, 34% as Spanish /i/). These results mirror previous findings involving native Spanish speakers [5].

**Table 1:** Percentage assimilation of SSBE vowels to Spanish vowels and goodness of fit ratings.

		English Stimuli								
		/æ/	/ɑ:/	/aɪ/	/ɛ/	/eɪ/	/ɜ:/	/ɪ/	/i:/	/ʌ/
Spanish Responses	/a/ %	<b>98</b>	<b>77</b>					14		<b>90</b>
	GR	<b>6.1</b>	<b>4.9</b>					2.9		<b>5.8</b>
/ai/ %				<b>84</b>	9					
	GR			<b>6.0</b>	5.6					
/e/ %				<b>99</b>		<b>73</b>	<b>65</b>			
	GR			<b>5.2</b>		<b>4.2</b>	<b>5.9</b>			
/ei/ %					<b>90</b>					6
	GR				<b>6.1</b>					4.4
/i/ %							34	<b>93</b>		
	GR						5.5	<b>6.5</b>		
/o/ %			20				11			8
	GR		5.0				4.3			5.2
/oi/ %				15						
	GR			5.8						
/u/ %										
	GR									

Following [8], a composite fit index score (FI) for each pair was calculated by multiplying the identification score by the goodness rating in order to take both measures into account. Fig. 1 displays the FI for each modal L2-L1 assimilation, ordered from highest (closest similarity between L2 and L1 phone) to lowest score: /æ/-/a/, /i:/-/i/, /eɪ/-/ei/, /ʌ/-/a/, /ɛ/-/e/, /aɪ/-/ai/, /ɜ:/-/e/, /ɑ:/-/a/, and /ɪ/-/e/.

**Figure 1:** Fit index scores obtained for each modal response in the PAT.



## 2.4 Discussion

Given the high assimilation rates (FI of 5 or higher) obtained for SSBE /æ/, /i:/, /e/, /ʌ/, /ɛ/ and /a/ to Spanish /a/, /i/, /ei/, /a/ and /ai/, respectively, it may be difficult for learners to detect differences between the target and the L1 vowels. Thus these vowels are likely to be perceived and produced in terms of the L1 categories, at least initially [6]. The SSBE vowels /ɜ:/, /ɑ:/, and /ɪ/ obtained FIs (as Spanish /e, a, e/, respectively) of 4 or lower. Differences between these vowels and native vowels may be more readily detected by learners given enough input and experience, and eventually they may be categorized more authentically [6].

The results illustrate several cases where two or more vowels were assimilated to a single L1 vowel. In PAM-L2's terms [1], English /æ/ and /ʌ/ exemplify a single-category (SC) assimilation as the two SSBE vowels were strongly assimilated to Spanish /a/. This scenario predicts difficulty of discrimination. SSBE /ɑ:/ was also predominantly assimilated to Spanish /a/ but with notably lower scores than /æ/ or /ʌ/, resulting in PAM-L2's category goodness (CG) assimilation (/ɑ:/ vs. /æ, ʌ/ with regards to Spanish /a/). Similarly, SSBE /ɛ/ is a much closer match to Spanish /e/ than SSBE /ɪ/ and /ɜ:/ are, also constituting CG assimilations. PAM-L2 suggests greater likelihood of accurate discrimination between two members of a CG assimilation (e.g., /ɛ/ and /ɜ:/ with respect to Spanish /e/) than between two members of a SC assimilation (e.g., /æ/ and /ʌ/ with respect to Spanish /a/).

Given these results four pairs of target language vowels were selected to test the predictions: SSBE /æ-ʌ/, /ɜ:-e/, /ɜ:-ɑ:/ and /ɪ-i:/. Only four pairs were selected as they allowed testing for different scenarios and also for the sake of brevity as this experiment formed part of a larger study including a variety of other measures. Two of the pairs involved two target vowels assimilated to the same L1 vowels, namely, a SC assimilation (/æ-ʌ/ to Spanish /a/), and a CG assimilation (/ɜ:-e/ to Spanish /e/). Discrimination of the former is predicted to be poorer than of the latter. The other two pairs involved cases where each target vowel was assimilated predominantly to a different L1 vowel (TC assimilations): /ɪ/-i:/ (assimilated to Spanish /e/ and /i/, respectively, but with different degrees of assimilation), and /ɜ:-/ɑ:/ to Spanish /e/ and /a/, both with low FIs. Note that /ɪ/ was assimilated to Sp. /e/ 65% of the time, that is, below 70%, which has been proposed as the threshold for categorization [e.g., 18]. This would make the /ɪ/-i:/ pair an uncategorized-categorized (UC) type of assimilation.

Discrimination of both TC and UC cases is expected to be good.

In terms of directionality, each pair allowed us to test if listeners would perform differently depending on the order in which the stimuli were presented: the pairs /ʌ-æ/, /ɜ:-e/, /ɜ:-ɑ:/ and /ɪ-i:/ illustrate sequences from a less to a more peripheral vowel position, while the sequences /æ-ʌ/, /e-ɜ:/, /ɑ:-ɜ:/ and /i:-ɪ/ imply going from a more to a less peripheral vowel.

## 3. EXPERIMENT 2. DISCRIMINATION OF L2 VOWELS

### 3.1 Participants

The participants were same as in Experiment 1.

### 3.2. Discrimination task

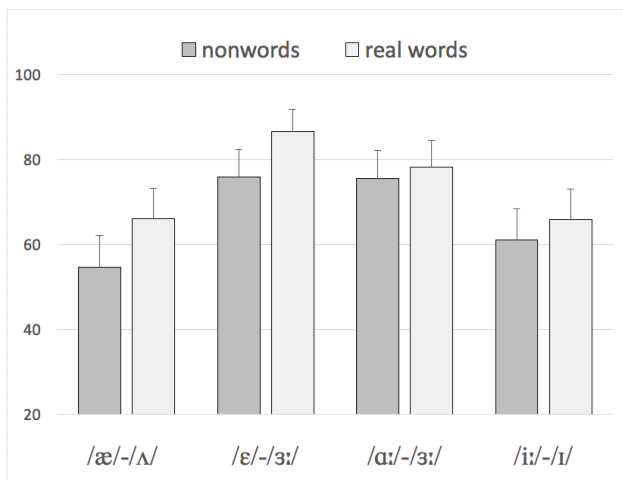
The discrimination test was a categorical AX same/different task involving the SSBE vowel pairs /æ-ʌ/, /ɜ:-e/, /ɜ:-ɑ:/ and /ɪ-i:/, which constituted the different-category pairs, and the corresponding same-category pairs. The stimuli were from a previous study [3] and consisted of 32 real words and 32 non-words elicited from two native speakers of SSBE (1 female, 1 male). All word stimuli were CVC words where both consonants were obstruents. The inter-stimulus interval (ISI) was 1.15 s. For each vowel pair, there were two talker combinations (T1-T2, T2-T1) and two possible orders (e.g., /æ-ʌ/, /ʌ-æ/). The total number of trials was 96 (48 same-category trials, 48 different-category trials). The tasks were administered using Praat [2].

### 3.3 Discrimination results

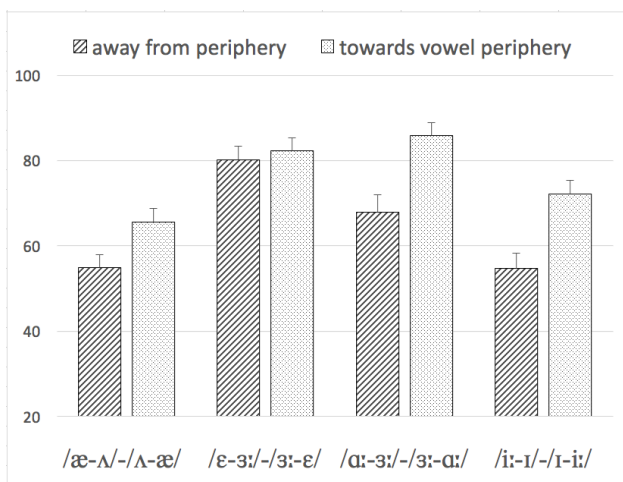
Discrimination results were analyzed in terms of percentage correct responses and are shown in two separate figures: Fig. 2 shows the results per vowel pair and type of word, Fig. 3 presents the results per vowel pair and direction of the contrast. A GLMM analysis was conducted on the accuracy scores with word type (real vs. nonword), vowel pair (4 pairs) and direction of contrast (more to less peripheral, less to more peripheral) as fixed factors and examining the potential interactions. The analysis yielded a main effect of word type [ $F(1, 691) = 24.67, p < .001$ ], vowel pair [ $F(3, 691) = 47.29, p < .001$ ] and direction of contrast [ $F(1, 691) = 66.72, p < .001$ ]. There was a significant vowel pair by direction of contrast interaction [ $F(3, 691) = 6.15, p < .001$ ], but the interactions involving type of word did not reach significance. Thus, discrimination was significantly more accurate for real words than for nonwords (74% vs. 67%), and a directionality from

less to more peripheral resulted in significantly greater accuracy (77% vs. 64%). Pairwise comparisons revealed that /e-ɜ:/ and /ɑ:-ɜ:/ were more accurately discriminated than /æ-ʌ/ and /i:-i/, and that directionality was significant for /æ-ʌ/, /i:-i/ and /ɑ:-ɜ:/, but not for /e-ɜ:/.

**Figure 2:** Discrimination accuracy per vowel contrast and word type.



**Figure 3:** Discrimination accuracy per vowel contrast and direction of contrast.



### 3.4 Discussion

The results of the discrimination tasks provide evidence of the effects of type of word and direction of contrast. There was no interaction involving type of word, indicating that the target vowels were always better discriminated when presented in real words than when embedded in nonwords. These results are in agreement with previous studies suggesting that the perception of nonnative phonemic contrasts is facilitated by lexical knowledge and word familiarity [11, 19], and underscoring the role of lexical representations in the perception of segmental contrasts [7, 16].

The overall significant effect of direction of contrast illustrates a perceptual asymmetry in line with previous findings involving infant and adult L2 perception and lends support to the NRV framework [14]. An asymmetry was found with /æ-ʌ/, /i:-i/ and /ɑ:-ɜ:/, but not in the case of /e-ɜ:/. Tyler et al. [18] hypothesize that asymmetries may not be expected with cases of TC, as this type of assimilation would behave as a native phonological contrast. On the other hand, cases of SC or CG assimilations, where both target phones are perceived as belonging to the same L1 category, would be amenable to asymmetries as listeners have to rely on phonetic information. Our results do not seem to go in this direction as asymmetries were more evident with the TC pairs (/i:-i/ and /ɑ:-ɜ:/) than with the other pairs.

Regarding PAM-L2's predictions, the greater accuracy scores obtained for the TC pair /ɑ:-ɜ:/ and the CG pair /e-ɜ:/ relative to the lower scores for the SC /æ-ʌ/ go in the expected direction. The results for /i:-i/, however, do not conform to the predictions. Being a TC (or a UC) assimilation, we would have expected this pair to be better discriminated than the SC or CG pairs. Difficulty perceiving this SSBE contrast by Spanish learners of English has been previously reported [e.g., 4, 15]. Further, different L2-L1 assimilation patterns have been reported for English /i:-i/ in previous studies involving L1 Spanish/Catalan bilinguals [4, 15]. It is possible that other factors are at play regarding this contrast such as the relative weighting of different acoustic cues [12], or individual differences in the perceived similarity between L1 and L2 vowels [18].

### 4. CONCLUSIONS

This study has reported the results of two experiments exploring the effects of lexical knowledge, direction of contrast and L1-L2 perceptual relations on the ability to discriminate L2 vowels. The results support the role of lexical representations in L2 perception, as discrimination in real words was more accurate than in nonwords. Perceptual asymmetries emerged generally in accordance with the predictions of the NRV framework, and the results for three of the four pairs seemed to support the PAM-L2 predictions. Nevertheless, the study has some limitations as only a subset of possible vowel contrasts was tested, and no native SSBE speakers were tested as control data. In addition, the relationship between L2 sound categorization and discrimination at the level of the individual needs to be examined, analyzing if individual differences in L1-L2 assimilation patterns are reflected in L2 discrimination. These issues will be addressed in future research.

This research was supported by Grant No. FFI2017-88016-P from the Spanish Ministry of Economy and Competitiveness and Grant No. 2017SGR34 from the Catalan Government.

## REFERENCES

- [1] Best C.T., Tyler M.D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Bohn, O-S, Munro, M. J. (eds), *Language Experience in Second Language Speech Learning*. Amsterdam: John Benjamins, 13–34.
- [2] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [computer program]. Version 6.0.43, retrieved from <http://www.praat.org/>.
- [3] Carlet, A. 2017. *L2 perception and production of English consonants and vowels by Catalan speakers: the effects of attention and training task in a cross-training study*. Unpublished Doctoral Dissertation. Barcelona: Universitat Autònoma de Barcelona.
- [4] Cebrian, J. 2006. Experience and the use of duration in the categorization of L2 vowels. *Journal of Phonetics*, 34, 372-387.
- [5] Cebrian, J. 2019. Perceptual assimilation of British English vowels to Spanish monophthongs and diphthongs. *Journal of the Acoustical Society of America*, 145 (1), EL52-EL58.
- [6] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. In: W. Strange (ed), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, York Press, 223-277.
- [7] Ganong, W. F. 1980. Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6(1):110-125.
- [8] Guion, S. G., Flege, J. E., Akahane-Yamada, R., Pruitt, J. S. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *JASA*, 107, 5, 1, 2711–2724.
- [9] Kuhl, P. K., Iverson, P. 1995. Linguistic experience and the "perceptual magnet effect." In: Strange, W, (ed.) *Speech perception and linguistic experience: Issues in cross-language research*. (pp. 121-154). York Press: Baltimore, MD.
- [10] Larsson, J.P., Constán, F. V., Sebastián-Gallés, N., Deco, G. 2008. Lexical plasticity in early bilinguals does not alter phoneme categories: I. Neurodynamical modeling. *Journal of Cognitive Neuroscience*, 20, 76-94.
- [11] Mora, J. C. 2005. Lexical knowledge effects on the discrimination of non-native phonemic contrasts in words and nonwords by Spanish/Catalan bilingual learners of English. In *Proceedings of ISCA Workshop on Plasticity in Speech Perception*, 43-46.
- [12] Morrison, G. S. 2009. L1-Spanish speakers' acquisition of the English /i/-/ɪ/ contrast II: Perception of vowel inherent spectral change. *Language and Speech*, 52(4), 437–462.
- [13] Piske, T., I.R.A. MacKay & J.E. Flege. 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, 191-215.
- [14] Polka, L., Bohn, O.-S. 2011. Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, 39(4), 467-478.
- [15] Rallo Fabra, L., Romero, J. (2012). Native Catalan learners' perception and production of English vowels. *Journal of Phonetics*, 40(3), 491–508.
- [16] Solé, M. J. 2013. Lexical vs phonological categories in L2 learners. Paper presented at the Phonetics and Phonology in Iberia 2013 conference. Lisbon, Portugal.
- [17] Strange, W. 2007. Cross-language similarity of vowels. Theoretical and methodological issues. Bohn, O. S., Munro M. J. (eds). *Language Experience in Second Language Speech Learning*. John Benjamins, 15-34.
- [18] Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. 2014. Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, 71(1), 4–21.
- [19] Yamada, R. A., Tohkura, Y., Kobayashi, N. 1997. Effect of word familiarity on non-native phoneme perception: Identification of English /r/, /l/, and /w/ by native speakers of Japanese. In James, A. and Leather, J. (eds.) *Second Language Speech. Structure and Process*, Mouton de Gruyter, Berlin. pp. 103-117.

---

<sup>1</sup> Monophthongs and diphthongs were included in the PAT following previous studies showing the role of diphthongs in cross-linguistic perception [5].