Minimal effects of L2 phonetic training on non-naïve learners' perception of cross-linguistic similarity

Minimal effects of training on perceived similarity

Juli Cebrian[1], Angélica Carlet[2], Núria Gavaldà[3], and Celia Gorba[1]

[1] Universitat Autònoma de Barcelona | [2] Charles Darwin University | [3] Universitat de Barcelona

Accuracy of L2 categorization is linked to the degree of perceived similarity between L1 and L2 categories. The effectiveness of high variability phonetic training in improving L2 perception is widely reported. Still, the effect of training on cross-linguistic perceived similarity is largely unexplored. Thirty-eight Catalan/Spanish bilingual learners of English underwent L2 vowel identification and discrimination training and were tested on their perceptual assimilation of English vowels. Although perceptual training improved both vowel identification and discrimination, training generally did not affect cross-linguistic perceived similarity. The only exception was a tendency for English /ɪ/ to show a decrease in assimilations to Spanish /e/ and an increase in assimilations to Spanish /i/, possibly as a result of enhanced metalinguistic knowledge about the target language. Longer training regimes or different types of tasks or stimuli may be necessary to influence the perception of cross-linguistic similarity.

Keywords: cross-linguistic perceived similarity, high variability phonetic training (HVPT), vowel categorization, L2 perception, longitudinal study

1. Introduction

L2 speech theories and speech perception models often draw on the notion of cross-linguistic (henceforth, CL) similarity to account for and make predictions about likelihood of authentic target-like categorization of target language sounds (e.g., the Perceptual Assimilation Model [PAM-L2; Best & Tyler, 2007],

https://doi.org/10.1075/jslp.25055.ceb | Published online: 18 December 2025
Journal of Second Language Pronunciation 12:1 (2026), pp. 131–145. ISSN 2215-1931 | E-ISSN 2215-194X
Available under the CC BY 4.0 license. © John Benjamins Publishing Company

the Speech Learning Model [SLM-r; Flege & Bohn, 2021], the Second Language Linguistic Perception Model [L2LP; Elvin & Escudero, 2019]). For instance, the SLM-r explains that the greater the perceived difference between native (L1) and target (L2) sounds, the more likely it is that separate (non-L1) categories are established for target sounds, and consequently are perceived and produced accurately. PAM-L2 and L2LP describe a series of assimilation patterns (i.e., L2 to L1 perceptual mappings) and learning scenarios that depend on the degree of CL similarity. Importantly, CL perceived similarity may change as a result of exposure to the target language (Best & Tyler, 2007; Flege & Bohn, 2021), so that the greater the experience with the L2, the greater the ability to distinguish L2 from L1 sounds. The goal of the current paper is to evaluate the effect of L2 experience, gained through perceptual training, on the perception of similarity between L1 (Spanish) and L2 (English) vowels.

CL similarity is often measured by means of a perceptual assimilation task (henceforth PAT), in which listeners identify stimuli in terms of L1 categories and provide a goodness-of-fit rating (e.g., Baigorri et al., 2019; Cebrian, 2019; Gong et al., 2017). For example, given an L2 stimulus like English /æ/, Spanish listeners are asked to indicate the closest L1 vowel choosing from a range of possible L1 category responses, and to provide a goodness-of-fit rating on a Likert scale (e.g., Cebrian, 2019). Some studies use a dissimilarity rating task in which listeners are presented with pairs of sounds (crucial trials involve L1-L2 pairs) and indicate on a Likert scale the degree of perceived (dis)similarity (e.g., Flege et al., 1994).

Few previous works have tested the effect of experience on CL perception. Some studies have addressed this issue by comparing naïve listeners and L2 learners, or L2 learners of different levels of experience. For instance, Ingram and Park (1997) found that Korean learners of Australian English perceived greater differences between L1 and L2 vowels than Korean monolinguals. Similarly, Levy (2009) found that L1 English naïve listeners and L2 French learners differed in their perceptual assimilation of certain French vowels, and that the naïve listeners gave overall higher goodness-of-fit ratings for L2-L1 pairs. Further, some studies report that more experienced L2 learners are more likely to judge an L2 sound as non-native than less experienced learners (Flege, 1991) or than naïve listeners (Rallo Fabra, 2005). On the other hand, Baigorri et al. (2019) reported that two groups of Spanish L1 English L2 speakers differing in age of arrival to the US provided similar L2 to L1 assimilation patterns for American English vowels, and the early bilinguals provided somewhat higher assimilation scores for modal responses (i.e., the most frequent responses), showing a more consistent L2 to L1 mapping. However, the goodness-of-fit ratings provided for the modal assimilations were significantly lower for the early bilinguals, in line with Levy (2009). This was interpreted as evidence that early L2 learning heightens the ability to distinguish L1 and L2 vowels.

Some other studies have not shown a clear effect of years of experience in the L2 setting. Cebrian (2006) found no difference between Catalan learners of English living in Canada and Catalan learners of English in their home country (undergraduate English majors). Perhaps the two types of experience had similar outcomes as neither group was naïve or inexperienced. Still, Flege et al. (1994) observed that Spanish learners of English with varying lengths of residence in the United States (1.8–7 years) performed similarly in a task in which listeners had to rate the dissimilarity between an L1 and an L2 vowel. Further, Wagner and Baker-Smemoe (2013) reported no differences between naïve listeners and experienced L2 learners in their perceptual assimilation of Q'eqchi' Mayan stops to L1 English stops. Similarly, Mayr and Escudero (2010) elicited comparable assimilation patterns between German rounded vowels and English vowels among experienced and inexperienced English learners of German. These studies seem to suggest that L2 experience may not greatly affect the perception of similarity between L1 and L2/nonnative sounds.

Clearer evidence may be drawn from studies evaluating changes in cross-linguistic perception over the course of the L2 learning process, but few studies to date have assessed CL similarity in a longitudinal approach. For example, in a study involving 15 Spanish learners of English residing in the United States, Imai et al. (2004) observed only minimal changes in CL perception over a five-year period. Namely, English /u/ was perceived to be more different from Spanish /u/ after the first year. The remaining English vowels did not differ notably in perceived similarity to a Spanish category over time.

A possible way to assess the effect of L2 experience on CL perception longitudinally is to examine the effect of specialized phonetic training on perception, particularly in a foreign language learning context, where authentic target language input may be limited. High variability phonetic training (HVPT) can be a useful resource in this context as it exposes learners to highly variable stimuli (i.e., a variety of talkers, stimuli, and phonetic contexts), and thus helps learners to abstract away from talker-specific speech characteristics and to attend to relevant cues to L2 contrasts. The results of numerous studies support the efficacy of phonetic training for enhancing the perception and production of L2 sounds (see Bradlow, 2008; Thomson, 2018; Uchihara et al., 2025, for overviews).

To our knowledge, only two studies have investigated the effect of HVPT on the perception of cross-linguistic similarity. Iverson and Evans (2009) examined the effect of five 45-minute sessions of identification training on the perceptual assimilation of L2 vowels to L1 categories, in addition to L2 identification, by 13 L1 Spanish and 13 L1 German learners of English. The results did not show changes in the modal assimilation patterns, although assimilations became more consistent (i.e., higher assimilation percentages) after training for some of the vowels. In

a study by Gong et al. (2017), twenty Mandarin Chinese speakers with no prior knowledge of Spanish were trained by means of identification tasks and tested on their identification and perceptual assimilation of 18 Spanish consonants. The results indicated that HVPT did not affect the perception of similarity of Spanish consonants that had a consistent match in the L1. By contrast, training affected cases without a clear L1 counterpart in different ways (becoming more dispersed or more consistent after training). These studies seem to indicate that L2 exposure may enhance the perception of differences between L1 and L2 sounds particularly for the most dissimilar L2 sounds.

The current study aims to assess the potential effect of different HVPT tasks on the perception of similarity between L1 and L2 vowels. CL similarity is measured by means of a PAT performed before and after training. The current paper differs from Iverson and Evans (2009) and Gong et al. (2017) in that perceptual training is not limited to identification tasks but also includes discrimination tasks. Thus, in addition to assessing the effect of HVPT on CL perceived similarity, the study contrasts the effect of the two perceptual tasks. The present study is related to a broader line of research contrasting the effectiveness of different perceptual training tasks and reporting that both identification training and discrimination training improved L2 vowel perception (Carlet & Cebrian, 2022; Cebrian et al., 2024). The discrimination task used was a categorical AX discrimination task in which same-category trials involve physically different stimuli from the same category (e.g., two stimuli involving the English vowel /æ/ produced by two different speakers). A categorical discrimination task may require listeners to identify each of the stimuli they are asked to discriminate, hence involving the same level of phonological encoding as identification tasks. Still, Cebrian et al. (2024) found that while both identification and categorical discrimination tasks improved discrimination of L2 vowels equally, identification training improved L2 vowel identification to a greater extent than discrimination training did. This was explained by methodological differences between the tasks involving the number of trials and the type of feedback provided. First, identification training consisted of twice as many trials as discrimination training, as the same number of stimuli were used in both tasks (in pairs in a categorical discrimination task, and individually in an identification task). Second, discrimination feedback indicated whether the stimuli were the same or different, while identification feedback provided the identity of the stimulus heard. The authors concluded that these differences may have resulted in a greater effect of identification training on identification. Regarding the current study, if the degree of phonological processing is the determining factor, both identification and discrimination tasks may equally affect CL perceived similarity. However, if the number of trials and type of feedback are more relevant, identification may have a greater effect than discrimination on CL perceived similarity.

2. Methodology

2.1 Participants

Thirty-eight Spanish/Catalan bilingual speakers (mean age: 19.4; 33 females), who were first-year students of English Studies at Universitat Autònoma de Barcelona, Spain, participated in the study. They had spent a maximum of two months in an English-speaking country and reported no hearing problems. The participants were randomly distributed into an identification training group (henceforth, IDG, $n=13$), a discrimination training group (DISG, $n=14$), and a control group (CG, $n=11$).

2.2 Stimuli

Training stimuli contained seven Southern Standard British English (SSBE) vowels /iː, ɪ, e, ɜː, æ, ʌ, ɑː/ found to constitute challenging vowel contrasts for native Catalan/Spanish speakers (Cebrian, 2019; Rallo Fabra, 2005). The stimuli consisted of 84 nonwords (e.g., jeet, jit, dadge, dudge), elicited from four native speakers of SSBE (2 female), sourced from a previous study (see Cebrian et al. (2024) for more details). The stimuli used for the PAT were drawn from an earlier study (Cebrian, 2019) and consisted of bVt words containing the target English vowels /iː, ɪ, e, æ, ʌ, ɜː, ɑː, aɪ, eɪ/, produced by three different male native speakers of SSBE. All recordings took place in a soundproof chamber at a university in England. To minimize crosslinguistic differences in consonant production, the bVt stimuli were edited to span from the release of the /b/ to the /t/ closure.

2.3 Procedure

Participants were tested before training (pretest) and after training (posttest), and received a small stipend as compensation. After the posttest, the control group underwent a training regime consisting of both discrimination and identification tasks and completed a second posttest afterwards (posttest2). Training for the experimental groups was carried out by means of a 7-alternative forced-choice identification task and of a categorical same/different AX discrimination task. Participants in both groups were presented with the same number of stimuli: 480 trials in each identification session (one stimulus per trial) and 240 trials in each discrimination session (a pair of stimuli per trial). They received immediate feedback (correct answer) after each individual trial as well as general feedback (% correct) at the end of each training session. The total number of training sessions was six. Testing involved identification and discrimination (not examined in this

paper, see Cebrian et al. (2024)) and perceptual assimilation, as described below (see also Table 1).

The training for the experimental groups consisted of six 30-minute sessions that took place over several weeks at the first author's university laboratory. The software used was TP (Rauber et al., 2011). Identification training consisted of a 7-alternative forced-choice identification task. Upon hearing a stimulus, listeners selected one of seven response options presented on a computer screen. The response options consisted of a vowel symbol and two example words, namely /æ/ ash/mass, /a:/ arm/palm, /e/ less/west, /3:/ earth/first, /I/ fish/his, /i:/ cheese/leaf, and /^/ sun/thus.[1] The categorical AX discrimination task used with the discrimination group included four different-category pairs: /æ/-/ʌ/, /ɪ/-/iː/, /ɜː/-/e/, /ɜː/-/ɑː/. There were 120 different-category trials and 120 same-category (e.g., /æ/-/æ/, /ʌ/-/ʌ/) trials per discrimination session. In all pairs, each word was produced by a different speaker, and the order of the vowels and the speakers was counterbalanced. Upon hearing a stimulus, listeners responded by selecting either "same" or "different" on a computer screen.

The PAT was administered before and after training. Participants were presented with English stimuli and had to identify them in terms of their L1 (Spanish) categories. They then had to provide a goodness rating on a 7-point scale where 1 represented a poor example of the vowel and 7 was a native-like example. The response options were the orthographic forms of five Spanish vowels and three diphthongs <i, e, a, o, u, ai, ei, oi> together with one monosyllabic word per vowel, also presented in written form, which were *di, se, da, do, tu, hay, rey, hoy*, respectively, meaning 'say,' 'I know,' 'give,' 'do (musical note),' 'you,' 'there is,' 'king,' and 'today'. The total number of trials was 108 (9 vowels x 3 talkers x 2 tokens x 2 repetitions). PAT was completed using Praat (Boersma & Weenink, 2018).

Table 1. Study design

	Pretest	**Training (6 sessions)**	**Posttest**	**Training (6 sessions)**	**Posttest 2**
Groups	IDG DISG CG	IDG DISG	IDG DISG CG	CG	CG
Tests/ Tasks	7FCID cat AX DIS **PAT**	ID DIS	7FCID cat AX DIS **PAT**	DIS+ID	7FCID cat AX DIS **PAT**

Note. 7FCID = 7-alternative forced choice identification task, cat AX DIS = categorical same/different AX discrimination task, PAT = perceptual assimilation task, IDG = identification training group, DISG = discrimination training group, CG = control group.

3. Results

Training had a positive effect on both L2 vowel identification and discrimination and resulted in generalization and retention of learning (see Cebrian et al. (2024)). Regarding the results of the PAT performed before and after training, the percentage of time each Spanish vowel response was selected for each English vowel stimulus and the corresponding mean goodness rating (GR) were calculated. The results are presented in Figure 1, which shows, for each English target vowel, the L1 vowels that were selected above chance level (12.5%, i.e., 100% divided by eight possible response alternatives), by each group in each test (the goodness-of-fit ratings are presented in Table A1 in the Appendix). As can be observed, at pretest, most English (E) vowels were mapped predominantly on to a single Spanish (S) vowel, namely E /æ, ʌ, ɑː/ to S /a/, E /iː/ to S /i/, E /ɪ, e, ɜː/ to S /e/, E /eɪ/ to S /ei/ and E /aɪ/ to S /ai/. Most assimilations were in fact above a categorization level of 70% (Tyler et al., 2014), except for /ɪ/, assimilated to /e/ 51% of the time and to /i/ 48% of the time by DISG, and /ɜː/ and /aɪ/, perceived as /e/ and /ai/, respectively, 68% of the time by DISG. These patterns of CL similarity replicate the results obtained in previous studies involving Spanish L1 speakers and SSBE vowels, e.g., Cebrian (2019). There appears to be minimal variation in the way English vowels were assimilated to L1 categories as a result of training, since assimilation percentages and goodness ratings remain stable across pretest, posttest, and posttest 2. The only exceptions are E /ɪ/, whose assimilations as S /e/ decreased and as S /i/ increased after training for IDG and CG, and the assimilation of E /ɑː/ to S /a/, which decreased slightly after training for DISG.

In order to examine the effect of training on the perception of similarity between L1 and L2 vowels, a linear mixed model was conducted on two dependent variables: the percentage assimilation to the modal response (in boldface in Table A1) per participant and assimilation pattern (i.e., L2 to L1 mapping), and the corresponding GR. The fixed factors were group, testing time, and assimilation pattern and their two-way interactions. The best-fitting model included a random intercept for participant. IBM Corp (2017) software was used. The results are given in Table 2. For both measures, assimilation pattern reached significance. However, neither group nor test nor any of the two-way interactions yielded a significant result. The main effects of assimilation pattern reflect differences between target vowels in their degree of assimilation to L1 categories, as predicted from previous studies (Cebrian, 2019). In brief, despite some minor differences between groups, namely DISG's lower assimilation rates for English /ɪ/ as Spanish /e/ than the other two groups, overall, the three groups patterned very similarly across vowels and testing times.

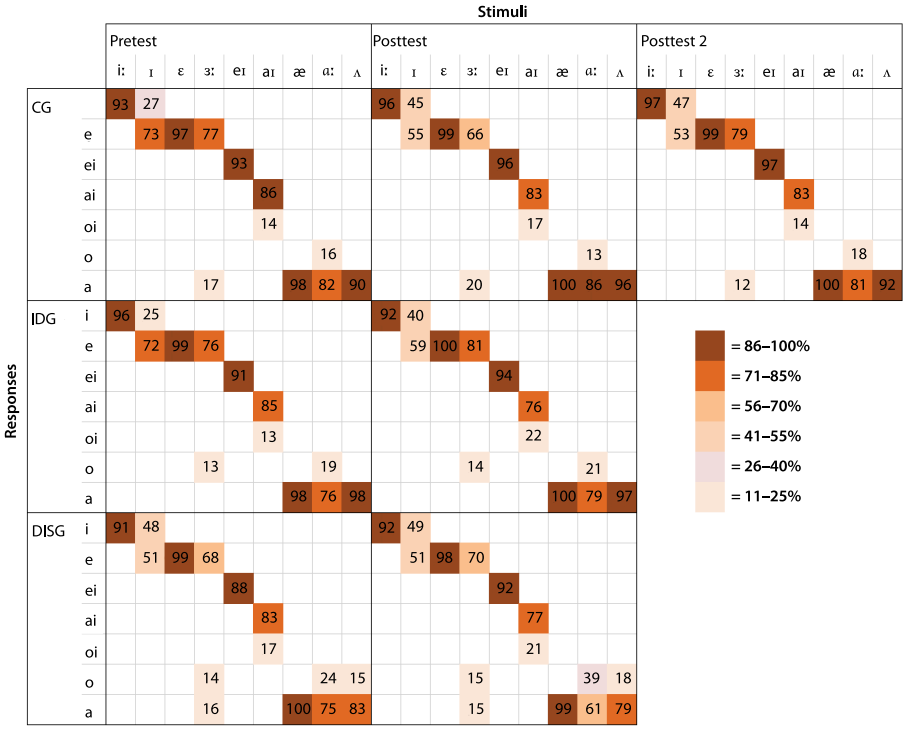


Figure 1. Percent assimilation for each English vowel stimulus (columns) to Spanish vowel responses (rows) yielded by each group at each testing time

Note: The horizontal axis shows the stimuli, the vertical axis shows the responses. IDG = identification training group, DISG = discrimination training group, CG = control group. Assimilations below chance level (12.5%) have been excluded (response <u> was chosen less than 12.5% of the time and is thus excluded).

4. Discussion and conclusions

HVPT did not seem to influence perceived similarity, as all groups performed very similarly at all testing times, except for some variation involving specific sounds like English /ɪ/, discussed below. This is true for both the L2 to L1 assimilation patterns and the GRs provided, as neither differed from pretest to posttest or posttest 2. The lack of an effect of L2 experience on CL perception is in line with some previous cross-sectional studies that compared groups differing in L2 experience (Flege et al., 1994; Cebrian, 2006; Wagner & Baker-Smemoe, 2013). However, it runs counter to previous findings that the greater the experience with the L2, the more likely it is that L2 sounds are identified as non-native (Flege, 1991; Rallo Fabra, 2005) or receive lower GRs (Levy, 2009; Baigorri

Table 2. Results of the linear mixed model. Tests of fixed effects. Assim. refers to the assimilation patterns, i.e., each target vowel and its corresponding modal L1 response

Factor/Interaction	Dependent variable	
	Percent assimilation	Goodness rating
Group	$F(2, 38.93) = 2.603$, $p = .087$	$F(2, 35.646) = .116$, $p = .891$
Time	$F(2, 537.43) = .387$, $p = .679$	$F(2, 513.32) = .422$, $p = .679$
Assim.	$F(6, 537.43) = 25.86$, $p < .001$	$F(6, 513.32) = 52.33$, $p < .001$
Group * time	$F(2, 537.43) = .060$, $p = .942$	$F(2, 513.32) = 2.33$, $p = .098$
Time * assim.	$F(12, 537.43) = .459$, $p = .938$	$F(12, 513.32) = 1.025$, $p = .424$
Group * assim.	$F(12, 537.43) = .885$, $p = .563$	$F(12, 513.32) = 1.152$, $p = .315$

et al., 2019). Regarding longitudinal research, two previous studies reported an effect of HVPT only on the consistency of assimilation, as assimilation percentages for modal responses increased for at least a subset of the sounds after training (Iverson & Evans, 2009; Gong et al., 2017). Gong et al. (2017), who tested naïve listeners, found that training had a significant effect, particularly on sounds that were not predominantly categorized as a single Chinese consonant. Both Iverson and Evans (2009) and Gong et al. (2017) tested a larger number of sounds than the current study, including several assimilations under the 70% categorization threshold (Tyler et al., 2014). The present results do not illustrate greater consistency of modal responses at posttest, perhaps because modal responses were already consistent at pretest, with most L2 vowels being assimilated to L1 vowels above 70%. Thus, the current participants may have already been too advanced for detectable changes. Such changes in assimilation patterns or consistency might instead be observed in learners with lower levels of experience. In line with Gong et al.'s (2017) findings, HVPT did not influence goodness-of-fit ratings in the current study, suggesting that training may have a limited effect on the prototypical representation of L1 categories.

The only notable variation observed in the present study involved E /ɪ/ (particularly for the IDG and CG groups), whose assimilation to S /e/ decreased after training, and the assimilation to S /i/ increased. This pattern has also been observed in previous studies (see Cebrian et al., 2021). The increase in assimilation of E /ɪ/ to S /i/, as opposed to S /e/, may be explained by the shift in perceived similarity: while assimilation may rely on acoustic-phonetic cues at initial stages of learning, it may become more phonological as learners gain experience with the L2 (Chang, 2019). This is particularly relevant in cases where perception of the target vowel is more easily influenced by metalinguistic and orthographic information, as in this case. E /ɪ/ is typically spelled with the letter <i>, which

stands for the vowel /i/ in Spanish. In addition, the /iː/-/ɪ/ contrast has a high functional load, and the two vowels are often referred to as long and short /i/ in L2 pronunciation instruction (Cebrian et al., 2021). Thus, this type of knowledge may be taken into account when judging cross-linguistic similarity, particularly once participants' attention has been drawn to the /iː/-/ɪ/ contrast through training or simply through general L2 exposure and instruction, as the performance of CG seems to indicate. Interestingly, DISG differed from CG and IDG in that there was no variation with the assimilation of E /ɪ/, which was evenly split between S /e/ and /i/ already at pretest. These differences between groups may simply be related to individual differences in the perception of similarity involving this English vowel. Cebrian (2019), for instance, found that individuals differed in the pattern of assimilation for this vowel (mostly as S /i/, mostly as S /e/ or split between /e/ and /i/). These different patterns may be due to individual differences in L1 category precision, which can affect the way L2 sounds are categorized in terms of L1 sounds (Flege & Bohn, 2021).

In any case, perceptual training had minimal or no effect on the perception of CL similarity. This finding runs counter to the expected effect of L2 experience on the ability to discern L2 from L1 sounds. Perhaps a longer training period, or a training method specifically focused on highlighting cross-linguistic differences, is necessary. Additionally, training may have a greater effect on CL perception at an earlier stage in L2 learning, as mentioned above. Alternatively, it is also possible that changes in the ability to categorize L2 sounds are unrelated to the perceived similarity between L2 and L1 sounds. In fact, as pointed out by Shinohara and Iverson (2018), HVPT studies generally demonstrate that training improves L2 perception, yet it remains unclear how training impacts L2 category formation. These authors explain that training may help learners to be more consistent and successful in using their existing categories to perceive their L2 sounds without necessarily altering the learners' internal representations of L2 categories. This is in line with Polka's (1992) observation that learners who undergo identification training may learn to identify L2 sounds accurately by paying attention to characteristics that may be helpful for differentiating non-native categories, but that may not be the properties attended to by native speakers (Shinohara & Iverson, 2018).

In conclusion, while the two perceptual training approaches (identification and discrimination) effectively improved both L2 vowel identification and L2 vowel discrimination (see Cebrian et al., 2024), both methods seemed to have little influence on the perceived similarity between L1 and L2 sounds. The only exception was the tendency observed for E /ɪ/ to experience a decrease in assimilation to S /e/ and an increase in assimilation to S /i/. This tendency may be linked to an enhanced metalinguistic awareness about the target E /iː/-/ɪ/ contrast, which may result in a more phonological perception of similarity (Chang, 2019). It is possible

that greater exposure to the target language is needed to change CL perception, or alternatively that changes in CL perception occur at earlier stages of learning. Further, a different type of training method, focused on directing learners' attention to cross-linguistic differences and similarities, may be needed to affect the perception of cross-linguistic similarity. These issues are left for future research.

Funding

This work was supported by Research Grant No. PID2021-122396NB-I00 from the Spanish Ministry of Science and Innovation, and the recognized research group 2021SGR00544 by the Catalan Agency for University and Research Grants (AGAUR).

Open Access publication of this article was funded through a Transformative Agreement with Universitat Autònoma de Barcelona.

References

Baigorri, M., Campanelli, L., & Levy, E. S. (2019). Perception of American English vowels by early and late Spanish–English bilinguals. *Language and speech*, *62*(4), 681–700.

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. Munro (Eds.), *Language Experience in Second Language Speech Learning — In honor of James Emil Flege* (pp. 13–34). John Benjamins Publishing Company.

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer* (Version 6.0.43) [Computer program]. http://www.praat.org/

Bradlow, A. R. (2008). Training non-native language sound patterns: Lessons from training Japanese adults on the English /r/-/l/ contrast. In J. G. Hansen Edwards, & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 287–308). John Benjamins Publishing Company.

Carlet, A., & Cebrian, J. (2022). The roles of task, segment type, and attention in L2 perceptual training. *Applied Psycholinguistics*, *43*(2), 271–299.

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, *34*(3), 372–387.

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.

Cebrian, J., Gorba, C., Gavaldà, N. (2021). When the easy becomes difficult: Factors affecting the acquisition of the English /i:/-/ɪ/ contrast. *Frontiers in Communication*, *6*, 1–17.

Cebrian, J., Gavaldà, N., Gorba, C. & Carlet, A. (2024). Differential effects of identification and discrimination training tasks on L2 vowel identification and discrimination. *Studies in Second Language Acquisition*, *46*, 1069–1093.

Chang, C. (2019). The phonetics of second language learning and bilingualism. *The Routledge Handbook of Phonetics* (pp.427–447). Routledge.

Elvin, J., & Escudero, P. (2019). Cross-linguistic influence in second language speech: implications for learning and teaching. In M. J. Gutierrez-Mangado, M. Martinez-Adrian, & F. Gallardo-del-Puerto (Eds.), *Cross-Linguistic Influence: From Empirical Evidence to Classroom Practice* (pp. 1–20).

Flege, J. E. (1991). The Interlingual Identification of Spanish and English Vowels: Orthographic Evidence. *The Quarterly Journal of Experimental Psychology*, *43*(3), 701–731.

Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-R). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp 3–83). Cambridge University Press.

Flege, J. E., Munro, M. J., & Fox, R. A. (1994). Auditory and categorical effects on cross-language vowel perception. *The Journal of the Acoustical Society of America*, *95*(6), 3623–3641.

Gong, J., Garcia-Lecumberri, M. L., & Cooke, M. (2017). Ab initio perceptual learning of foreign language sounds: Spanish consonant acquisition by Chinese learners. *System*, *66*, 142–155.

IBM Corp (2017). IBM SPSS Statistics for Windows. Version 25.0. Armonk, NY: IBM Corp.

Imai, S., Flege, J. E., Wayland, R. (2004, June 24–26). *Perception of cross-language vowel differences: A longitudinal study of native Spanish adults learning English* [conference poster]. 9th Conference on Laboratory Phonology, University of Illinois.

Ingram, J. C., & Park, S. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, *25*(3), 343–370.

Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America*, *126*(2), 866–877.

Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America*, *126*(5), 2670–2682.

Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism: Language and Cognition*, *13*, 279–297.

Polka, L. (1992). Characterizing the influence of native language experience on adult speech perception. *Perception & Psychophysics*, *52*(1), 37–52.

Rallo Fabra, L. (2005). Predicting ease of acquisition of L2 speech sounds. A perceived dissimilarity test. *Vigo International Journal of Applied Linguistics*, (2), 75–92.

Rauber, A. S., Rato, A., Kluge, D., & Santos, G. R. (2011). TP software.

Shinohara, Y. & Iverson, P. (2018). High variability identification and discrimination training for Japanese speakers learning English /r/–/l/. *Journal of Phonetics*, *66*, 242–251.

Thomson, R. I. (2018). High variability [pronunciation] training (HVPT): A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation*, *4*(2), 208–231.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual Assimilation and Discrimination of Non-Native Vowel Contrasts. *Phonetica*, *71*, 4–21.

Uchihara, T., Karas, M., & Thomson, R. (2025). High variability phonetic training (HVPT): A meta-analysis of L2 perceptual training studies. *Studies in Second Language Acquisition*, 1–34.

Wagner, K. O. C., & Baker-Smemoe, W. (2013). An investigation of the production of ejectives by native (L1) and second (L2) language speakers of Q'eqchi' Mayan. *Journal of Phonetics*, *41*(6), 453–467.

Appendix A

Table A1. Perceived similarity between L2 (English) vowels and L1 (Spanish) vowels and goodness-of-fit ratings

English target V	L1 response	Group	Pretest		Posttest		Posttest 2	
			% A.	GR	% A.	GR	% A.	GR
/æ/	/a/	IDG	98	6.3	100	6.6		
		DISG	100	6.1	99	6.3		
		CG	98	5.9	100	5.9	100	6.1
/ʌ/	/a/	IDG	98	5.9	97	6.2		
		DISG	83	5.9	79	5.8		
		CG	90	5.8	96	5.6	92	5.4
	/o/	DISG	15	5.4	18	5.3		
/iː/	/i/	IDG	96	6.4	92	6.6		
		DISG	91	6.5	92	6.4		
		CG	93	6.5	96	6.3	97	6.3
/ɪ/	/e/	IDG	72	5.7	59	6.1		
		DISG	51	6.3	51	6		
		CG	73	5.9	55	6.2	53	6.2
	/i/	IDG	25	4.9	40	5.7		
		DISG	48	6	49	6		
		CG	27	5.2	45	5.4	47	5.9
/ɛ/	/e/	IDG	99	5.3	100	5.9		
		DISG	99	5.3	98	5.1		
		CG	97	5	99	4.8	99	5.2

Table A1. *(continued)*

English target V	L1 response	Group	Pretest		Posttest		Posttest 2	
			% A.	GR	% A.	GR	% A.	GR
/ɜː/	/e/	IDG	76	4.2	81	3.9		
		DISG	68	4.2	70	3.4		
		CG	77	4.3	66	4	79	3.8
	/a/	DISG	16	2.5	15	2.1		
		CG	17	2.6	20	2.2	12	1
	/o/	IDG	13	3.8	14	3.2		
		DISG	14	4.7	15	2.4		
/ɑː/	/a/	IDG	76	4.6	79	4.6		
		DISG	75	5	61	5		
		CG	82	5.1	86	4.8	81	4.7
	/o/	IDG	19	5.3	21	5		
		DISG	24	5.1	39	4.8		
		CG	16	4.6	13	5.4	18	4.5
/eɪ/	/ei/	IDG	91	5.8	94	6		
		DISG	88	63	92	6		
		CG	93	6.1	96	6.3	97	6.2
/aɪ/	/ai/	IDG	85	5.8	76	5.7		
		DISG	83	6.2	77	6		
		CG	86	6	83	6	83	5.5
	/oi/	IDG	13	6.5	22	5.7		
		DISG	17	5.4	21	4.7		
		CG	13	5.8	17	5.8	14	5.5

Note: % A. = percent assimilation as L1 vowel, GR = mean goodness rating, IDG = identification training group, DISG = discrimination training group, CG = control group. Assimilations below chance level (12.5%) have been excluded.

Address for correspondence

Juli Cebrian
Departament de Filologia Anglesa i de Germanística
Universitat Autònoma de Barcelona
Facultat de Filosofia i Lletres, Edifici B, Campus UAB
Bellaterra, 08193
Spain
juli.cebrian@uab.cat
https://orcid.org/0000-0003-1547-3145

Co-author information

Angélica Carlet
Faculty of Arts and Society
Charles Darwin University
Danala Education and Community Precinct
angelica.carlet@cdu.edu.au

Celia Gorba
Departament de Filologia Anglesa i de Germanística
Universitat Autònoma de Barcelona
celia.gorba@uab.cat

Núria Gavaldà
Department of Modern Languages and Literatures and of English Studies
Universitat de Barcelona
nuria.gavalda@ub.edu

Publication history

Date received: 12 September 2025
Date accepted: 4 December 2025
Published online: 18 December 2025