Ampersand xxx (xxxx) 100109

Contents lists available at ScienceDirect

Ampersand

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

Predicting discrimination difficulty of Californian English vowel contrasts from L2-to-L1 categorization

Lucrecia Rallo Fabra [a,*], Michael D. Tyler [b]

[a] Interdisciplinary Speech & Language Acquisition Lab, Universitat de les Illes Balears, Beatriu de Pinós, despatx 6, Ctra, Valldemossa, km. 7,5, 07122, Palma de Mallorca, Spain

[b] Western Sydney University, Australia

A B S T R A C T

L1-Spanish-Catalan speakers often face difficulties when learning the English vowel system. PAM/PAM-L2 posits that native language experience shapes perception of L2 sounds, so discrimination of L2 contrasts can be predicted from L2-to-L1 categorization. The current study focused on whether discrimination of the nine Californian English vowel contrasts /i/-/ɪ/, /ɛ/-/ɪ/, /ɛ/-/ə/, /æ/-/ʌ/, /æ/-/ɑ/, /ɑ/-/ʌ/, /ə/-/ʌ/, /ə/-/ʊ/, and /u/-/ʊ/ can be accounted for by the perceptual assimilation of these vowels to L1-Catalan categories. Experienced L1-Spanish-Catalan learners of English performed categorization with goodness rating and AXB discrimination tasks. The vowels /i/, /u/, /æ/, and /ɑ/ were categorized, while /ɪ/, /ɛ/, /ʌ/, /ə/, /ɝ/, and /ʊ/ were not consistently mapped onto a single Majorcan Catalan category. Discrimination of the vowel contrasts /i/-/ɪ/, /u/-/ʊ/, /ɛ/-/ə/, and /ʊ/-/ə/ was very high (>90%). The contrasts /ɪ/-/ɛ/, /æ/-/ɑ/, and /æ/-/ʌ/ were also well discriminated (>80%). In contrast, /ʌ/-/ɑ/ and /ʌ/-/ə/ were discriminated less accurately. Percent overlap scores predicted discrimination accuracy at the group level. At the individual level, discrimination could only be predicted for the contrasts /i/-/ɪ/, /ɪ/-/ɛ/, /u/-/ʊ/, and /ɛ/-/ə/. The results are discussed in terms of three L2 speech learning models, PAM-L2, SLM-*r* and L2LP.

1. Introduction

The groundwork of categorical perception converges in that humans perceive many linguistic and non-linguistic events and objects categorically. For instance, in vision, observers can label the chromatic spectrum using four basic color terms – blue, green, yellow and red, or their combinations. Work by Eimas and colleagues (Eimas et al., 1987) provided empirical evidence that infants are born with categorical representations for speech sounds. These take the form of prototypes that guide the child's first utterances and can be later attuned via experience with specific languages. According to Medin and Barsalou (1987), categories can be divided into all-or-none categories and graded ones. The former imply that a set of features is shared by all or some members. In contrast, membership for graded categories is a matter of degree. Categories have also been found to be unstable between and within individuals (Barsalou, 1987; Medin and Barsalou, 1987) and between-category boundaries can be shifted depending on the language that is activated in the bilingual brain (e.g., Amengual, 2016; Mora and Nadeu, 2012; Williams, 1977).

For many years it has been widely accepted that language-specific patterns of perception learned in childhood were highly resistant to modification in adulthood. Adult listeners are known to have considerable difficulty perceiving non-native contrasts that are not functional in their native language (L1). Jenkins et al. (1995) referred to "perceptual foreign accents" that interfered with learning the L2 phonology. Nevertheless, perceptual reactivation of lost perceptual capacities is possible. Adult L2 learners with extensive immersion experience or intensive High-Variability-Phonetic-Training (HVPT) exhibit gains in the ability to perceptually differentiate the most difficult non-native contrasts (Bradlow et al., 1997; Carlet and Cebrian, 2019; Lively et al., 1994; Ortega et al., 2021).

1.1. Models predicting L2 speech learning

At present, three theoretical models have tried to account for L2 phonetic/phonological learning. The revised version of the Speech Learning Model (SLM-r) (Flege and Bohn, 2021) focuses on L2 learners in a naturalistic setting. Its basic tenet is that the learner's L1 and L2 phonetic systems interact in a common allophone-sensitive phonetic space. Unlike its predecessor, the SLM (Flege, 1995), the SLM-r proposes that accurate L2 perception is not a prerequisite for accurate speech production, but both abilities coevolve. Phonetic categories are understood to be mental representations that will guide the learners' L2 production and they can be established at the end stage of the L2 learn-

* Corresponding author. Interdisciplinary Speech & Language Acquisition Laboratory (ISLA-Lab), Department of Spanish, Modern & Classical Philologies University of the Balearic Islands, Beatriu de Pinós Building, Campus Univesitari Ctra, Valldemossa, km. 7,5, 07122, Palma de Mallorca, Spain.

E-mail address: lucrecia.rallo@uib.es (L. Rallo Fabra).

https://doi.org/10.1016/j.amper.2023.100109
Received 3 October 2022; Received in revised form 13 December 2022; Accepted 2 February 2023

ing process. These categories will either be new categories for sounds that do not exist in the L1 sound inventory or composite L1-L2 categories that will serve both languages. A novel aspect of the SLM-r is the "category precision" hypothesis that applies to both L2 speech perception and production. In perception, an individual's ability to discern the phonetic differences between an L1 and L2 sound will depend on the precision of her/his L1 categories. Category precision is operationalized as the "variability of acoustic dimensions measured in multiple productions of a phonetic category" (p. 36). In the case of vowels, category precision would involve little within-category variability and great between-category distances. Finally, the authors make a plea for further research testing whether differences in discernment of cross-language phonetic differences will impact the production and perception of L2 sounds.

PAM-L2 (Best and Tyler, 2007) was conceived to make predictions about the likelihood of L2 category acquisition based on how L2 phonemes are perceived at the learner's point of first contact with the L2. The model's predictions were originally proposed for functional monolinguals learning the L2 in a pervasively L2-speaking environment, but its principles are applicable to other learning situations, including in a foreign language (FL) setting (Tyler, 2019). Unlike the SLM-r, PAM-L2 considers L2 category acquisition at the phonological and phonetic levels and proposes that there is an optimal window for the perceptual learning of new L2 phonological categories. Ideally, this occurs before the learner has acquired a large vocabulary. Nevertheless, it is possible for learners with larger vocabularies to perceptually reattune to the L2 phonology (Bundgaard-Nielsen et al., 2011a, 2011b), particularly for L2 phonemes that are not perceived as clear instances on L1 categories at the initial stages of learning (Tyler, 2019). Even learners with accented L2 speech may eventually succeed at acquiring new L2 phonological categories because, for PAM-L2, there is no direct link between perception and production.

Thus, in line with prior work in L1 speech perception (Burnham et al., 2002; Strange, 2011; Werker and Curtin, 2005), PAM-L2 assumes that phonological categories may develop with vocabulary size and may be also reinforced by an alphabetic orthographic system. The likelihood of acquiring a new L2 phonological category depends on how the learner perceives L2 phone in relation to L1 categories, and also in relation to other L2 phonological categories. Considering pairs of L2 phonological categories, PAM-L2 distinguished four possible assimilation patterns of L2 categories onto L1 categories: two-category (TC), category goodness (CG), single category (SC), uncategorized-categorized (UC), and uncategorized-uncategorized (UU). More recently, Tyler (2019) elaborated on how perceived phonological overlap (Faris et al., 2016, 2018) impacts on L2 category acquisition when an L2 phoneme is initially perceived as uncategorized. Thus, two-category assimilations should be easily acquired by FL learners because the two phonemes of an L2 contrast are perceived as two distinct L1 categories. Category goodness assimilations will be challenging for the FL learner, especially if there is no direct correspondence between the written form and the phonological form. Perceptual learning for single-category assimilations will be unlikely in a FL setting. As for uncategorized phonemes, the degree of perceptual overlap will predict discrimination accuracy, which will be poor for completely overlapping contrasts and will gradually improve for partially and non-overlapping contrasts.

The revised L2LP (van Leussen and Escudero, 2015) proposes developmental trajectories for L2 learners on the basis of various learning scenarios, that is, from naïve inexperienced learners to advanced learners reaching nativelike levels of achievement. This is possible thanks to the implementation of a computational learning model inspired in the connectionist framework (Boersma, 1998). In line with PAM-L2, it focuses on sound contrasts and makes predictions on the development of L2 speech perception based on similarity. A clear difference with PAM-L2, is that in the L2LP, L2-to-L1 similarity relationships are operationalized as acoustic comparisons between the L2 contrast and the closest L1 equivalent. The model thus assumes a direct link between L2 perception and production. Another addition of the L2LP is that it extends the level of analysis to the lexical level, distinguishing between pre-lexical perception and lexical recognition. The two levels of processing can be viewed as sequential (strictly bottom-up) or interacting (allowing lexical feedback to lower-level perception).

1.2. Predicting discrimination of L2 vowel contrasts

Previous work in L2 speech perception has employed a quantitative method to operationalize predictions of L2 discrimination based on L2 categorization. For instance, Flege and MacKay (2004) used a classification overlap score to predict discrimination of English vowel contrasts from percentage of assimilation of these vowels by Italian early and late learners. The score is obtained by identifying any response categories that are selected for both vowels in a contrast and then summing the smaller of the two categorization percentages for each response category. Subsequent studies (Baigorri et al., 2019; Levy, 2009; Vasiliev, 2013) have applied the same "cross-language assimilation overlap" to predict how discrimination of English vowel pairs can be explained from assimilation patterns obtained for the same vowels. More recently, Faris et al. (2018) instead used a criterion-based approach, where contrasts were classified according to whether response categories selected above chance overlapped completely, partially, or not at all. This approach accounted for discrimination results more effectively than the overlap score, but the authors suggested that additional data should be obtained to confirm the efficacy of the approach.

Another approach to predicting discrimination difficulty of English vowel contrasts by native Spanish listeners argues that feature-based models of similarity can account for differences in L2 discrimination (Barrios et al., 2016). Specifically, they predict that Spanish advanced learners of English should discriminate the /æ/-/ɑ/ contrast more accurately than the /i/-/ɪ/ contrast because the Spanish feature [back] can be reused to distinguish /æ/ from /ɑ/ but the tense/lax distinction that characterizes the /i/-/ɪ/ contrast is absent in Spanish, therefore Spanish learners are unlikely to discriminate this contrast accurately. The results of an AX discrimination task provided no empirical support for these predictions, since no significant differences were found in the discrimination accuracy of these two contrasts.

The Natural Referent Vowel (NRV) framework proposed by Polka and Bohn (2011) accounts for differences in the discrimination of vowel contrasts from the acoustic standpoint. The authors argue that peripheral vowels have a special status in vowel perception, they are characterized with extreme articulatory-acoustic properties and act as natural referent vowels. These perception biases favoring peripheral vowels have been documented in infant speech perception but "may resurface in adults as they learn to map out a new vowel inventory in a second language" (p. 470), possibly when new L2 vowel categories fall in regions of the vowel space that are not firmly committed to the L1.

Finally, Bohn and Garibaldi (2021) compared the predictive power of three methods to assess the relationship between perceptual assimilation and discrimination of Danish vowels by Southern British English learners. The three methods under evaluation were i) categorization task with goodness ratings, ii) a graded discrimination task and, iii) acoustic comparisons of the L1 and L2 vowel spaces. The authors concluded that none of the three methods successfully predicted L2 discrimination.

1.3. The vowel systems of Spanish, Majorcan Catalan and Californian English

Spanish has a simple, symmetrical five-vowel system with two high vowels, /i/, /u/, two mid vowels /e/, /o/ and one low vowel /a/. Unlike English, all five vowels are pure non-diphthongized vowels and, according to Hualde (2005), vowel quality is quite stable across Spanish dialects. An exception is Andalusian Spanish, which has the open allophones [ɛ] and [ɔ] as a consequence of final /s/ elision. Although the Spanish vowel system has no phonological vowel reduction, Nadeu (2014) found that, in a normal speech rate, stressed vowels are longer relative to their unstressed counterparts. Another relevant finding for Spanish-Catalan speakers learning English involves the widely-accepted characterization of Spanish /a/ as a central vowel. Based on an exhaustive analysis on vowel asymmetry, Nadeu claims that it should be considered a front vowel both in Spanish and Eastern Catalan (EC).

The vowel system of Majorcan Catalan has eight vowels, the point vowels /a/, /i/, /u/, the two mid-vowel contrasts /e/-/ɛ/ and /o/-/ɔ/ and the central vowel /ə/. Cross-dialect studies of vowel change in Catalan (Recasens and Espinosa, 2006) reveal that, unlike other Catalan dialects, /ə/occurs in stressed and unstressed position in MC: "stressed schwa is realized as a mid-central vowel occupying a somewhat lower and more retracted position than unstressed /ə/" (p. 663). The realizations of /a/ and /ɔ/ in MC substantially differ from their counterparts in EC, specifically /a/ is more fronted relative to Eastern Catalan, MC /ɔ/ is lower and therefore more open than the canonical form in EC. In a follow-up study, Recasens and Espinosa (2009) found evidence of a vowel merger between /e/ and /ɛ/, caused by the front realization of /a/, which might have pushed /ɛ/ towards the /e/ space. This chain shift is documented for a minor dialect in MC, Felanitxer, but the authors did not report whether it extended to other dialectal areas in the island. In any case, the cross-dialect differences between EC and MC put MC learners of English in an advantageous situation compared to their EC peers. From the acoustic perspective, a fronted /a/ and a lower /ɔ/ make these L1 vowels more similar to the targets CE /æ/, /ɑ/.

The Californian English (CE) vowel space is much more crowded compared to that of Spanish or MC with 12 vowels: /i/, /ɪ/, /e/, /ɛ/, /æ/, /ə/, /ɝ/, /ʌ/, /ɑ/, /o/, /ʊ/ and /u/.[1] The characteristic traits of CE include, a merger of /ɑ/ and /ɔ/, lack of general tensing of /æ/, and a fronting and unrounding of /u/ (Hagiwara, 1997, 2005). Recent work investigating the sources of vowel change in the California Central Valley (D'Onofrio et al., 2019) have reported evidence of a compression of the vowel space in younger generations (aged 2–22) compared to older generations (aged 53–89), indicating that the distances between some vowel categories become smaller.

A cross-linguistic comparison of the vowel inventories of CE, MC and Sp must consider the role of Vowel-Inherent-Spectral-Change (VISC) in North American English (NAE), an acoustic parameter coined by Nearey and Assmann (1986) that challenged traditional accounts that labelled NAE vowels as monophthongal. VISC can be defined as a "systematic change in formant frequencies occurring during the vocalic nucleus" (Rogers et al., 2013, p. 232), thus acknowledging the dynamic nature of vowels in many NAE dialects. Nearey (2013) showed that VISC patterns vary across dialects and as a function of consonantal context, thus the data available allowed him to classify vowels in four subclasses according to their VISC trajectories. Thus, the high tense vowels /i/ and /u/ did not show significant movement as opposed to the lax vowels /ɪ/, /ɛ/, /æ/, and /ʊ/, that exhibit considerable VISC movement. From the perception perspective, VISC has been found to be a crucial cue for vowel identification by monolingual listeners but, as Rogers et al. (2013) posit, L2 learners might not pay attention to the type of information encoded by VISC, thus diminishing their ability to perceptually identify some of the target vowels accurately.

1.4. Empirical studies of vowel assimilation by L2-English L1-Spanish/L1-Catalan listeners

The perceptual assimilation of English vowels to Spanish (Sp) and/or Catalan (Cat) has been previously investigated for both L1-Spanish (Baigorri et al., 2019; Cebrian, 2019; Escudero and Chládková, 2010) and L1-Catalan listeners (Cebrian, 2021; Cebrian et al., 2010; Rallo Fabra, 2005). Some parallels between the results of these studies can be observed despite differences in their methodological approaches, including the number of L2 vowels tested, the number of the response categories, the type of stimuli used (synthetic vs. natural stimuli), the English variety of the vowel stimuli (American English vs. Southern British English), the L1 of the listeners (Spanish vs. Catalan) and their L2 experience (experienced vs. naïve). There seems to be agreement in the categorization of English vowels /i/ and /u/ as Sp/Cat /i/ and /u/, respectively. Both AE and BE /æ/ are consistently categorized as Sp/Cat /a/ with assimilation percentages that range from 87% to 100%. An exception is the mapping of AE /æ/ in Escudero and Chládková (2010), which was categorized as Sp /e/ in 99% of the cases. This discrepancy was probably due to the acoustic values used to create the synthetic vowel stimuli, which corresponded to the Northern Cities AE variety (Hillenbrand et al., 1998), which is characterized by the raising of /æ/. The categorization of AE /ɑ/ and SBE /ɑː/ onto Sp/Cat /a/ is also quite consistent across the studies with assimilation percentages that surpass 75%.

Cross-dialect differences in the English variety used to create the vowel stimuli are probably responsible for the different pattern observed for /ʌ/ in BE and AE. BE /ʌ/ is consistently assimilated to Sp /a/ by Spanish listeners (Cebrian, 2019; Escudero and Chládková, 2010) and also by Catalan experienced and naïve listeners (Cebrian, 2021; Cebrian et al., 2010). In contrast, AE /ʌ/ is uncategorized, with response categories split between Sp /a/ and /o/ (Baigorri et al., 2019; Escudero and Chládková, 2010). Similarly, the listeners' L1 vowel inventory, five vowels in Spanish /i e a o u/ and eight in Catalan /i e ɛ ə a ɔ o u/, probably account for the clear cross-language difference in the categorization of AE /o/ and SBE /ɔː/, which were categorized as Sp /o/ by Spanish listeners and Catalan naïve listeners but as Cat /ɔ/ by Catalan experienced listeners (Cebrian et al., 2010). Similarly, AE and SBE /ɛ/ were perceptually assimilated to Sp /e/ by Spanish listeners but as Cat /ɛ/ by Catalan listeners.

The case of /ɪ/ deserves special attention because there is considerable discrepancy across studies and it is not clear that this discrepancy can be attributed to cross-dialect differences in the vowel stimuli or cross-linguistic differences in the listeners' L1. This vowel has been categorized quite consistently as /e/ by Spanish and Catalan listeners with no or little L2 experience (Cebrian, 2021; Escudero and Chládková, 2010). Experienced Catalan listeners categorized it as Cat /i/ (Cebrian et al., 2010). Interestingly, both AE /ɪ/ and SBE /ɪ/ are uncategorized by Spanish experienced listeners (Baigorri et al., 2019) and Spanish naïve listeners (Cebrian, 2021; Escudero and Chládková, 2010). Given the disparity of the results, there seems to be a complex interaction of L2 experience and L1 background. L1-Spanish and Spanish-dominant bilinguals seem to perceive this vowel as "uncategorized" in PAM's terms. The response categories are balanced between /e/ and /i/. In contrast, for Catalan-dominant listeners /ɪ/ can be considered categorized, albeit the L1 response category varied depending on L2 experience. Naïve Catalan-dominant listeners categorize it as /e/ while experienced Catalan-dominants categorize it as Cat /i/.

Finally, the categorization of the AE central vowels /ɝ ə ɚ/, their SBE counterparts /ə ɜː/, and the lax vowel /ʊ/ cannot be compared because they were not the vowels of interestin the studies reviewed above. In sum, prior studies examining perceptual assimilation of AE/SBE vowels to Spanish/Catalan vowels indicate that both the L1 dialect and the English variety of the stimuli influence the assimilation patterns of L2 vowels to L1 vowels.

[1] IPA equivalences in Received Pronunciation (RP): /e/ = /eɪ/, /ɛ/ = /e/, /ɝ/ = /ɜː/. CE does not have the low rounded vowel /ɒ/ found in RP.

1.5. Empirical studies in L2 vowel discrimination by Spanish-Catalan learners

There is extensive evidence that Spanish/Catalan learners of English experience serious difficulties learning the vowel system of the target language. Rallo Fabra and Romero (2012) investigated perception and production of Catalan-English and English-English vowel contrasts by Catalan learners in a formal setting. They administered an oddity discrimination task to test the learners' sensitivity to seven Catalan-English (C-E) contrasts, /i/-/i/, /i/-/ɪ/, /ɛ/-/ɛ/, /ɛ/-/æ/, /a/-/æ/, /a/-/ʌ/, /u/-/u/ and four English (E) contrasts /i/-/ɪ/, /ɛ/-/æ/, /ɑ/-/ʌ/ and /u/-/ʊ/. Discrimination A′ scores were poor (lower than 0.5) for the C-E pairs /ɛ/-/ɛ/, /ɛ/-/æ/, /a/-/æ/, /a/-/ʌ/, moderate for C-E /i/-/i/, /u/-/u/ and good for C-E /i/-/ɪ/. The results of the English contrasts followed a similar pattern, moderate discrimination for the tense-lax vowel contrasts (/i/-/ɪ/, /u/-/ʊ/) but poor discrimination for the two contrasts involving the mid-low vowels /ɛ/-/æ/ and /ɑ/-/ʌ/. These findings were later replicated by Safronova (2016).

In the case of Spanish, there have been a few studies testing L1-Spanish speakers' perception of the English low vowels. For instance, Baigorri et al. (2019) examined perceptual assimilation and discrimination of American English (AE) vowels by two groups of Spanish-English bilinguals (early vs. late) and a group of English monolinguals. They found that the two bilingual groups faced serious difficulties discriminating the vowel contrasts /æ/-/ɑ/, /ʌ/-/ɑ/, /ʌ/-/æ/, although discrimination accuracy was higher by the early-bilinguals compared to the late-bilinguals. The relation between perceptual assimilation and discrimination was examined on the basis of PAM-L2. The authors concluded that future research should include the AE central vowels /ə/, /ɜ˞/, and /ɚ/.

Barrientos (2021) examined perception of the AE /ʌ/-/ɑ/ contrast in two groups of L1-Spanish speakers varying in English experience and a group of native English (NE) speakers. The results of an AX discrimination revealed that experience influenced listeners' sensitivity to this contrast. Experienced non-native speakers' d' scores did not significantly vary from native English listeners' values. Discrimination of the English-Spanish (E-S) contrasts /ɑ/-/a/ and /ʌ/-/a/ was poor even for the NE listeners (d' scores of 1 or below). This was taken as evidence that the AE /ʌ/-/ɑ/ contrast is an instance of a category-goodness assimilation in PAM terms. Additional discrimination data from acoustic continua revealed that experience had a positive effect in the discrimination of the two target vowels. In contrast, perceptual mapping of /ʌ/ and /ɑ/ to Spanish /a/ seemed not to be altered by experience with the L2.

2. The present study

This paper investigates the perception of Californian English (CE) vowels by Spanish-Catalan bilinguals in an instructional setting. Our main goal is to test the predictions of the PAM-L2, which posits that discrimination of L2 phonological contrasts can be predicted from the assimilation patterns of L2 vowels to L1 vowels. We claim that similarity relationships between CE and MC contrasts cannot be drawn on the basis of acoustic comparisons of the vowel inventories of both languages because (1) we would need to assume that perceptual and production abilities are aligned, something that has not been proven so far (Flege and Bohn, 2021) and, (2) the high level of cross-dialect vowel variability, both in CE and MC makes L2-to-L1 acoustic comparisons unreliable. In turn we intend to establish which CE vowel contrasts are more difficult to discriminate by MC learners in order to plan appropriate intervention schemes that tackle problematic areas.

Two experiments were designed: a categorization task in which participants were asked to map the CE vowels /i/, /ɪ/, /ɛ/, /æ/, /ʌ/, /ə/, /ɑ/, /ʊ/, and /u/ to the Majorcan Catalan (MC) vowel categories /i/, /e/, /ɛ/, /a/, /ə/, /ɔ/, /o/, /u/, and an AXB task testing the same participants' ability to discriminate the CE vowel contrasts /i/-/ɪ/, /ɛ/-/ɪ/, /ɛ/-/ə/, /æ/-/ʌ/, /æ/-/ɑ/, /ɑ/-/ʌ/, /ə/-/ʌ/, /ə/-/ʊ/, and /u/-/ʊ/. We intend to extend the existing literature in the field in various ways. To our knowledge, it is the first study that examines categorization of AE vowels to MC vowels adopting a whole-system approach in line with recent trends in L2 vowel perception (Bundgaard-Nielsen et al., 2011a, 2011b; Faris et al., 2018). The choice of CE for the vowel extends to the existing perceptual assimilation data for British English (Cebrian, 2019, 2021), Mid-Atlantic AE vowels (Baigorri and Levy, 2018), and Northern Cities AE vowels (Escudero and Chládková, 2010).

2.1. Method

2.1.1. Vowel stimuli

Five English L1 speakers from the California Bay area with a mean age of 21 years produced 10 repetitions of the vowel monophthongs /i/, /ɪ/, /ɛ/, /æ/, /ʌ/, /ə/, /ɜ˞/, /ɑ/, /u/, /ʊ/, the diphthongs /aʊ/, /aɪ/ and the vowel + /r/ sequences /ɔr/, /ar/, /ʊr, /ɛr/, /ɪr/ in the context of the nonword /hVbə/, which has been found to minimize the effects of coarticulation (Baigorri et al., 2019; Strange et al., 2007). The diphthongs and vowel + /r/ sequences were used as fillers and thus were not included in the analysis. All five speakers were undergraduate students from non-Hispanic households and were raised in the vicinities of San Francisco and Sacramento. The L1 speakers were recorded in a sound-proof cabin at the Phonetics Lab of the University of California Santa Cruz. The stimuli were randomly presented visually on a computer screen. In some cases, keywords were provided to aid the speakers with the production. The best four tokens from one speaker were selected for inclusion in the test on the basis of acoustic analysis and auditory judgment (Fig. 1). We discarded the other four speakers on these grounds: audible puffs of air, vocal fry, spirantization of /b/, devoicing of /b/ and lowering of /ə/.

2.1.2. Participants

The final sample consisted of 43 Spanish-Catalan bilinguals (mean age = 21.68, *SD* = 5.44, 38 females, 5 males). All participants reported growing-up in either Catalan-speaking (*n* = 13) or Spanish-speaking households (*n* = 32). Seven additional participants were tested but their data were removed because they were not Spanish-

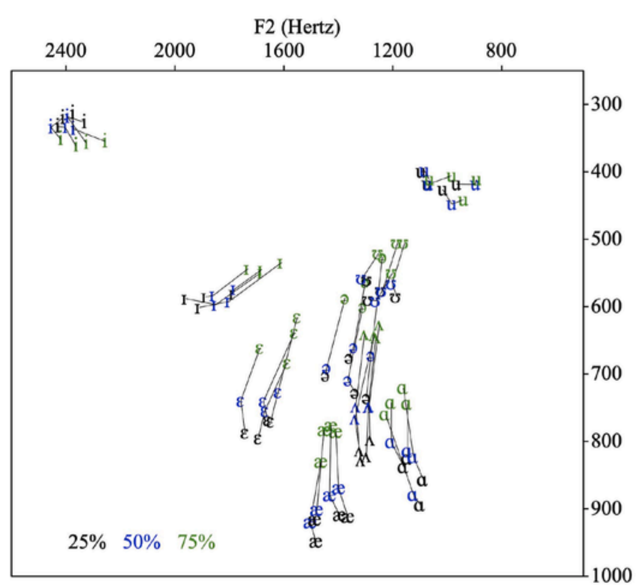

Fig. 1. F1 and F2 values of the h_ba nonwords used in the categorization and AXB experiments measured at 25%, (black), 50% (blue), and 75% (green) of the vowel segment duration. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Catalan bilinguals ($n = 4$), failed to follow the instructions ($n = 2$), or had a speech/language disorder ($n = 1$). The Catalan-dominant participants were exposed to MC from birth and reported speaking this language at home. The Spanish-dominant participants had been exposed to Catalan from the age of 2 or 3 and spoke both languages on a regular basis. The participants were experienced EFL learners with a B2+/C1 level of English, according to the *Common European Framework of Reference for Languages*, and received course credit for their participation.

2.1.3. Procedure

Participants were tested on campus in groups of two or three students and performed two perceptual tasks in the same session, an AXB discrimination task followed by a categorization task. They were told that they would hear three *h_ba* nonwords and they were instructed to choose whether the vowel in the first (A) or third (B) utterance was the same as the vowel in the middle utterance (X). The task was divided in nine separate blocks of 48 trials each, each block testing a different vowel contrast and presented in a counter-balanced order. There were 48 AXB trials for each contrast (12 of each AXB trial type: AAB, ABB, BBA and BBA). Each element of the AXB trial was a different token and the tokens were counterbalanced such as that each of the four tokens appeared an equal number of times in each position. In the categorization task, listeners were instructed to select the vowel label that most closely matched the vowel in each stimulus utterance, using a grid of Catalan vowels written in IPA script. There were eight monophthongs (a ɛ e i o ɔ u ə) and 24 permissible two-vowel combinations (e.g., ai au). Sample MC words were provided in orthographic form to aid the listeners with the task. After selecting the label, they rated the goodness of fit using a 7-point scale (1 = a poor fit, 7 = a perfect fit). If the vowel they heard did not fit on any of the response options, they could use three additional options: (1) *Not a speech sound* (2) *Other speech sound*. (3) *Unknown speech sound* and were requested to type a description of what they heard. The whole testing session lasted about 70 min, including a 10-min break between the two tasks but it varied as a function of the listener's ability to respond rapidly. If a listener surpassed the time limit of 2 s, a “time out” warning message appeared on screen prompting the listener to respond faster and the trial was presented a second time. The responses of the “time out” trials were not included in the analysis.

2.2. Results

2.2.1. Categorization

The mean percentages and goodness ratings of the CE vowel categorization by Spanish-Catalan listeners averaged across listeners are pooled in Table 1. Following previous work in L2 perception (Tyler et al., 2014), a given vowel was considered “categorized” if it was consistently mapped onto an L1 vowel category 70% of the time or above. Overall, listeners categorized CE vowels /i/, /u/, /æ/, and /ɑ/ quite consistently with assimilation percentages ranging from 74% to 94%. The remaining vowel categories, /ɪ/, /ɛ/, /ʌ/, /ə/, /ɝ/, and /ʊ/ were not consistently mapped onto a single MC category. However, the results for the uncategorized vowels must be examined in more detail because some of them are subject to substantial individual variability.

Analysis of the results at the individual level revealed high between-subject variability in the responses of some of the uncategorized vowels. For instance, /ɪ/ was quite consistently categorized as /e/ with percentages of 70% or higher by 18 participants. Ten participants categorized this vowel less consistently with percentages ranging from 50% to 70%. In these cases, the /ɛ/ response category was selected, suggesting that the L1 /e/-/ɛ/ contrast could be in the process of merging for some of the participants. These between-subject differences suggest that CE /ɪ/ is categorized as MC /e/ for around half of the cohort.

The case of CE /ɛ/ is interesting. If we take the 70% categorization benchmark, 9 participants mapped this vowel as MC /ɛ/ and twelve as MC /e/. If we lower the categorization benchmark to 50%, 14 listeners perceptually assimilated CE /ɛ/ to MC /ɛ/ and 19 listeners assimilated it to MC /e/. Closer inspection of individual data revealed that L1 dominance in either Spanish or Catalan partially accounted for this difference. Catalan-dominants and Catalan-Spanish balanced bilinguals used the /ɛ/ more often than Spanish-dominants, who had a preference for the /e/ label.

Table 1
Mean percent categorization of the 9 Californian English vowel monophthongs in terms Majorcan Catalan vowel categories. Responses categorized above 70% are in boldface and mean goodness ratings out of 7 are in parentheses for above-chance labels only. Responses selected below chance are in italics. Percentages have been rounded to the nearest whole number.

	Majorcan Catalan Response Category								
Stimulus Vowel	i	e	ɛ	a	ɔ	o	u	ə	other
i	**93** (6.3)	*3*						*1*	*1* ei
ɪ	*8*	59 (5.3)	16 (4.7)				*1*	13 (4.3)	*1* ei
ɛ		41 (5.0)	34 (5.1)	*3*				21 (4.8)	
æ			*3*	**94** (6.0)	*1*			*1*	*1* au
ʌ		*1*	*3*	66 (5.6)	*2*			23 (5.0)	*3* au
ɑ			*3*	**74** (5.6)	*9*	*2*		8 (5.0)	*3* au
ʊ		*1*	*1*		14 (4.3)	38 (4.6)	21 (4.9)	*8*	*8* ou *2* ue *1* au *1* eu *1*uɛ *1* uo
u							**94** (6.3)		*3* ou *1* iu
ə		*6*	8 (4.2)	23 (4.8)	*2*	*1*	*1*	56 (5.1)	*1* ɛu

The central vowels /ʌ/ and /ə/ also deserve special attention. The average categorization percentages do not reach the 70% benchmark. However, this benchmark was achieved by 23 of the participants, who mapped CE /ʌ/ as MC /a/ quite consistently. Another group of 9 also categorized CE /ʌ/ as MC /a/, albeit less consistently, with categorization percentages that ranged from 50% to 70%. As for CE /ə/, we find a similar trend, 22 participants perceived it as close to MC /ə/ 70% of the times or higher, another 4 listeners also perceived CE /ə/ as its MC counterpart /ə/ but less consistently with percentages of 62.5% or 66.7% and lower goodness ratings (3.7–5.6).

Based on the categorization percentages tallied in Table 1, we can make predictions of discrimination accuracy following PAM-L2 (Best and Tyler, 2007; Tyler, 2019). The PAM-L2 assimilation patterns and the predictions of ease/difficulty of discrimination are shown in Table 2. Due to insufficient prior empirical data, it is difficult to make very specific predictions. We speculate that the ranking of ease of discrimination could be the following: CG > UC-N > UC-P > UC-C > UU-N > UU-P > UU-C, where CG stands for “category-goodness” assimilation, UC stands for uncategorized-categorized assimilation and UU for uncategorized-uncategorized assimilation. These three patterns of assimilation are followed by N, P or C, indicating no overlap, partial overlap or complete overlap, respectively (Faris et al., 2018).

In line with previous studies in L2 perception (Baigorri et al., 2019; Levy, 2009; Vasiliev, 2013), we also calculated modal overlap scores based on the percentage of perceptual overlap of the overall assimila-

Table 2
PAM assimilation patterns and % overlap scores for the nine vowel contrasts tested.

Contrast	PAM Assimilation pattern	% Overlap score
/i-ɪ/	UC-N	13.3
/ʊ-ə/	UU-N	15.8
/ɛ-ə/	UU-P	38.9
/ʊ-u/	UC-P	26.6
/ɪ-ɛ/	UU-C	70.9
/æ-ɑ/	CG	80.1
/æ-ʌ/	UC-P	72.4
/ʌ-ɑ/	UC-C	83.0
/ʌ-ə/	UU-C	53.5

tion percentages shown in Table 1. This metric allowed us relate the overall overlap scores with the PAM assimilation patterns (Table 2). Additionally, due to the high inter-subject variability we also calculated an individual overlap score for each participant and for each of the nine vowel contrasts (Table 3). For instance, on the /æ-ʌ/ contrast, a given listener categorized CE /æ/ as MC /a/ and /ə/ categories on 70.8% and 29.2% of the trials, respectively. The same listener categorized CE /ʌ/ as MC /a/, /ə/, and /ɔ/ on 37.5%, 37.5%, and 20.8% of the trials. The overlap score for this listener's vowel contrast is calculated by summing the lower percentages when the two vowels of the contrast were categorized to the same MC categories, that is /a/ = 37.5%, /ə/ 29.2% and /ɔ/ = 0%, which results in an overlap score of 66.7%. The descriptive statistics for the individual overlap scores are shown in Table 3.

Table 3
Mean overlap scores from individual participants.

Contrast	Mean	*SD*	Minimum	Maximum
/i-ɪ/	9.9	18.7	0.0	79.2
/ʊ-u/	24.5	25.2	0.0	83.3
/ɛ-ə/	29.8	21.5	0.0	79.2
/ʊ-ə/	13.7	13.2	0.0	58.3
/ɪ-ɛ/	59.7	27.3	4.2	100.0
/æ-ɑ/	77.6	18.9	8.3	100.0
/æ-ʌ/	69.5	24.2	16.7	100.0
/ʌ-ɑ/	50.1	15.8	16.7	83.3
/ʌ-ə/	29.8	21.5	0.0	79.2

2.2.2. AXB discrimination

Discrimination accuracy was calculated from the total number of correct trials out of a total of 48 for each of the CE vowel contrasts. A value of 100% indicates excellent discrimination, values close to 50% indicate poor discrimination. As shown in Fig. 2, listeners discriminated seven out of the nine vowel contrasts tested quite accurately with scores averaging 80% or higher. Discrimination of the tense/lax vowel contrasts /i/-/ɪ/ and /u/-/ʊ/ approached ceiling with mean scores of 97.8% (SD = 2.6%) and 94.3% (SD = 6.5%), respectively. Five contrasts, namely /ɛ/-/ə/, /ʊ/-/ə/, /ɪ/-/ɛ/, /æ/-/ɑ/, and /æ/-/ʌ/ were moderately well discriminated, averaging 92.6 (SD = 7.4%), 90% (6.3%), 85.6% (SD = 11.1%), 84.2% (SD = 0.139) and 81.8 (SD = 12.2%). In contrast, /ʌ/-/ɑ/ and /ʌ/-/ə/ were discriminated less accurately, with mean discrimination scores of 67.2% (SD = 11.7%) and 65% (SD = 13.8%), respectively.

A series of Kolmogorov-Smirnov normality tests revealed that discrimination of four of the contrasts tested exhibited non-normal distributions (/i/-/ɪ/: $D = 0.256$, $p < .05$; /ɪ/-/ɛ/: $D = 0.139$, $p < .05$; /æ/-/ʌ/: $D = 0.115$, $p = .160$; /æ/-/ɑ/: $D = 0.161$, $p < .05$; /ʌ/-/ɑ/: $D = 0.101$, $p = .317$; /u/-/ʊ/: $D = 0.218$, $p < .05$; /ɛ/-/ə/: $D = 0.201$, $p < .001$; /ʌ/-/ə/: $D = 0.104$, $p = .284$; /ʊ/-/ə/: $D = 0.300$, $p < .05$). Therefore, a non-parametric *Friedman's* ANOVA was run to test differences in discrimination accuracy between the nine vowel contrasts. A significant effect of contrast was found: $\chi^2(8) = 215.61$, $p < .001$. Stepwise step-down comparisons yielded the following discrimination ranking from easiest to discriminate to more difficult to discriminate: /i/-/ɪ/ > (/u/-/ʊ/, /ɛ/-/ə/) > (/ʊ/-/ə/, /ɪ/-/ɛ/, /æ/-/ɑ/) > /æ/-/ʌ/> (/ʌ/-/ɑ/, /ʌ/-/ə/).

Spearman rank-order correlations were run at the group and individual levels to test whether differences in discrimination accuracy could be predicted from the overlap scores in the categorization of each vowel. At the group level, overall discrimination accuracy for each vowel contrast was negatively correlated with percent overlap ($r_s(9) = -0.767$, $p < .05$). We also found significant negative correlations between each individual participant's discrimination accuracy and overlap scores for /i/-/ɪ/ ($r_s(43) = -0.302$, $p < .05$), /ɪ/-/ɛ/ ($r_s(43) = -0.391$, $p < .005$), /u/-/ʊ/ ($r_s(43) = -0.389$, $p < .005$) and /ɛ/-/ə/ ($r_s(43) = -0.286$, $p < .05$), indicating that discrimination increased as percent of overlap decreased. However, the correlations for the other five contrasts failed to reach significance: /æ/-/ʌ/ ($r_s(43) = -0.152$, $p = .165$), /æ/-/ɑ/ ($r_s(43) = -0.007$, $p = .482$), /ʌ/-/ɑ/ ($r_s(43) = 0.003$, $p = .493$), /ʌ/-/ə/ ($r_s(45) = -0.142$, $p = .181$) and /ʊ/-/ə/ ($r_s(43) = -0.049$, $p > .377$).

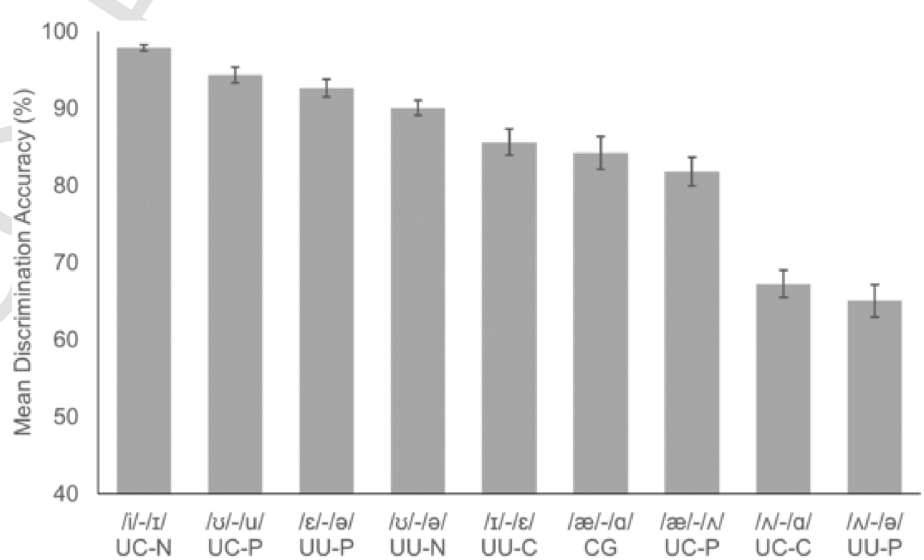

Fig. 2. *Mean discrimination accuracy of the nine CE vowel contrasts tested. Scores range from poor* (60%) *to very good discrimination* (100%).

2.3. Discussion

One of the aims of this study was to establish the categorization patterns of CE vowels to MC vowel categories by a group of Spanish-Catalan listeners with English experience following a "whole system approach". Listeners categorized CE vowels /i/, /u/, /æ/ and /ɑ/ quite consistently with percentages that ranged from 70% to 100%. The CE vowels /ɪ/, /ɛ/, /ʌ/, /ə/, and /ɝ/ were uncategorized in PAM-L2's terms but differed in degree of perceived phonological overlap. The highly consistent categorization of CE /i/, /u/, and /æ/ could also be explained in terms of phonological theory. As argued by Nadeu (2014), the corner vowels /i/, /a/, and /u/ have a privileged status due to their position in the vowel space, are more acoustically stable, and are not subject to vowel reduction processes. The inconsistent categorization of CE /ɪ/ as MC /e/ could be explained by the instability of the /e/-/ɛ/ contrast in various Catalan dialects due to the influence of Spanish (Mora and Nadeu, 2012). A closer inspection of the data at the individual level revealed that 90% of the listeners who categorized CE /ɪ/ as MC /ɛ/, were Spanish-dominant so we speculate that MC /e/ and /ɛ/ might be perceptually overlapping categories for these listeners. Further to this, the Catalan mid vowels /ɛ/, /e/, and /ə/ are subject to high cross-dialect variability (Recasens and Espinosa, 2006), specifically the merger of /e/ and /ɛ/ in a minor subdialect of MC (Calvo Barreiro, 2021; Torres-Tamarit and Hamann, 2021) would somehow account for the cross-dialect difference in the categorization of English /ɛ/ in Central Catalan (Cebrian, 2021; Cebrian et al., 2010) and Majorcan Catalan. The possibility exists that, in line with the SLM-r, listeners vary in terms of category precision, "an endogenous factor that is potentially linked to individual differences in auditory acuity, early stage (precategorical) auditory processing, and auditory working memory" (p.36) (Flege and Bohn, 2021). This would explain the listeners' difficulty to differentiate their own L1 categories and thus could not categorize some of the CE vowels accordingly.

However, the categorization patterns just reported cannot be explained on the basis of language experience alone. The vowel inventory of MC does not include /æ/ or /ɑ/, yet these were quite consistently categorized. In turn, the central vowel /ə/ is found in MC in stressed and unstressed position but listeners did not categorize this vowel consistently. The poor categorization of /ə/ and /ʌ/ can be atributted to the merger of these two vowels in CE and a lack of L1 category precision because MC /ə/ is often defined as "targetless schwa" due to its expansion in the acoustic space. The possibility of a perceptual bias towards more peripheral vowels, as suggested by Polka and Bohn's (2011) Natural Referent Vowel (NRV) framework, should also be considered. These biases have been found in studies of infant vowel perception and "may resurface in adults as they learn to map out a new vowel inventory in a second language" (p. 470). Finally, VISC did not seem to have much weight in the categorization of CE vowels by MC listeners. All the CE vowel stimuli, except for /i/ and /u/, exhibited a systematic change of frequency, yet listeners were able to categorize some of these vowels quite consistently, especially /æ/ and /ɑ/.

2.3.1. Predictions of discriminability

The second aim of the study was to test whether discrimination accuracy of the CE vowel contrasts could be predicted from the categorization patterns of the L2 vowel categories in terms of the L1 categories. This prediction was tested by calculating the percent overlap score at the group and individual levels following previous studies in L2 perception (Baigorri et al., 2019; Flege and Mackay, 2004; Levy, 2009). At the group level, the significant negative correlation found between discrimination scores and percentage of overlap support the predictive power of the overlap score. However, when we examine the individual results for each one of the contrasts separately, discrimination can only be predicted from the overlap in four of the contrasts tested, namely, /i/-/ɪ/, /ɪ/-/ɛ/, /u/-/ʊ/ and /ɛ/-/ə/. In these cases, the significant correlations found account for 8%–15% of the variance but the lack of significant correlations for /æ/-/ʌ/, /æ/-/ɑ/, /ʌ/-/ɑ/, /ʌ/-/ə/, and /ʊ/-/ə/ question the predictive power of this metric.

In terms of the PAM-L2, the contrast /æ/-/ɑ/ was the only instance of a category-goodness (CG) pattern of discrimination. As predicted, discrimination accuracy was moderate-good averaging 84.2%. Three contrasts, namely /i/-/ɪ/, /æ/-/ʌ/, and /u/-/ʊ/, were classified as uncategorized-categorized (UC), with different degrees of perceptual overlap, no overlap for /i/-/ɪ/ and partial overlap for /æ-ʌ/ (72.5%) and /u/-/ʊ/ (27.3%). Again, the discrimination scores met PAM-L2's predictions of high accuracy for UC contrasts with no/partial overlap. The low discrimination scores obtained for /ʌ/-/ɑ/ and /ʌ/-/ə/ are consistent with the complete overlap between the two categories of each contrast, which averaged 82.6% and 53.3%, respectively. Finally, discrimination accuracy for the completely overlapping /ɪ/-/ɛ/ contrast was higher than would be predicted by PAM-L2.

The relationship between perceptual assimilation patterns and discrimination accuracy also meets one of the tenets of the L2LP model (van Leussen and Escudero, 2015), since the two vowels of best-discriminated contrasts /i/-/ɪ/, /u/-/ʊ/, and /ɛ/-/ə/ are quite distant from one another in the acoustic vowel space. In contrast, the two vowels in the poor-discriminated contrasts /ʌ/-/ɑ/ and /ʌ/-/ə/ overlap one another in the vowel space. Contrary to what has been accepted so far, cross-language similarity can be also assessed acoustically despite the evidence that acoustic measures might not reflect what listeners perceive (Flege et al., 2021) (p. 33).

Just like in the categorization task, some of the between-contrast differences in vowel discrimination could be explained in terms of the NRV. Participants showed little sensitivity to the two vowels in the /ʌ/-/ɑ/ and and /ʌ/-/ə/ contrasts. Interestingly, listeners showed higher sensitivity to /æ/-/ʌ/ compared to /ʌ/-/ɑ/, it seems that a perceptual bias towards peripheral front vowels could partially account for these results. The poor discrimination of the /ʌ/-/ɑ/ replicates prior findings (Baigorri et al., 2019; Barrientos, 2021) and they can be extended by a recent study of vowel discrimination by Danish EFL learners (Bohn and Garibaldi, 2021). The case of /ʌ/-/ə/ seems to support one of the tenets of the SLM-r when it claims that "the mapping of L2 sounds to L1 sounds occurs at the level of position-sensitive allophones, not phonemes" (p. 13). The schwa vowel occurs in stressed and unstressed position in MC, so participants probably merged CE /ʌ/ and /ə/ as a sole category and could not discriminate between them, despite the fact that /ə/ only occurs in stressed position in English.

2.3.2. Establishment of new phonological categories

L2 categories develop with linguistic exposure and interaction with speakers of the language (Strange, 2011). This would partially explain why the participants performed much better in the AXB task compared to the participants in prior studies testing discrimination of English vowels by listeners in an instructed-learning context (Carlet and Cebrian, 2019; Fouz-González, 2020; Rallo Fabra and Romero, 2012). Various factors might have accounted for these differences. First, we should consider the role of methodological factors. Rallo Fabra and Romero (2012) used an oddity task and vowel stimuli that had been produced by AE speakers from different dialectal regions in the US, which probably added extra difficulty to the task. Second, the linguistic profiles of the participants in the present study are not the same as in our previous one. The former received a substantial amount of authentic input from the media and from native speakers visiting the island during the holiday season, which might provide the closest conditions to a naturalistic setting. As noted by Flege (2008), exposure to authentic input enhances L2 perceptual learning and eventually leads to the establishment of long-term mental representations for the L2 vowel categories.

2.3.3. Methodological limitations and suggestions for further research

Task demands have been found to highly influence perception performance (Werker and Curtin, 2005). The possibility exists that the "whole system" approach adopted in the categorization task has the trade-off that the number of response options increases the level of difficulty and, consequently, triggers more dispersion of results (Flege, 2021). Further to this, work by Strange and colleagues (Strange et al., 1998, 2001) has shown that vowel categorization is sensitive to various factors, among them, consonant and prosodic context. For instance, Japanese listeners with very limited exposure to English were more consistent categorizing American English [ɛ æː ɑː ʌ] when these were presented in disyllabic pseudowords embedded in sentences than when the vowels were presented in citation-form disyllables. Strange (2011) concluded that perceptual assimilation patterns are better established with speech materials that better reflect the conditions of perception of continuous speech. We suggest that future research should consider the possibility of testing categorization of L2 vowels to L1 vowels in a two-step process. The first step should follow the whole system approach adopted in the present study and prior research (Bundgaard-Nielsen et al., 2011a, 2011b; Faris et al., 2018). The second step should include the L2 sounds that were not consistently categorized onto an L1 vowel category (i. e. below 70%) and limit the response categories to the above-chance responses. Future research avenues should also investigate the weight of VISC in categorization consistency by L2 listeners of different L2 backgrounds.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by grants FFI 2017-84479-P from the Spanish Ministry of Economy, Industry and Competitiveness and PID 2020-117804 GB-100 from the Spanish Ministry of Science and Innovation. We are indebted to Mark Amengual, Clara Vega Munar and Miquel Àngel Palanco for their assistance with data collection and analysis.

References

[Amengual, 2016] M. Amengual, Cross-linguistic influence in the bilingual mental lexicon : evidence of cross-linguistic influence in the bilingual mental lexicon : evidence of cognate effects in the phonetic production and processing of a vowel contrast, Front. Psychol. 7 (2016) 617.

[Baigorri et al., 2019] M. Baigorri, L. Campanelli, E.S. Levy, Perception of American–English vowels by early and late Spanish–English bilinguals, Lang. Speech 62 (2019) 681–700.

[Baigorri and Levy, 2018] M. Baigorri, E.S. Levy, Perception of American – English vowels by early and late Spanish – English bilinguals, Lang. Speech 62 (2018) 681–700.

[Barrientos, 2021] F. Barrientos, On segmental representations in second language phonology: A perceptual account, Sec. Lang. Res. 39 (2021) 259–285.

[Barrios et al., 2016] S. Barrios, N. Jiang, W.J. Idsardi, Similarity in L2 Phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts, Second Lang. Res. 32 (2016) 367–395.

[Barsalou, 1987] L.W. Barsalou, The instability of graded structure: Implications for the nature of concepts, in: U. Neisser (Ed.), Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization, Cambridge University Press, 1987, pp. 101–140.

[Best et al., 2007] C.T. Best, M.D. Tyler, Nonnative and second-language speech perception: commonalities and complementarities, in: O.-S. Bohn, M.J. Munro (Eds.), Lang. Exp. Second Lang. Speech Learn. Honor James Emil Flege, John Benjamins, Amsterdam, 2007, pp. 13–34.

[Bohn and Garibaldi, 2021] O.-S. Bohn, C.L. Garibaldi, Examining the success of three kinds of cross-language similarity in predicting English listeners' discrimination of Danish vowels, J. Acoust. Soc. Am. 150 (2021) A44–A44.

[Bradlow et al., 1997] A.R. Bradlow, D.B. Pisoni, R. Akahane-Yamada, Y. Tohkura, Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production, J. Acoust. Soc. Am. 101 (1997) 2299–2310.

[Bundgaard-Nielsen et al., 2011a] R.L. Bundgaard-Nielsen, C.T. Best, M.D. Tyler, Vocabulary size matters: the assimilation of second-language Australian English vowels to first-language Japanese vowel categories, Appl. Psycholinguist. 32 (2011a) 51–67.

[Bundgaard-Nielsen et al., 2011b] R.L. Bundgaard-Nielsen, C.T. Best, M.D. Tyler, Vocabulary size is associated with second-language vowel perception performance in adult learners, Stud. Second Lang, Acquisitions 33 (2011b) 433–461.

[Burnham et al., 2002] D. Burnham, M. Tyler, S. Horlyck, Periods of speech perception development and their vestiges in adulthood, Trier, in: P. Burmeister, T. Piske, A. Rohde (Eds.), An Integr. View Lang. Dev. Pap. Honor Henning Wode, 2002, pp. 281–300.

[Calvo Barreiro, 2021] A. Calvo Barreiro, Dominància lingüística i realització de les vocals mitjanes del valencià central, Universitat de les Illes Balears, 2021.

[Carlet et al., 2019] A. Carlet, J. Cebrian, Assessing the effect of perceptual training on L2 vowel identification, generalization and long-term effects, in: A.M. Nyvad, M. Hejná, A. Højen, A.B. Jespersen, M.H. Sørensen (Eds.), A Sound Approach to Lang. Matters Honor Ocke-Schwen Bohn, Dept. of English, School of Communication & Culture, Aarhus University, 2019, pp. 91–119.

[Cebrian, 2019] J. Cebrian, Perceptual assimilation of British English vowels to Spanish monophthongs and diphthongs, J. Acoust. Soc. Am. 145 (2019) EL52–EL58.

[Cebrian, 2021] J. Cebrian, Perception of English and Catalan vowels by English and Catalan listeners: A study of reciprocal cross-linguistic similarity, J. Acoust. Soc. Am. 149 (2021) 2671–2685.

[Cebrian et al., 2010] J. Cebrian, J.C. Mora, C. Aliaga-García, Assessing crosslinguistic similarity by means of rated discrimination and perceptual assimilation tasks, in: M. Wrembel, M. Kul, K. Dziubalska-Kolaczyk (Eds.), Proceedings of the 6th international symposium on the acquisition of second language speech, New Sounds, Peter Lang, Frankfurt, Germany, 2010, pp. 77–82.

[D'Onofrio et al., 2019] A. D'Onofrio, T. Pratt, J. Van Hofwegen, Compression in the California Vowel Shift: Tracking generational sound change in California's Central Valley, Language Variation and Change 31 (2) (2019) 193–217, https://doi.org/10.1017/S0954394519000085.

[Eimas et al., 1987] P.D. Eimas, J.L. Miller, P.W. Jusczyk, On infant speech perception and the acquisition of language. - PsycNET, in: S. Harnad (Ed.), Categ. Percept. Groundwork Cogn., Cambridge University Press, Cambridge, 1987, pp. 161–195.

[Escudero and Chládková, 2010] P. Escudero, K. Chládková, Spanish listeners' perception of American and Southern British English vowels, J. Acoust. Soc. Am. 128 (2010) EL254–L259.

[Faris et al., 2016] M.M. Faris, C.T. Best, M.D. Tyler, 139, 2016, pp. 1–EL5.

[Faris et al., 2018] M.M. Faris, C.T. Best, M.D. Tyler, Discrimination of uncategorised non-native vowel contrasts is modulated by perceived overlap with native phonological categories, J. Phonetics 70 (2018) 1–19.

[Flege and Mackay, 2004a] I.R.A. Flege, J.E. Mackay, Perceiving vowels in a, stud. Second lang, Acquisitions 26 (2004a) 1–34.

[Flege1995] Flege, J.E., 1995. Second language speech learning: Theory, findings, and problems. In: Strange, W. (Ed.), Sp. Perc. Ling. Exp. Issues Cross-Lang. Res. York Press, Baltimore, pp. 233–276.

[Flege et al., 2008] J.E. Flege, Give input a chance!, in: T. Piske, M. Young-Scholten (Eds.), Input matters in SLA, Multilingual Matters, Bristol, 2008, pp. 175–190.

[Flege and Wayland, 2021] J.E. Flege, New methods for second language (L2) speech research, in: R. Wayland (Ed.), Second Lang. Speech Learn. Theor. Empir. Prog., Cambridge University Press, Cambridge, 2021, pp. 119–156.

[Flege et al., 2021] J.E. Flege, O.-S. Bohn, The revised speech learning model (SLM-r), in: R. Wayland (Ed.), Second Lang. Speech Learn. Theor. Empir. Prog., Cambridge University Press, Cambridge, 2021, pp. 3–83.

[Flege and MacKay, 2004b] J.E. Flege, I.R.A. MacKay, Perceiving vowels in a second language, Stud. Second Lang, Acquisitions 26 (2004b) 1–34.

[Fouz-González, 2020] J. Fouz-González, Using apps for pronunciation training: An empirical evaluation of the English File Pronunciation app, Lang. Learn, Technol. 24 (2020) 62–85.

[Hagiwara, 1997] R. Hagiwara, Dialect variation and formant frequency: The American English vowels revisited, J. Acoust. Soc. Am. 102 (1997) 655–658.

[Hagiwara, 2005] R. Hagiwara, Monophthongs and formant movement in North American English, J. Acoust. Soc. Am. 118 (2005) 2037.

[Hillenbrand et al., 1998] J. Hillenbrand, L.A. Getty, M.J. Clark, K. Wheeler, Acoustic characteristics of American English vowels, J. Acoust. Soc. Am. 97 (1998) 3011–3099.

[Hualde, 2005] J.I. Hualde, The sounds of Spanish, Cambridge University Press, Cambridge, 2005.

[van Leussen and Escudero, 2015] J.W. van Leussen, P. Escudero, Learning to perceive and recognize a second language: the L2LP model revised, Front. Psychol. 6 (2015) 1000.

[Jenkins et al., 1995] J.J. Jenkins, W. Strange, L. Polka, Not everyone can tell a "rock" from a "lock": Assessing individual differences in speech perception, in: D.J. Lubinski, R.V. Dawis (Eds.), Assessing individual differences in human behavior: New concepts, methods, and findings, Davies-

Black Publishing, Palo-Alto, CA, USA, 1995, pp. 297–325.
[Levy, 2009] E.S. Levy, On the assimilation-discrimination relationship in American English adults' French vowel learning, J. Acoust. Soc. Am. 126 (2009) 2670–2682.
[Lively et al., 1994] S.E. Lively, D.B. Pisoni, R.A. Yamada, Y. Tohkura, T. Yamada, Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories, J. Acoust. Soc. Am. 96 (1994) 2076–2087.
[Medin et al., 1987] D.L. Medin, L.W. Barsalou, Categorization processes and categorical perception. - PsycNET, in: S. Harnad (Ed.), Categ. Percept. Groundwork Cogn., Cambridge University Press, Cambridge, 1987, pp. 455–490.
[Mora and Nadeu, 2012] J.C. Mora, M. Nadeu, L2 effects on the perception and production of a native vowel contrast in early bilinguals, Int. J. Biling. 16 (2012) 484–500.
[Nadeu, 2014] M. Nadeu, Stress- and speech rate-induced vowel quality variation in Catalan and Spanish, J. Phonetics 46 (2014) 1–22.
[Nearey et al., 2013] T.M. Nearey, Vowel inherent spectral change in the vowels of North American English, in: G.S. Morrison, P.F. Assmann (Eds.), Vowel Inherent Spectr. Change, Springer Berlin, Heidelberg, 2013, pp. 49–85.
[Nearey and Assmann, 1986] T.M. Nearey, P.F. Assmann, Modeling the role of inherent spectral change in vowel identification, J. Acoust. Soc. Am. 80 (1986) 1297–1308.
[Ortega et al., 2021] M. Ortega, I. Mora-Plaza, J.C. Mora, Differential effects of lexical and non-lexical high-variability phonetic training on the production of L2 vowels, in: A. Kirkova-Naskova, J. Fouz-González, A. Henderson (Eds.), English Pronunciation Instr. Res. Insights, John Benjamins, Amsterdam, 2021, pp. 327–356.
[Polka and Bohn, 2011] L. Polka, O.S. Bohn, Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development, J. Phonetics 39 (2011) 467–478.
[Rallo Fabra, 2005] L. Rallo Fabra, Predicting ease of acquisition of L2 speech sounds. A perceived dissimilarity test, VIAL, Vigo Int. J. Appl. Linguist. 2 (2005) 75–92.
[Rallo Fabra and Romero, 2012] L. Rallo Fabra, J. Romero, Native Catalan learners' perception and production of English vowels, J. Phonetics 40 (2012) 491–508.
[Recasens and Espinosa, 2006] D. Recasens, A. Espinosa, Dispersion and variability of Catalan vowels, Speech Commun. 48 (2006) 645–666.
[Recasens and Espinosa, 2009] D. Recasens, A. Espinosa, Dispersion and variability in Catalan five and six peripheral vowel systems, Speech Commun. 51 (2009) 240–258.
[Rogers et al., 2013] C.L. Rogers, M.M. Glasbrenner, T.M. DeMasi, M. Bianchi, Vowel inherent spectral change and the second-language learner, in: G.S. Morrison, P.F. Assmann (Eds.), Vowel Inherent Spectr. Chang., Springer Berlin, Heidelberg, 2013, pp. 231–259.
[Safronova2016] Safronova, E., 2016. The role of cognitive ability in the acquisition of second language perceptual phonological competence. Unpublished PhD thesis, Universitat de Barcelona, Barcelona, Spain.
[Strange, 2011] W. Strange, Automatic selective perception (ASP) of first and second language speech: A working model, J. Phonetics 39 (2011) 456–466.
[Strange et al., 2001] W. Strange, R. Akahane-Yamada, R. Kubo, S.A. Trent, K. Nishi, Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners, J. Acoust. Soc. Am. 109 (2001) 1691–1704.
[Strange et al., 1998] W. Strange, R. Akahane-Yamada, R. Kubo, S.A. Trent, K. Nishi, J.J. Jenkins, Perceptual assimilation of American English vowels by Japanese listeners, J. Phonetics 26 (1998) 311–344.
[Strange et al., 2007] W. Strange, A. Weber, E.S. Levy, V. Shafiro, M. Hisagi, K. Nishi, Acoustic variability within and across German, French, and American English vowels: phonetic context effects, J. Acoust. Soc. Am. 122 (2007) 1111–1129.
[Torres-Tamarit and Hamann, 2021] F. Torres-Tamarit, S. Hamann, Canvi lingüístic a Eivissa: Noves dades sobre la vocal neutra - YouTube, Semin. Virtuals Del GRESIB 2020-2021 (2021). https://youtu.be/F0rThFOlxhw.
[Tyler2019] Tyler, M.D., 2019. PAM-L2 and phonological category acquisition in the foreign language classroom. In: Nyvad, A.M.Hejná, M.Højen, A.Jespersen, A.B.Sørensen, M.H.A (Eds.) Sound Approach to Lang. Matters Honor Ocke-Schwen Bohn. Dept. of English, School of Communication & Culture, Aarhus University, pp 607-630.
[Tyler et al., 2014] M.D. Tyler, C.T. Best, A. Faber, A.G. Levitt, Perceptual assimilation and discrimination of non-native vowel contrasts, Phonetica 71 (2014) 4–21.
[Vasiliev2013] Vasiliev, P., 2013. The initial state for Californian English learners of Spanish and Portuguese vowels. Unpublished PhD thesis, University of California, Los Angeles, USA.
[Werker and Curtin, 2005] J. Werker, S. Curtin, PRIMIR: A developmental framework of infant speech processing, Lang. Learn. Dev. 1 (2005) 197–234.
[Williams, 1977] L. Williams, The perception of stop consonant voicing by Spanish-English bilinguals, Percept. Psychophys 21 (1977) 289–297.