PRODUCTION AND PERCEPTION OF ENGLISH LOW-MID VOWELS BY SPEAKERS OF COLOMBIAN SPANISH IN AN ENGLISH LANGUAGE IMMERSION ENVIRONMENT

Gina Marcela Pineda Mora[1], Jenna Conklin[2], Kelly Johanna Vera Diettes[1], Olga Dmitrieva[2]

Universidad Nacional de Colombia[1], Purdue University[2]
gmpinedam@unal.edu.co, jconkli@purdue.edu, kjverad@unal.edu.co, odmitrie@purdue.edu

ABSTRACT

The study examines the perception and production of the English vowels [æ] and [ʌ] by native speakers of Colombian Spanish who are experienced L2 English speakers. Participants recorded three minimal pairs and performed a multiple-choice identification task. Acoustic data were compared to participants' Spanish vowels and monolingual English speakers' productions. Perceptual results demonstrated that although [æ] and [ʌ] had fewer correct identifications than other English vowels, the difference was not statistically significant. When incorrectly identified, [æ] and [ʌ] were typically confused with each other or another low vowel, [ɑ]. In production, although participants realized [æ] and [ʌ] as acoustically distinct, both were statistically different from the native targets. Participants' English [æ] was not statistically different from their native Spanish [a] in either height or backness. The results suggest that English low-mid vowels present a moderate challenge to experienced Spanish learners immersed in an English language environment.

Keywords: L2 English, L1 Spanish, low-mid vowels, acoustics, perceptual identification.

1. INTRODUCTION

The acquisition of second language (L2) phonology is one of the most difficult tasks L2 learners face [14]. It is generally believed that the relative difficulty of acquiring L2 phonological categories and contrasts is determined by the position of L2 speech sounds in the acoustic-perceptual space relative to the L1 sounds [1, 7, 8]. Specifically, L2 sounds that are acoustically distinct but perceptually similar to a given L1 category are expected to be perceptually assimilated to the L1 category and subsequently be rendered in a non-native fashion in production [8, 9, 10, 13]. Moreover, if two L2 categories are assimilated to a single L1 sound, discrimination between the L2 categories is also expected to be impeded [2]. These theoretical predictions have found empirical support in multiple studies of L2 acquisition. For example, several investigations have shown that English tense-lax contrasts in vowels, such as [i]-[ɪ], present a challenge for Spanish-speaking learners, who tend to assimilate both vowels to Spanish [i] [8].

The present investigation explores the acquisition of American English (AE) vowels [æ] and [ʌ], which have received less attention in the literature. The low back vowel [ɑ] was not targeted in our investigation due to the widespread dialectal variation in its realization and distribution [19]. The low-mid vowel space is more crowded in AE than in Spanish, with AE contrasting front [æ], mid-central [ʌ], and back [ɑ], where Spanish exhibits only a low central [a] [4, 17]. None of these AE low-mid vowels are acoustically identical to Spanish [a]; thus, acquiring the three-way contrast may present an acquisitional challenge to learners, depending on the perceptual similarities to Spanish vowels. Previous findings indicate that Spanish learners of English tend to perceive the English vowels [æ] and [ʌ] as good exemplars of native vowels, most frequently [a], although the exact assimilation patterns depend on what dialect of English learners are exposed to. [8]'s findings suggest that Spanish learners assimilate English [æ] to Spanish [a], instead of perceiving the English sound as a distinct L2 category (see also [13]). [6] reports that speakers of Peruvian Spanish assimilated Southern British English [æ] and [ʌ] to Spanish [a]. AE [æ] was assimilated to Spanish [e] and AE [ʌ] was variably perceived as Spanish [a] or [o]. The same population of speakers perceived Canadian English [æ] as Spanish [a] [7].

Given these cross-linguistic patterns of assimilation, we can expect that Spanish learners will have difficulties identifying and producing English [æ] and [ʌ] in a native-like fashion. However, we are not aware of a study which directly compared these two vowels in the perception and production of Spanish-speaking learners. The present study was designed to address this gap.

2. METHODS

2.1. Participants

Thirty native speakers of Colombian Spanish (9 F; 21 M) were tested on the campus of a major Midwestern university in the U.S. The mean age of participants was 28.3 years (SD=8.5), ranging from 19 to 49 years. Length of residence in the U.S. averaged 31.7 months; average age of arrival was 25.5 years. Participants reported, on average, a 'good' command of English pronunciation and listening (5 on a 7-point scale from 'very poor' to 'native-like').

Control production data was provided by seventeen monolingual native speakers of midwestern AE (14 F; 3 M). The mean age was 26 years, ranging from 18 to 57.

2.2. Elicitation prompts

The English words for the production task consisted of three CVC minimal pairs (*cap-cup, cat-cut, cab-cub*) and 13 fillers with comparable monosyllabic structure but different vowels. In all English words, the medial vowel was surrounded by stop consonants (e.g., *keep, cot*). English words were embedded in the carrier phrase *I say* ____ *again*. The monolingual English control group produced all but one of the target English words in citation form embedded in a list with 118 fillers.

The Spanish words were chosen to include all five Spanish vowels (three words per vowel) in closed syllables comparable to those of English monosyllabic words (stop-vowel-stop syllables carrying the primary stress), e.g., *cactus, boicot*. Multisyllabic words were used because of the limited availability of stops in coda position in Spanish [15]. Spanish words were embedded in the carrier phrase *Digo* ____ *ahora* ('I say ____ now').

In the perceptual identification task, fifteen minimal CVC triplets contrasted the vowels [æ], [ʌ] and a third vowel, e.g., *ban-bin-bun*. Each member of the triplet was spliced into the same carrier sentence for presentation, e.g. *Have you seen my cup/cape/cap?*. Seven native speakers of AE (3 F; 4 M) read the words used in the perceptual experiment, resulting in 45 sentences. (Every triplet was read by a single reader.) Different sentences were used for different triplets.

2.3. Procedure

Colombian learners of English performed a sentence reading task in Spanish and in English. The order was counterbalanced across participants. The control group of English speakers read the target words in citation form.

Elicitation prompts were presented one by one on a computer screen. All participants were instructed to read each prompt in a natural manner at a comfortable pace. Recordings were conducted in a sound-attenuated booth using an AudioTechnica AE4100 cardioid microphone and a TubeMP preamp. Items were randomized across trials; each participant saw each item three times. Colombian participants were given six seconds to read each sentence, and a 30 second break was offered approximately every 17 words. Monolingual English participants were given two seconds per word, with self-timed breaks offered after each full block.

Following the production task, Colombian participants took part in a four-choice identification task implemented in ExperimentMFC [3]. Sentences were presented on a computer screen in randomized order with target words replaced by an underscore (e.g. *Have you seen my ____?*) and presented auditorily via Sennheiser HD 380 Pro headphones. Each sentence was presented three times to each participant for a total of 135 trials, with breaks provided every 45 sentences. Each sentence was accompanied by four response choices: the three members of the tested triplet and an 'I don't know' option. Other vowels were not offered as a response option because they did not always combine with the given phonetic context to form a real word that was semantically appropriate for the question. No control group was tested perceptually because of the expected 'at ceiling' performance.

2.4. Measurements

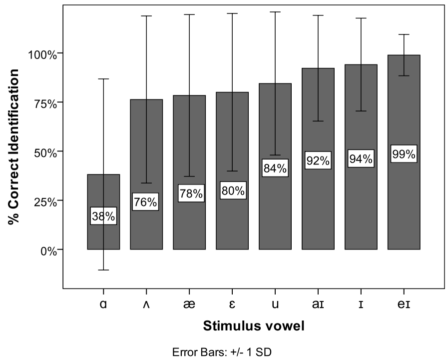
Measurements of the first, second, and third formant frequencies (F1, F2, and F3) were extracted at vowel midpoint using an automated procedure in Praat [3]. Values were checked for outliers and corrected manually when needed. To reduce individual differences, the Bark Difference Metric normalization was applied to F1 and F2 [22].

3. RESULTS

3.1. Perceptual identification task

Fig. 1 demonstrates the percent correct identification per stimulus vowel. Since filler vowels were presented with lower and unequal frequencies compared to [æ] and [ʌ], their identification results are not equally reliable. Nevertheless, the data in Fig. 1 suggests that identification was largely successful, eliciting 80% correct or higher for all vowels except [æ] and [ʌ], which hover just under 80% correct, and [ɑ], which was dramatically unsuccessful with only about 40% correct.

Figure 1: Percent correct identification per stimulus vowel.



Identification response, coded as correct or incorrect (‘I don’t know’ was coded as incorrect), was analyzed using a Generalized Linear Mixed Model (binary logistic regression) implemented in SPSS 24.0 [18] with Stimulus Vowel as a fixed factor and random intercepts for Subject, Subject by Stimulus Vowel, and Subject by Item. Deviation contrasts were also tested for the Stimulus Vowel factor (with sequential Sidak correction). The results showed that there was no significant difference in accuracy rates between [æ] and [ʌ] (β = .141, SE = .149, t = 0.95, p = .342). As deviation contrasts showed, neither the accuracy rate for [ʌ] nor that for [æ] differed from the overall mean (β = -.021, SE = .019, t = -1.114, p = .709 and β = .000, SE = .017, t = .024, p = .981, respectively). It is worth noting that the accuracy rate for [ɑ] was significantly below the overall mean (β = -.451, SE = .033, t = -13.702, $p < .001$).

Table 1 demonstrates the response choices for the most frequently misidentified vowels [æ], [ʌ], and [ɑ]. Response options with frequency under 5% are jointly coded as ‘other’.

Table 1: Response choices and their frequency for vowels [æ], [ʌ], and [ɑ].

		Response Vowel			
		[ʌ]	[ɑ]	[æ]	other
Stimulus Vowel	[ʌ]	76%	7%	12%	4%
	[ɑ]	47%	38%	13%	2%
	[æ]	15%	1%	78%	6%

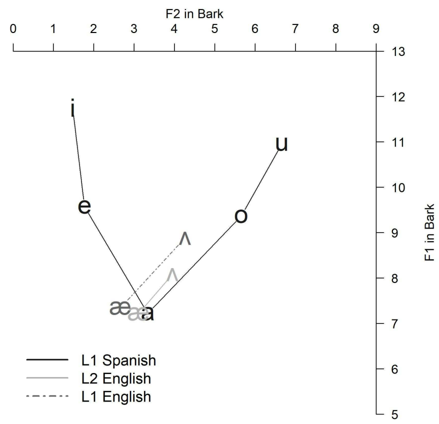
It is apparent that when [æ] and [ʌ] are misperceived, it is usually as each other. Low back [ɑ] is confused with both, but most frequently with [ʌ] (in fact, it was misidentified as [ʌ] more frequently than it was identified correctly).

Percentages of [æ], [ʌ], [ɑ], and ‘other’ responses were calculated per participant and analysed in a separate Univariate ANOVA for each stimulus vowel with Response as an independent variable, followed by post-hoc pairwise comparisons between the levels of Response (with Sidak adjustment). Response was a significant factor in both models ($F(3, 116) = 403.137$, $p < .001$ and $F(3, 116) = 342.667$, $p < .001$). The results of the pairwise comparisons for [æ] showed that the correct response was significantly more frequent than all misidentifications and that misidentification as [ʌ] was significantly more frequent than other errors. Pairwise comparisons for [ʌ] showed that the correct response was significantly more frequent than all misidentifications and misidentification as [æ] was significantly more frequent than ‘other’, but there was no significant difference between the frequency of [æ] and [ɑ] errors.

3.2. Production Task

Fig. 2 demonstrates the location of the English vowels [æ] and [ʌ] pronounced by Colombian learners and native speakers of English within the Spanish vowel space of Colombian participants.

Figure 2: Spanish (dark solid line) and English (light solid line) vowels produced by Colombian learners of English and English vowels produced by native speakers of American English (dashed line).



It is apparent that the English vowels of Colombian learners are lower and more central than native English vowels. Learners’ [æ] overlaps with Spanish [a].

Bark-transformed F1 and F2 (B(F1) and B(F2)) of learners’ and native English speakers’ vowels [æ] and [ʌ] were analysed in a Linear Mixed Model with Native Language (Spanish vs. English) and Vowel ([æ] vs. [ʌ]) as fixed factors and random intercepts for Subject and Item. (Separate models were implemented for B(F1) and B(F2).) The results showed that Vowel was a significant factor in both models: B(F1) and B(F2) were higher for [ʌ] than for [æ], predictably indicating that [ʌ] was higher and more retracted than [æ] ($F(1, 6.151) = 6.151$, $p = .001$ and $F(1, 9.872) = 104.369$, $p < .001$). The effect of Native Language across the two vowels was not significant in either model. The interaction between Native Language and Vowel was significant in both models ($F(1, 587.684) = 6.953$, $p = .009$ and $F(1, 583.767) = 16.073$, $p < .001$). To determine the source of the interactions, each vowel was tested separately for the effect of Native Language on B(F1) and B(F2). The results showed that B(F2) of learners' [æ] was significantly higher than B(F2) of native speakers' [æ], indicating a less fronted vowel ($F(1, 279.274) = 10.757$, $p = .001$). B(F1) and B(F2) of learners' [ʌ] were significantly lower than those for native speakers' [ʌ], indicating a lower and more fronted vowel ($F(1, 295.962) = 17.025$, $p < .001$ and $F(1, 294.557) = 4.683$, $p = .031$). In each case, learners' [æ] and [ʌ] appear to be moving towards Spanish [a], which is lower and more central than English [æ] and [ʌ].

Learners' English vowels were also compared to their own Spanish [a] using Linear Mixed Models with Vowel ([æ] vs. [ʌ] vs. [a]) as a fixed factor and random intercepts for Subject, Item, and Subject by Vowel (B(F1) and B(F2) were analyzed in separate models), followed by pairwise comparisons with Sidak correction. The results showed that Vowel was a significant factor in both models ($F(2, 14.2) = 11.008$, $p = .001$ and $F(2, 8.95) = 4.915$, $p = .036$). Pairwise comparisons determined that learners' vowels [æ] and [ʌ] were significantly different from each other both in terms of height ($p = .003$) and backness ($p = 0.042$). Learners' English [ʌ] was also significantly higher than their Spanish [a] ($p = .003$). However, their English [æ] did not differ significantly from their Spanish [a] in height or in backness.

4. DISCUSSION

Perceptual results demonstrated that participants tended to be somewhat less accurate at identifying [æ] and [ʌ] than at identifying other English vowels, but not significantly so. Interestingly, only the vowel [ɑ] triggered a significant decrease in identification accuracy, but since [ɑ] was not a target of the experiment, this result can only be interpreted as tentative. Nevertheless, when [æ] and [ʌ] were misidentified, they were confused with each other more frequently than with other vowels, suggesting that Colombian learners occasionally perceive them as exemplars of the same category. The production data shed further light on this possibility.

While learners produced English [æ] and [ʌ] as acoustically distinct, both vowels were statistically different from native targets (which could be partly attributable to different elicitation methods). The direction of difference from English targets suggests that both vowels are pulled towards Spanish [a], becoming lower and more central. In fact, learners' [æ] was not statistically different from their Spanish [a], and their [ʌ] differed from [a] only in height.

Acoustic convergence between learners' English [æ] and [ʌ] and their Spanish [a] suggests that these three vowels are not fully separated into distinct categories in the cognitive phonetic space of Colombian learners. Taking into account previous findings [8], it is plausible that perceptual assimilation of English [æ] and [ʌ] into the Spanish [a] category is the source of this acoustic convergence. If this is the case, the same cross-linguistic assimilation pattern can explain the confusability between [æ] and [ʌ] in the perceptual results.

As frequently noted in previous literature [20, 21], perceptual assimilation between L1 and L2 sounds is not always attributable to acoustic proximity. As Fig. 2 shows, English [ʌ] is quite distant from Spanish [a]. One possibility is that English low-mid vowels are phonologically assimilated to Spanish [a], despite acoustic distinctiveness [2], due to the 'inability' to form additional categories of low vowels in the L1-based cognitive grammar of the interlanguage [5]. This phonological assimilation could be further encouraged by the orthographic links between the two languages, including cognates (e.g., *canal*), wherein English [æ] is realized in spelling as <a> (e.g., *bat*), the same grapheme used for Spanish [a]. As research shows, L2 orthography plays an important role in guiding (and misguiding) acquisition of L2 phonology [16, 23].

Thus, despite being highly experienced and proficient speakers of English, Colombian learners face challenges related to perception and production of AE low-mid vowels. Learners' differences from native speakers were more apparent in production than in perception, where, despite occasional mutual confusability, [æ] and [ʌ] were identified with high accuracy. The data therefore suggest that these Colombian learners may be more advanced in their perception than in their production skills, as predicted by the perception-production link hypothesis [11, 12].

5. REFERENCES

[1] Barrios, S., Jiang, N., Idsardi, W. J. 2016. Similarity in L2 phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts. *Second Language Research* 32(3), 367–395.

[2] Best, C. T., Tyler, M. D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In: Bohn, O.-S., Munro, M. (eds). *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege.* Amsterdam: John Benjamins, 13–34.

[3] Boersma, P., Weenink, D. 2001. Praat. a system for doing phonetics by computer. Glot International 5:9/10, 341-345.

[4] Bradlow, A. R. 1995. A comparative acoustic study of English and Spanish vowels. *The Journal of the Acoustical Society of America* 97(3), 1916–1924.

[5] Brown, C. A. 1998. The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14(2), 136–193.

[6] Escudero, P., Chládková, K. 2010. Spanish listeners' perception of American and Southern British English vowels: Different initial stages for L2 development. *The Journal of the Acoustical Society of America* 128(5), EL254–EL260.

[7] Escudero, P., Vasiliev, P. 2011. Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *The Journal of the Acoustical Society of America* 130(5), EL277–EL283.

[8] Flege, J. E. 1991. The interlingual identification of Spanish and English vowels: Orthographic evidence. *The Quarterly Journal of Experimental Psychology* 43(3), 701–731.

[9] Flege, J. E. 1995. Second-language speech learning: Theory, findings, and problems. In: Strange, W. (ed). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research.* Timonium, MD: York Press, 233–277.

[10] Flege, J. E. 2003. Assessing constraints on second-language segmental production and perception. In: Meyer, A., Schiller, N. (eds). *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities.* Berlin: Mouton de Gruyter, 319–355.

[11] Flege, J. E., Bohn, O.-S. 1997. Perception and production of a new vowel category by adult second language learners. In: James, A., & Leather, J. (eds). *Second-Language Speech: Structure and Process (Studies on Language Acquisition)* 13, 53.

[12] Flege, J. E., Bohn, O.-S., Jang, S. 1997. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25(4), 437–470.

[13] 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.

[14] Flege, J. E., Munro, M. J., MacKay, I. R. 1995. Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America* 97(5), 3125–3134.

[15] Harris, J. 1969. *Spanish Phonology*. Cambridge, MA: MIT.

[16] Hayes-Harb, R., Brown, K., Smith, B. L. 2017. Orthographic input and the acquisition of German final devoicing by native speakers of English. *Language and Speech* 61(4), 547–564.

[17] Hillenbrand, J., Getty, L. A., Clark, M. J., Wheeler, K. 1995. Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America* 97(5), 3099–3111.

[18] IBM Corp., 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk (NY): IBM Corp.

[19] Ostalski, P. 2009. Back vowels in British and American English. *Participations and Assessments* 5(4), 105–118.

[20] Strange, W., Bohn, O.-S., Nishi, K., Trent, S. A. 2005. Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *The Journal of the Acoustical Society of America* 118(3), 1751–1762.

[21] Strange, W., Bohn, O.-S., Trent, S. A., Nishi, K. 2004. Acoustic and perceptual similarity of North German and American English vowels. *The Journal of the Acoustical Society of America* 115(4), 1791–1807.

[22] Traunmüller, H. 1997. Auditory scales of frequency representation. http://www.ling.su.se/staff/hartmut/bark.htm. Retrieved 14 December 2018.

[23] Vera Diettes, K. J. 2014. Fenómenos de reducción vocálica por hablantes colombianos de inglés como l2: Un estudio acústico. *Forma y función* 27(1), 11–43.