# SCHWA TO /A/: DEVELOPMENT OF UNSTRESSED VOWELS IN L2 SPANISH

Robert Esposito

Department of Spanish and Portuguese, Rutgers University
rme70@scarletmail.rutgers.edu

## ABSTRACT

Native (L1) English second language (L2) Spanish learners acquiring the Spanish vowel system are tasked with adjusting entrenched vowel categories and morphophonological rules. For example, allophonic variation of stressed and unstressed vowels in English is not present in Spanish, but the triggering environments for the variation is present in Spanish. As such, learners of Spanish must come to ignore those triggering environments during perception and production. The present study examines the production of unstressed Spanish /a/ by 10 L1 English L2 Spanish learners during a seven-week domestic immersion program in a semi-longitudinal study. Preliminary acoustic analysis revealed that within the first week, participants fronted and lowered their production of unstressed /a/, assumed here to correspond to less centralization and more target-like production. The findings are situated within current models of second language speech learning and provide often-sought longitudinal data on an understudied phenomenon.

**Keywords:** phonetics, phonology, second language acquisition, Spanish, English

## 1. INTRODUCTION

Adult native English (L1) learners of Spanish as a second language (L2) face difficulties producing target-like Spanish vowels. Each of the five Spanish vowels /i e a o u/ is produced in a vowel space that corresponds to multiple English vowel categories, so learners must ignore phonetic differences relevant to their L1. English vowels also undergo morphophonological processes which do not exist in Spanish. As such, L1 English L2 Spanish learners begin the endeavor of learning Spanish with expectations from their L1, which shape their perception and production and lead to non-target-like production of vowels. For example, a learner may produce the Spanish word /ˈka.sa/ as non-target-like [ˈka.sə] due to a morphophonological centralization process in English in which unstressed vowels tend to be reduced to a centralized vowel. The present study investigates the production of Spanish unstresed /a/ in 10 L1 English L2 Spanish beginner learners during a seven-week domestic immersion program in a semi-longitudinal design. Instead of extracting formant values at solely the midpoint of the vowel, acoustic analysis focuses on the trajectory of formant values throughout the temporal duration of the vowel to obtain a more holistic view of vowel centralization. Results will provide information about typical development of L1 English L2 Spanish beginner learners with growing L2 exposure and how learners adjust their production in the face of growing evidence against a process present only in the L1.

## 2. LITERATURE REVIEW

### 2.1. Spanish and English Vowels

Reported variation in Spanish varieties should not be ignored in general [1, 2, 3], but vowel production has been found to be relatively uniform across varieties [4, 5], whereas descriptions of English vowel systems vary more widely [6, 7]. Spanish can be said to have a subset of English vowel phonemes /i e a o u/, but the phonetic realizations of these phonemes are not identical.

A major difference to highlight between the two systems is the impact of prosodic effects, such as lexical stress, on the realization of vowels. One of the major acoustic correlates of lexical stress cross-linguistically is vowel reduction in unstressed syllables [8]. Spanish has been found to have minimal centralization of vowels in unstressed syllables [9, 4], representing a possibly universal low-level phonetic phenomenon [10]. On the other hand, English presents both low-level phonetic reduction like Spanish, as well as a language-specific, morphophonologic, categorical, rule-based process in which vowels in lexically unstressed syllables are realized as centralized vowels [ɪ] or [ə] [11]. For example, the first vowel in the English word ‘atom’ is realized as [æ] in the tonic syllable, but when realized in a non-tonic syllable like in the derived word ‘atomic’, the vowel is obligatorily reduced to [ə].

### 2.2. L1-L2 Transfer Effects

L1 English L2 Spanish learners typically do not present major issues acquiring phonological vowel contrasts [12], but the phonetic and morphophonological differences between their L1 and L2 may result in non-target-like production [13, 14, 15, 16, 17]. The acquisition path of Spanish vowels by L1 English L2 Spanish learners can be accurately described by a number of frameworks such as the Speech Learning Model revised [18], the Perceptual Assimilation Model [19], or, as will be explored here, the Second Language Linguistic Perception (L2LP) model [20, 21, 22, 23, 24].

L2LP proposes that learners of a L2 create a new grammar via *full copying/full access*. Accordingly, difficulties in acquiring L2 speech sounds are accounted for by phonetic similarities, differences, or perceived equivalences to contrasts present in the L1. The learner must adjust their perception grammar to account for L2 input. Furthermore, it is maintained that perception precedes and motivates production [25]. Based on this framework, it is expected linguistic processes present in the L1 may undesirably surface in the L2, but that these transfer effects may be reduced with increasing proficiency and exposure to the L2. For example, [26] found that L1 English L2 Spanish late learners over a seven-week home-immersion program shifted perception and production of voice onset time categorical boundaries from initially English-like towards more Spanish-like values.

As mentioned, English presents vowel allophony in lexically stressed and unstressed positions not present in Spanish. In a cross-sectional study, [15] found that L1 English L2 Spanish late learners of Spanish, divided into "intermediate" and "advanced" proficiency groups, produced unstressed Spanish vowels as more centralized than L1 Spanish controls. Although valuable, [21] urges for the inclusion of longitudinal data from beginning language learners. Longitudinal L2 acquisition data has provided evidence, for example, that the formation of L2 phonetic categories can occur abruptly at early stages, and they are particularly susceptible to cross-linguistic influence [25].

Furthermore, [15], similar to many other vowel production studies, collected formant values at only the vowel midpoint. The midpoint is most likely the least centralized portion of the vowel, and thus may not present the full story. For example, [27] calculates a "spectral centroid", a measure calculated from formant values at 5 equidistant temporal locations in the vowel. Similarly, [28] suggests collecting formant values at equidistant temporal locations in the vowel and to model the vowel's trajectory instead of calculating a single-point value.

## 3. THE PRESENT STUDY

In sum, previous research has pointed to the difficulties of L1 English L2 Spanish target vowel production, particularly when it comes to unstressed vowels due to L1-L2 transfer effects. Available studies typically present cross-sectional designs and measure only single-point values. From an L2LP perspective, cross-sectional designs only tell part of the story, and longitudinal data will reveal how the L2 grammar is progressively adjusted as the learner receives more input. Methodologically, a more holistic view of vowels via the collection of various formant values along the vowel's temporal realization may reveal fine-grain phonetic patterns not seen with a single-point value. As such, the present study tracks the production of unstressed /a/ in 10 L1 English L2 Spanish beginner late learners during a seven-week domestic immersion program in a semi-longitudinal design. With this data, the following research questions (RQs) are hoped to be answered:

1. Do L1 English L2 Spanish learners centralize unstressed /a/?
2. Do L1 English L2 Spanish learners reduce centralization of unstressed /a/ as they progress through a 7-week domestic immersion program?

Based on the L2LP [20, 21, 22, 23, 24], it is expected that learners will produce unstressed /a/ due to lexical stress effects present in the L1 [11, 14, 15]. It is hypothesized that centralization of unstressed /a/ will reduce over the seven-week period [26], reflecting gains in proficiency and increased exposure to the L2 (although the present study cannot differentiate individual contributions of these effects). Furthermore, as has been seen in research on other segmentals [25], it is hypothesized that participants may demonstrate abrupt, as opposed to gradual, development in target-like L2 production of unstressed /a/.

## 4. METHODOLOGY

### 4.1. Participants

Ten L1 English L2 Spanish late learners (females = 6) were recruited for the present study. They reported no prior experience with any foreign languages, nor had they spent a significant amount of time in a foreign country. Participants were enrolled in the Middlebury Language School program's beginner classes, placed via a placement test and interviews with two faculty members of the program. The Middlebury

Language School program is a seven-week domestic immersion program where students are encouraged and quasi-obligated to use their target language at all times. During the seven-week program, participants received high amounts of L2 input (from native and non-native speakers) and reported minimal L1 use.

### 4.2. Materials

Participants completed a delayed repetition task. The following sections describe the target words, the auditory stimuli, and recordings.

#### *4.2.1. Target Words*

The analyzed materials consisted of 18 real and nonce words. Each word was a bisyllabic paroxytone with /a/ as the final vowel. While the first syllable variably contained an onset, the second syllable always contained an onset (e.g., ada, ita, gaka, gota).

#### *4.2.2. Auditory Stimuli*

A twenty-nine year old native female spanish speaker from Cádiz, Spain produced the auditory stimuli. The 18 items were listed on a sheet of paper and the speaker read the list aloud in the carrier phrase 'X es la palabra' (*X is the word*). A Shure SM10A dynamic head-mounted microphone was used to record the items. A Sound Devices MM-1 pre-amplifier boosted the signal and sent it to a laptop computer where it was recorded using Praat at a 44.1 kHz sample rate with a 16-bit quantization [29]. The recording took place in a sound attenuated booth in the Arizona Applied Phonetics Laboratory at the University of Arizona.

#### *4.2.3. Recordings*

Participants were recorded in a quiet classroom on site at Middlebury College. The same procedure described for the auditory stimuli was used.

### 4.3. Acoustic Analysis

Participant recordings were transcribed and segmented automatically via the Montreal Forced Aligner [30]. Incorrect segmentations (e.g., unaligned or misidentified phones) were manually corrected. The Praat script Fast Track [31] was then used to extract formant values at 3 ms intervals. As recommended by the author, the script was set to optimize extraction of F1 and F2, as those are the formants of interest. Due to vowel durational differences, vowel duration was normalized and binned into 11 bins. The average F1 and F2 values were then calculated for each bin.

### 4.4. Procedure

Participants completed the first experimental session on the second or third day of the program; further experimental sessions took place every Sunday, for a total of 7 sessions. PsychoPy2 presented the experimental stimuli randomly. Stimuli were presented aurally via the native Spanish speaker recordings in the first session. Participants listened to and repeated the 18 items 3 times. In the following sessions, participants were visually presented the stimuli. After listening producing the item, participants pressed a button on a keyboard to advance to the next item. Each participant provided the dataset with 378 items for a total of 3780 tokens (18 words * 3 repetitions * 7 weeks * 10 participants). The participants finished the task in approximately 10 minutes.

### 4.5. Statistical Analysis

The present analysis is descriptive. Vowel trajectories were averaged across all participants for each of the seven weeks.

## 5. RESULTS

Fig. 1 displays the averaged unstressed /a/ trajectories of the 10 participants over the seven-week domestic immersion program. Visually, it can be seen that by the second session (a week into the program), participants' unstressed /a/ trajectories started from a more fronted point and ended at a lower point compared to the first session. This adjustment was maintained throughout the 7 weeks.

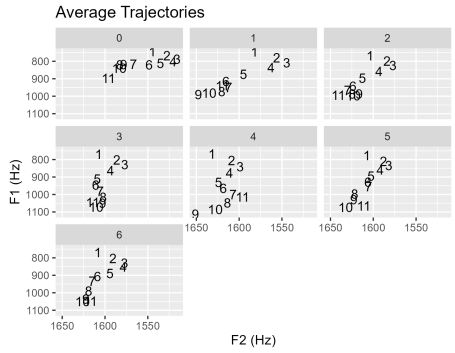


**Figure 1:** Averaged trajectories of unstressed /a/ from 10 participants over seven weeks. Each individual point represents the linear, binned temporal moment in the vowel (i.e., 1 = start of vowel, 11 = end of vowel).

## 6. DISCUSSION

The present work examined the production of unstressed /a/ in L1 English L2 Spanish beginner learners in a semi-longitudinal design. Based on L2LP [20, 21, 22, 23, 24], it was expected that these learners would present centralization of unstresesd /a/ due to L1-L2 transfer [11, 14, 15], but transfer effects would be mitigated with increasing L2 proficiency and exposure.

Preliminary data suggests that in the first session, participants produced more centralized vowels than in later sessions. This can be seen in the fronting and lowering seen in later sessions. As such, the current data tentatively supports the hypothesis that L1 English L2 Spanish speakers would initially present centralized unstressed /a/ in Spanish, as well as develop more target-like uncentralized /a/ as the program progressed. In fact, the development occurred within the first week and was maintained throughout all sessions, suggesting an abrupt adjustment [25].

## 7. LIMITATIONS AND FUTURE DIRECTION

There a number of improvements that can be made to the current study.

First, there is available data on the participants' productions of stressed /a/ that should be analyzed in comparison to their unstressed /a/. A measure such as Euclidean distance [28] can be used as a more robust measure of centralization.

Second, formant values were collected every three miliseconds, vowel durations were normalized, and formant values were binned. This is not a common practice, and a more robust methodology should be taken. For example, collecting formant values at some amount of equidistant points along the vowel [28, 27].

Third, no statistical analysis was performed. Vowel trajectory presents an opportunity to use a Generalized Additive Mixed Model (GAMM), which allows for modelling of non-linear data such as vowel trajectories [28]. By modeling the vowel trajectories, a deeper understanding of where meaningful differences occur may arise.

Lastly, the L2LP posits that perception precedes production [22], but the current data cannot provide evidence on this topic. In a case like this, perceptual boundaries of /a/ versus /ə/ could be investigated in L1 English L2 learners [32].

## 8. CONCLUSION

The present study analyzed the production of unstressed /a/ in 10 L1 English L2 Spanish beginner learners in a domestic immersion program in a semi-longitudinal design over seven weeks. The study deepens our understanding of how L1 morphophonological processes not present in the L2 may arise in L2 production, as well as presenting a less common way of analyzing vowel production by looking at the entire spectral envelope, instead of only a single-point value. The longitudinal data suggests that lexical stress effects in the L1 can be abruptly curtailed in the L2 with growing L2 proficiency and exposure.

### 8.1. References

## 9. REFERENCES


[1] E. W. Willis, "An initial examination of southwest spanish vowels." *Southwest Journal of Linguistics*, vol. 24, no. 1-2, pp. 185–199, 2005.

[2] J. M. Lipski, "Tracing mexican spanish/s:: A cross-section of history," *Language Problems and Language Planning*, vol. 18, no. 3, pp. 223–241, 1994.

[3] M. C. Alba, "Accounting for variability in the production of spanish vowel sequences," in *Selected proceedings of the 9th Hispanic Linguistics Symposium*. Cascadilla Press Somerville, MA, 2006, pp. 273–285.

[4] T. N. Tomás, *Documentos lingüísticos del alto Aragón*. Syracuse University Press, 1957, vol. 4.

[5] A. Quilis, "Tratado de fonología y fonética españolas," *(No Title)*, 1993.

[6] P. Ladefoged, "A course in phonetics (5ª edición), boston, ma: Thomson wadsworth," ISBN 1-4130-2079-8, Tech. Rep., 2006.

[7] A. R. Bradlow, "A comparative acoustic study of english and spanish vowels," *The Journal of the Acoustical Society of America*, vol. 97, no. 3, pp. 1916–1924, 1995.

[8] M. Gordon and T. Roettger, "Acoustic correlates of word stress: A cross-linguistic survey," *Linguistics Vanguard*, vol. 3, no. 1, p. 20170007, 2017.

[9] E. Martínez Celdrán, "Fonética:(con especial referencia a la lengua castellana)," *Teide*, 1984.

[10] V. Kapatsinski, S. Easterday, and J. Bybee, "Vowel reduction: a usage-based perspective," *Rivista di Linguistica*, vol. 32, no. 1, pp. 19–42, 2020.

[11] J. Bybee, *Phonology and language use*. Cambridge University Press, 2003, vol. 94.

[12] G. S. Morrison, "Perception and production of spanish vowels by english speakers," in *Proceedings of the 15th international congress of phonetic sciences*, vol. 2003. Causal Productions Adelaide, 2003, pp. 1533–1536.

[13] A. C. Aldrich, *Acquisition of L2 phonology: An acoustic analysis of the centralization of L2 Spanish/a/in adult L1 English-speaking learners*. Brigham Young University, 2014.

[14] K. Cobb, *La pronunciación de vocales átonas en español: La aplicación de reglas fonológicas por parte de hablantes no-nativos del español*. The University of Arizona, 2009.

[15] K. Cobb and M. Simonet, "Adult second language learning of spanish vowels," *Hispania*, vol. 98, no. 1, pp. 47–60, 2015.
[16] A. Iruela, "Adquisición del vocalismo español por holandeses: Análisis en estilo semiespontáneo," *Journal of Experimental Phonetics*, vol. 9, pp. 135–181, 1997.
[17] M. R. Menke and T. L. Face, "Second language spanish vowel production: An acoustic analysis." *Studies in Hispanic & Lusophone Linguistics*, vol. 3, no. 1, 2010.
[18] J. E. Flege, K. Aoyama, and O.-S. Bohn, "The revised speech learning model (slm-r) applied," *Second language speech learning: Theoretical and empirical progress*, pp. 84–118, 2021.
[19] C. T. Best, M. Tyler, O. Bohn, and M. Munro, "Nonnative and second-language speech perception," *Language experience in second language speech learning*, vol. 17, pp. 13–34, 2007.
[20] P. Escudero and P. Boersma, "Bridging the gap between l2 speech perception research and phonological theory," *Studies in second language acquisition*, vol. 26, no. 4, pp. 551–585, 2004.
[21] P. R. Escudero Neyra, *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Utrecht University & LOT, 2005.
[22] P. Escudero, "Second-language phonology: The role of perception," in *Phonology in context*. Springer, 2007, pp. 109–134.
[23] P. Escudero *et al.*, *The linguistic perception of similar L2 sounds*. na, 2009.
[24] J.-W. Van Leussen and P. Escudero, "Learning to perceive and recognize a second language: The l2lp model revised," *Frontiers in psychology*, vol. 6, p. 1000, 2015.
[25] J. V. Casillas, "Phonetic category formation is perceptually driven during the early stages of adult l2 development," *Language and Speech*, vol. 63, no. 3, pp. 550–581, 2020.
[26] ——, "The longitudinal development of fine phonetic detail in late learners of spanish," Ph.D. dissertation, The University of Arizona, 2016.
[27] E. Jacewicz, R. A. Fox, and J. Salmons, "Vowel change across three age groups of speakers in three regional varieties of american english," *Journal of phonetics*, vol. 39, no. 4, pp. 683–693, 2011.
[28] S. Coretta and J. V. Casillas, "A tutorial on generalised additive mixed effects models for bilingualism research," *Linguistic Approaches to Bilingualism*, 2024.
[29] P. Boersma and D. Weenink, "Praat: doing phonetics by computer," 2025, computer program. Retrieved 7 November 2025. [Online]. Available: https://praat.org
[30] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi," in *Proc. Interspeech 2017*, 2017, pp. 498–502.
[31] S. Barreda, "Fast track: fast (nearly) automatic formant-tracking using praat," *Linguistics Vanguard*, vol. 7, no. 1, 2021.
[32] J. D. Miller, "Auditory-perceptual interpretation of the vowel," *The journal of the Acoustical society of America*, vol. 85, no. 5, pp. 2114–2134, 1989.