Adult Second Language Learning of Spanish Vowels

Katherine Cobb, Miquel Simonet

Hispania, Volume 98, Number 1, March 2015, pp. 47-60 (Article)

Published by Johns Hopkins University Press
DOI: https://doi.org/10.1353/hpn.2015.0026

Adult Second Language Learning of Spanish Vowels

Katherine Cobb
University of Arizona

Miquel Simonet
University of Arizona

Abstract: The present study reports on the findings of a cross-sectional acoustic study of the production of Spanish vowels by three different groups of speakers: 1) native Spanish speakers; 2) native English intermediate learners of Spanish; and 3) native English advanced learners of Spanish. In particular, we examined the production of the five Spanish phonemic monophthongs (/i e a o u/) in two prosodic positions: stressed and unstressed. Since Spanish and English vary in the nature of the difference between stressed and unstressed vowels, a logical question is: How does the relationship between stressed and unstressed vowels differ across the three groups of participants? Do the learners transfer English patterns to their Spanish? The findings are discussed in terms of their significance for current models of second language speech learning and provide important information regarding the paths of pronunciation development in the second language acquisition of Spanish.

Keywords: acoustics/acústica, phonetics/fonética, second language acquisition/adquisición de segundas lenguas, stress/acento, vowels/vocales

Introduction

Spanish has five phonemic vowels (/a e i o u/), as evidenced by the following word list: *paso*, *peso*, *piso*, *poso*, *puso* (Hualde, 2005: 120–34, among others). General American English, on the other hand, has around fourteen vocalic phonemes, including monophthongs and diphthongs, as evidenced by the following word list: *beat*, *bit*, *bait*, *bet*, *bite*, *bat*, *bout*, *pot*, *bought*, *boy*, *boat*, *but*, *put*, *boot*, although dialectal differences are enormous (Giegerich 1992: 47; Stockwell and Bowen 1965: 86–115; among others). Whereas a native speaker of Spanish learning English would have to learn a more crowded and complex vowel system, a native speaker of English learning Spanish would have to learn a simpler and sparser one. Although the two learning routes are very different from each other, they both present serious obstacles.

Adult native speakers of Spanish learning English experience enormous difficulties in their acquisition of English vowel phonemes. In most cases, obstacles come from having to learn an English-specific vowel contrast; that is, the obstacle is due to having to partition the vowel space corresponding to a single Spanish vowel into two (or more) subsets, corresponding to two (or more) different English vowels. For instance, Spanish-speaking learners of English are limited in their production and perceptual discrimination of the English /i/-/ɪ/ contrast (Morrison 2008, 2009). Casillas (2012) has recently shown that the English /ɑ/-/æ/ contrast presents even more serious difficulties to these learners than the /i/-/ɪ/ does. This is an instance of phonetic as well as phonological learning in the sense that learners need to acquire two phonemes with their corresponding lexical sets and acoustic distributions. Best's (1995) Perceptual Assimilation Model or PAM proposes that patterns of interlingual assimilation (equivalence classification of L2 to L1 vowels) predict sound discrimination in cross-language listening. Discriminatory constraints may, in turn, predict acquisition difficulties in L2 speech learning (Best and Tyler 2007). For instance, a Spanish-speaking learner of English may assimilate both English /i/ and English /ɪ/ to Spanish /i/; for as long as this assimilation continues, discrimination and learning of the English /i/-/ɪ/ contrast will be blocked. On the other hand, acquiring the English /i/-/ɛ/ contrast may prove easier for Spanish speakers, as they may assimilate English /i/ to Spanish /i/ and English /ɛ/ to Spanish /e/ (Casillas 2012).

Native English speakers learning Spanish have a challenge of a very different nature. In their case, instead of having to partition their native vowel space into smaller bins, the task consists of learning to ignore phonetic differences that may be relevant for English but are not so for Spanish. The interlingual assimilation of the five Spanish vowels will most likely create one-to-one correspondences between these vowels and the native English vowels. For instance, Spanish /i/ may be assimilated to English /i/ (or /ɪ/), Spanish /e/ may be assimilated to English /ɛ/ (or /eɪ/), Spanish /a/ may be assimilated to English /ɑ/ (or /æ/), Spanish /o/ may be assimilated to English /oʊ/ and, finally, Spanish /u/ may be assimilated to English /u/ (or /ʊ/). If each Spanish vowel corresponds to a different English vowel, the PAM predicts high discriminatory abilities in the L2 and, therefore, relatively easy phonological acquisition (Best 1995). It might be relatively easy to learn the system of contrasts and the lexical sets associated with the phonemes. However, the acquisition of phonetic detail (i.e., native-like acoustic distributions) is an issue of a different sort.

Flege's (1995) Speech Learning Model or SLM is a theoretical framework that explicitly addresses the L2 acquisition of phonetic detail. According to the SLM, success in phonetic development is determined by the acoustic-perceptual difference between L2 sounds and their equivalent L1 sounds. The SLM shares with the PAM the understanding that L2 learners and naïve listeners assimilate novel or L2 sounds to their closest L1 sound. The SLM goes slightly beyond this and claims that, perhaps contrary to intuition, the larger the difference between a novel sound and its equivalent L1 sound the easier it will be for learners to overcome this pattern of interlingual assimilation and, therefore, the easier it will be for them to develop a separate sound category for the L2 sound. On the other hand, smaller differences between L1 and L2 sounds may make it difficult for learners to overcome interlingual assimilation, which will in turn block their progress. This model, thus, explicitly addresses the nature of new category formation. In sum, in addition to interlingual phonological assimilations, the phonetic distance between the assimilated categories may affect the patterns of new category formation and the retention of a nonnative accent. Therefore, studying the acquisition of Spanish vowels by native speakers of English is of high theoretical relevance for our understanding of speech learning in L2 acquisition.

The acquisition of Spanish vowels has received relatively little attention in applied linguistics literature (see Menke and Face 2010; cf. Simonet 2012 for an overview). This is especially evident if we take into consideration the attention other phonetic features, such as voice onset time (VOT) in voiceless stops, have received (e.g., Zampini and Green 2001). The opinion persists that phonetic variation in Spanish is to be found mostly in the consonants and not in the vowels. It is to be expected that English-speaking learners of Spanish will not experience any major difficulties in acquiring a phonological representation of Spanish vowel contrasts (Morrison 2003). However, this is not to say that the phonetic substance of the vowel phonemes produced by these learners will match that of native Spanish speakers. This is obvious to many native Spanish speakers who converse with second language learners whose native language is English. The present study seeks to capture the nature of these differences with an acoustic investigation.

An additional and important aspect of the present study is that it explores potential acoustic variation in the vowels of Spanish that might be due to (English) prosodic effects, lexical stress in particular. Stressed as well as unstressed vowels are compared across three groups of participants—native Spanish speakers, native English intermediate learners of Spanish, and native English advanced learners of Spanish—in order to explore the effects of the well-known English phenomenon of 'unstressed vowel reduction.' The latter term has at least two meanings. It refers to 1) a phonetic, performative process according to which vowels are hypoarticulated with respect to neighboring vowels as a function of their prosodic position, such as in a lexically unstressed syllable; and 2) a morphophonological, categorical, rule-based process according to which a symbolic allophone is selected as a variant of a phoneme in specific prosodic positions, such as in a lexically unstressed syllable (Barnes 2006; Fourakis 1991; Lindblom 1963). In other words, while the first process triggers low-level phonetic differences between stressed and unstressed vowels, the second process leads to morphophonological alternations; while the first process might be universal and grounded on articulatory patterns, the second one is language-specific and grounded on the morphology (Bybee 2001). For instance, the second vowel in apply ([aɪ]) is not available in application ([ɪ] or [ə]), the second vowel in emphatic ([æ]) is not available in emphasis ([ɪ] or [ə]). In English, these vowel alternations are conditioned by the position of stress in an exceptionless process. English, therefore, has a process of morphophonological unstressed vowel reduction. Spanish, on the other hand, does not: the second vowel in *énf**a**sis* ([a]) is the same as the vowel in *enf**á**tico* ([a]).

With this in mind, the current study seeks to answer the questions: Do native speakers of English acquiring Spanish tend to apply the vowel reduction rule of their native language to their second language? Do these patterns differ as a function of the vowels involved?

Flege and Bohn (1989) investigated the production of stressed-unstressed vowel alternations in English (emphasis~emphatic) by native speakers of Spanish. It was found that Spanish learners of English tended to display a more modest pattern of vowel reduction; that is, there were smaller acoustic differences between stressed and unstressed vowels in English alternations in the speech of Spanish learners of English than in the productions of the English controls, as expected. Notice, however, that Flege and Bohn examine unstressed vowel reduction from the opposite direction we do in the present study. To our knowledge, this aspect of second language speech learning has received very little attention.

Menke and Face (2010) analyzed the acoustics of stressed and unstressed Spanish vowels as produced by a large group of speakers divided in four groups: native speakers of Spanish, and beginning, intermediate and advanced learners of Spanish whose L1 was English. Their study, although very large in scope, was limited to testing for statistical significance (hit, fail) between stressed and unstressed vowels for the four groups of subjects separately. Their results were inconclusive. First, they found, for native speakers, that stressed and unstressed /e/ and /u/ differed in fronting. No other vowels (or dimensions) differed as a function of stress for these subjects. For the beginners, these authors found effects of stress in at least one dimension for /e a u/ but not for /o i/. For the intermediate learners, it was found that stressed and unstressed /i e a u/ differed from each other, in at least one dimension, but this was not the case for /o/. Finally, advanced learners showed effects of stress, in at least one dimension, for all the vowels. In sum, it would appear that, with experience, learners become 'worse,' not 'better'; that is, differences between native controls and learners increase in a scalar progression as a function of experience. Furthermore, the multiplicity of statistical tests in this study may have led to some random results. It is our opinion that this issue merits further research.

Methods

Participants

A total of 15 female volunteers were recruited for participation and divided into three groups: 1) a group of native Spanish speakers or controls (N = 5); 2) a group of native speakers of English, proficient late L2 learners of Spanish (N = 5); and 3) a group of English native speakers, intermediate late L2 Spanish learners (N = 5). The two learner groups differed in a number of ways: chronological age, length of experience with Spanish (L2), and amount of L2 use in daily life. Importantly, they did not differ in age at onset of acquisition of Spanish. The learners in the two groups began acquiring Spanish in late adolescence (range = 11–18). None of the learners in either of the two groups claimed to have developed a functional command of Spanish (a command allowing them to carry on a meaningful conversation in the language) before 19 years of age. The subjects in the two learner groups had studied (or were in the process of studying) Spanish in a formal college setting. All the speakers were recruited from the University of Arizona community and were recorded in Tucson, Arizona. Their chronological age ranged from 20 to 38. All the participants were female because it is well known that sex affects acoustics—females and males differ in the size of the vocal tracts, which in turn affects their voice. In order not to add a further confounding factor, we limited our research to one gender.

The native Spanish speakers (controls) were born and raised in various Spanish-speaking countries (two in Spain, one in Peru, and two in Mexico)—dialectal differences were not expected for the vowels. They were brought up by Spanish monolingual families and grew up in a monolingual Spanish environment although they had studied English in school during adolescence. At the time of the recordings they all resided and worked in Tucson, Arizona. None of the control participants had moved to the United States, or to any other English-speaking environment, prior to age 20. When they were recorded for the present experiment, they had been living in Arizona between 3 and 8 years and had developed an advanced command of English. They used Spanish and English daily in their personal and/or professional life. Although they are bilingual Spanish-English speakers, they are considered native speakers of Spanish for the purposes of the present study. Here, they are referred to as controls or natives.

The advanced second-language learners of Spanish were born and raised in the United States. They were brought up by English-speaking monolingual families and were functionally monolingual until age 19 or later. Their age of first exposure to Spanish ranged between 11 and 18 years of age. Their first exposure to Spanish took place in a formal setting, such as a high school or college classroom. All five had studied abroad in a Spanish-speaking country. At the time of the recordings they were all pursuing a graduate degree in Spanish at the University of Arizona and thus used Spanish and English daily in both their personal and professional lives.

The intermediate second-language Spanish learners were also born in the United States to English-speaking monolingual families. They were brought up in a monolingual environment and were exposed to Spanish for the first time in a formal setting during adolescence. None claimed to have developed a functional command of Spanish before age 19. At the time of the recordings they were taking intermediate (third-year sequence) Spanish language courses at the University of Arizona and used Spanish several days a week but typically only in an academic setting.

Materials

The materials consisted of words that contained the five Spanish vowels in both stressed and unstressed positions. The chosen words are included in Table 1. The chosen words were controlled for the following three phonological characteristics: 1) the timbre of the target vowel (five values = /i e a o u/); 2) the lexical stress configuration of the word (two values = target vowel in stressed syllable vs. target vowel in syllable preceding lexically stressed syllable, such as the pre-stressed syllable); 3) the place of articulation of the consonant preceding the target vowel. While factors 1) and 2) were systematically manipulated, factor 3) was simply controlled for and not manipulated. In other words, the consonant preceding the target vowel was always a coronal consonant: /t d r l n s/. Variation in the vowels triggered by the nature of the surrounding consonants was not investigated. It is to be expected that coronal consonants will have significant co-articulatory effects on the target vowels due to their gestural requirements. However, since all target vowels were surrounded by coronal consonants in all cases, this effect is expected to be reasonably similar across target vowels. Each one of the five phonemic Spanish vowels was represented by five Spanish lexical items in each of the two stress configurations. Fifteen distractors were added to the materials.

Table 1. Spanish materials (target vowels appear in bold)

Stressed vowels	Unstressed vowels
/a/	/a/
habl**a**do	t**a**rifa
tir**a**no	s**a**rape
tir**a**da	r**a**cista
sen**a**do	n**a**cido
pas**a**do	l**a**nero
/e/	/e/
mod**e**lo	r**e**ceta
esc**e**na	t**e**resa
mor**e**na	s**e**dante
mad**e**ra	r**e**cibo
mon**e**da	d**e**dico
/i/	/i/
met**i**do	d**i**nero
ment**i**ra	t**i**rada
cas**i**ta	l**i**naje
vis**i**ta	d**i**recta
lad**i**no	n**i**dito
/o/	/o/
sobret**o**do	t**o**nada
par**o**dia	s**o**lapa
verd**o**so	r**o**dilla
perez**o**so	l**o**neta
ad**o**ra	d**o**lida
/u/	/u/
fract**u**ra	t**u**rista
ast**u**ta	s**u**reño
sut**u**ra	r**u**lero
rot**u**ra	d**u**dado
quemad**u**ra	l**u**neta

Recordings

The participants recorded each target item three times: 5 vowels × 5 words per vowel type × 3 repetitions × 2 lexical stress configurations = 2,250 data observations or tokens (plus 15 distractors × 3 repetitions, which were discarded). For this study we decided to use controlled, laboratory speech; that is, speakers were recorded while producing a set of materials that was carefully chosen and was provided to them. This type of speech, although arguably less natural than spontaneous speech-in-interaction, such as that present in a meaningful conversation, allows for the control of the materials so that balanced data may be collected. The benefit of laboratory speech for an exploratory investigation is that it allows for the collection of balanced numbers of tokens in controlled phonological environments, which, in turn, permits a reliable quantitative examination of the relevant factors.

Although most phonetic studies use read-aloud speech, we decided to use an experimental procedure that did not require reading, and it thus potentially reduced effects caused by differences in reading fluency amongst the participants. These differences are not negligible in our study since native speakers and advanced and intermediate learners are compared. The procedure we chose is called 'delayed repetition' and has been widely used in studies of L2 speech learning (Guion 2003; among many others). Its most important advantage is that it allows for the collection of controlled materials while eliminating the need to ask subjects to read words or sentences, which introduces particular types of errors, especially in non-fluent readers. In our procedure the speakers listened to an auditory model of each of the target words in a carrier phrase *María dice* . . . , then listened to the question, *¿Qué dice María?* and, finally, they repeated the target word in its carrier phrase *María dice*. . . . The sentences varied not only in the target words but also in the proper names used in the carrier phrases: *María*, *Juan*, *Alberto*, *Luisa*, etc. In this way speakers are asked to retain in memory the target word plus a distracter. Notice that between the time in which speakers hear the target word and the time in which they produce it, a few seconds have elapsed: the time for the question *¿Qué dice María?* (that they hear in the auditory model) and the carrier phrase *María dice* . . . (that they produce). The auditory model listened to by the speakers is hypothesized to activate the mental lexicon of the participants. The delay in the repetition of the target word is hypothesized to make the productions of the participants depend mostly (or exclusively) on their long-term representation of the target word (their mental lexicon or phonological representation) or at least the target phonemes. In this way, although the task provides an auditory model, it is believed to require more than mere imitation or acoustic shadowing.

As mentioned above, the participants produced three iterations of each target word. These were produced in three different blocks as follows. In block 1, the participants heard the auditory models and then repeated the sentences as explained above. Additionally, the participants were able to read the target words on a computer screen simultaneously if they wanted to do so. Blocks 2 and 3 differed from block 1 only in that no visual model of the target words was simultaneously presented to the speakers; thus, the participants depended exclusively on the auditory models. Blocks 2 and 3 were, therefore, not different in nature except for the fact that each of the three blocks contained a different randomization of the word list. Each speaker heard different randomizations in each of the three blocks. The auditory models were recorded by a native speaker of Spanish, born and raised in Spain, who read the sentences from a printed list.

The study was run in a sound-attenuated booth in the Applied Phonetics & Psycholinguistics Laboratory of the University of Arizona. Speakers were run individually. A dynamic head-worn microphone was used to record speech (Shure SM10A). Speech productions were digitized using a solid-state Marantz digital recorder (44.1 kHz, 16-bit). The presentation of the auditory (and visual) models was controlled via a desktop computer running Praat (Boersma 2001) and presented to the participants via a monitor and professional-quality headphones (AKG 77). The participants controlled the pace of the presentation of the auditory models with a computer keyboard.

Acoustic Analyses

As it is standard in acoustic studies of vowel timbre, first (F1) and second formant (F2) values were extracted from each vowel token. The first formant non-linearly indexes vowel height while the second formant indexes vowel fronting or backing, also non-linearly. Since humans perceive acoustic frequency logarithmically, we followed standard procedures in vowel acoustics and log-transformed the raw frequency values (Chládková et al. 2011; Escudero et al.

2009, among others). We used Praat's built-in function to transform Hz values into Bark units, a logarithmic scale. (Hereafter we refer to Bark-transformed F1 and F2 simply as F1 and F2.)

Approximate vowel steady states were located by hand and marked by clicking with the cursor on the corresponding spectrogram. Segmentation (location of midpoints) was carried out by carefully examining synchronized sound waves and spectrograms using the software package Praat. Vowel midpoints were identified by finding the region in the middle of the vowel where less formant movement was observable (where it is hypothesized that the speaker has reached the intended articulatory target for the specific vowel token) or, in case no stability was observable, the peak F1 value in the approximate middle of the vowel region.

Formant values were initially extracted in Hz using a Gaussian-like window with a length of 25 milliseconds centered on the segmented time point. We used the standard procedure in Praat (the Burg algorithm: *To Formant (burg) . . .*) to automatically detect formant values with the maximum number of formants set to five. A procedure was developed to minimize computer tracking errors. The procedure is a simplified adaptation of the method used in Chládková et al. (2011) and in Escudero et al. (2009): a Praat script extracted F1 and F2 twenty times from each vowel token, with twenty different equidistant formant ceilings between 4,500 Hz and 6,400 Hz. Subsequently, a script calculated the standard deviations for each vowel category, each ceiling parameter, and each speaker separately. Optimal formant values for each vowel category and each speaker were those that had the lowest standard deviation as compared to the same vowel tokens when their formants had been extracted with other ceiling parameters. This is based on the assumption that higher standard deviations when using different ceiling parameters are caused by tracking errors or ill-detected outliers. Escudero et al. (2009) show that this procedure (or a similar one) effectively reduces noise caused by tracking errors in the formant-detection algorithm, which are not uncommon in automatic formant tracking.

The theoretical dataset consisted of 4,500 values—F1 and F2 of 2,250 vowel tokens. A total of 35 tokens were not entered as data due to recording errors of various sorts, including creak and devoicing. Thus, the final dataset consisted of 2,215 vowel tokens. By-subject ANOVAs were chosen as the basis for all statistical explorations. By-subject, by-condition medians were selected (over iterations and items) so that each speaker provided at most 20 data observations (10 F1 and 10 F2). This resulted in 300 observations, each a median of several productions. This prepares the data for mixed-model ANOVAs, and it eliminates outliers.

Results

Figures 1, 2, and 3 plot the mean log-transformed formant values plus one-standard-deviation (1-SD) ellipses for the native controls, the advanced, and the intermediate learners, respectively. Stressed and unstressed vowels are plotted for each group. Stressed vowels are marked with an acute accent mark and unstressed vowels are not marked in any special manner.

The first step was to submit the median formant values to two separate mixed-model ANOVAs of medians with 'vowel' and 'stress' as within-subject factors and 'speaker group' as a between-subjects factor, one analysis per formant. These models are necessary to justify further explorations of the dataset in subsets. The ANOVA in which F1 values were explored revealed significant effects of 'vowel,' $F\,(4,47) = 459.44$, $p < 0.001$, and a significant 'vowel' by 'stress' interaction $F\,(4,48) = 71.44$, $p < 0.001$. The effects of other factors or interactions were found to be negligible. Regarding F2 values, the corresponding ANOVA yielded significant effects of 'vowel,' $F\,(4,47) = 1025.47$, $p < 0.001$, and 'stress,' $F\,(1,11) = 61.52$, $p < 0.001$, as well as significant 'vowel' by 'stress,' $F\,(4,48) = 47.86$, $p < 0.001$) and 'vowel' by 'group,' $F\,(8,47) = 5.84$, $p < 0.001$ interactions. A three-way interaction was not found.

These statistical tests on median formant values suggest that stressed and unstressed vowels differ acoustically, and that vowel category and speaker group modulate this difference. Also,

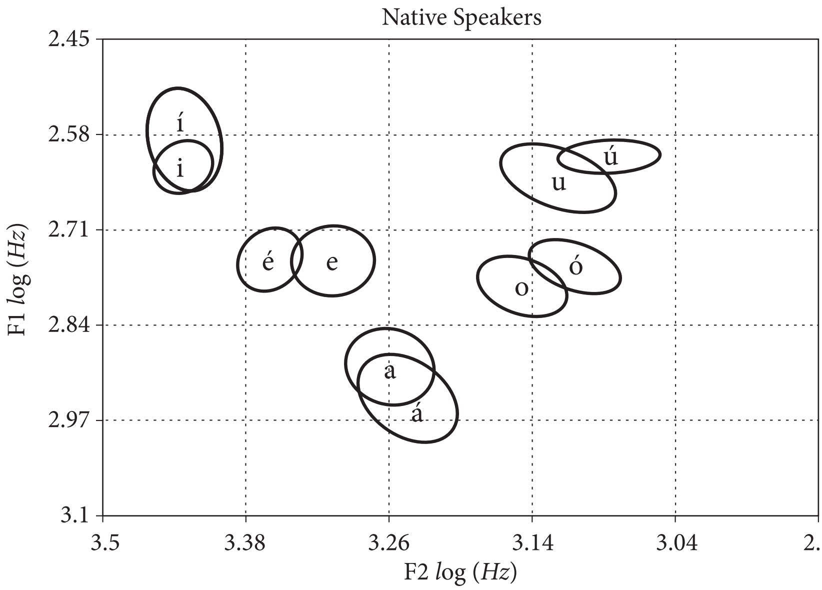

Figure 1. Spanish F1 × F2 stressed and unstressed vowel space produced by the native Spanish controls

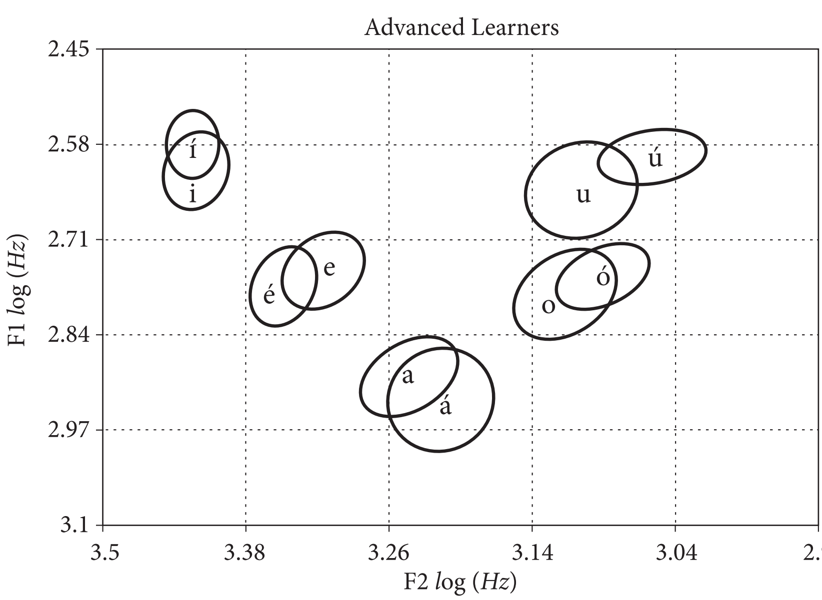

Figure 2. Spanish F1 × F2 stressed and unstressed vowel space produced by the advanced learners

the acoustics of vowels differ across groups with respect to fronting but not height, and the effects of stress are dependent on vowel category in fronting. The next step was to investigate the possible effects of stress and speaker group, as well as their potential interaction, for the five Spanish vowels separately.

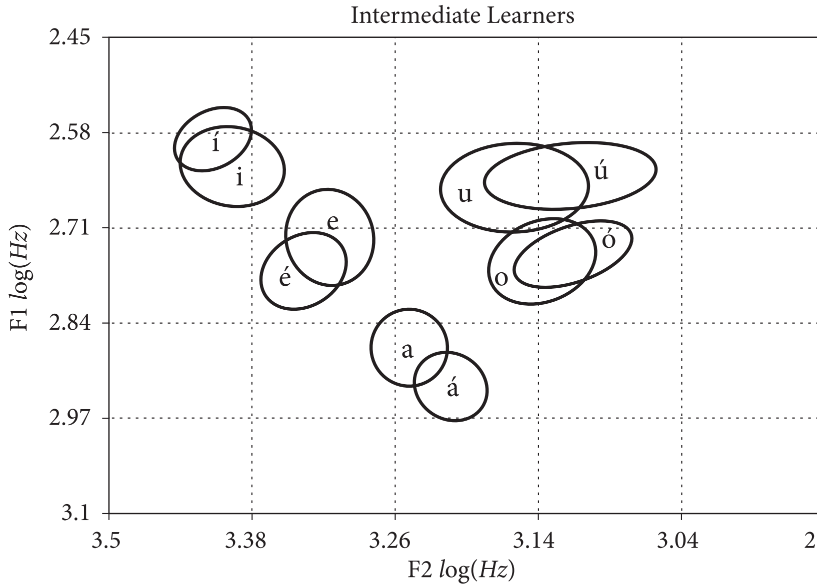

Figure 3. Spanish F1 × F2 stressed and unstressed vowel space produced by the intermediate learners

Vowel /i/

The formant medians corresponding to /i/ tokens, stressed and unstressed, by the three subject groups were submitted to two separate ANOVAs, one per formant, with 'stress' as the within-subjects factor and 'speaker group' as the between-subjects factor. The ANOVA that analyzed F1 values revealed significant 'stress' effects: F (1,12) = 42.52, $p < 0.001$. No main effects of 'speaker group' were found. We did not find an interaction between 'stress' and 'group.' This means that stressed /i/ tokens have a significantly lower F1 if stressed than if unstressed; that is, stressed /i/ vowels are higher in the vowel space than unstressed /i/ tokens. This effect is consistent across speaker groups, including the native controls.

Regarding F2, no main effects and no interactions were found. In fact, 'stress' approached, but did not reach, significance, F (1,12) = 3.65, $p = 0.08$. In other words, stressed and unstressed /i/ vowels differ mostly in their height (higher in the vowel space when stressed than when unstressed). The degree of fronting is similar across stressed and unstressed /i/ vowels. No differences between the three groups were found with respect to how they produced /i/ or how they implemented unstressed vowel reduction for /i/, which consists of lowering unstressed /i/ with respect to stressed /i/.

Vowel /e/

Identical statistical tests were used to examine the acoustics of /e/. F1 values differed as a function of 'stress,' F (1,12) = 10.979, $p < 0.01$, but not 'speaker group.' A significant interaction between 'stress' and 'group' was found, F (2,12) = 6.27, $p = 0.01$. Since no main effects of 'group' were found, we conclude that /e/ tokens did not differ significantly across speaker groups and, therefore, learners and native controls, overall, use /e/ vowels with similar formant frequencies. On the other hand, since a two-way interaction was found, it is reasonable to believe that the main effects of stress differ as a function of speaker group; that is, the speaker groups implement the difference between stressed and unstressed vowels differently. Planned comparisons explored the interaction. The interaction was evidently due to the fact that, while neither the native Spanish speakers, F (1,4) = 0.56, $p > 0.1$, nor the advanced L2 learners, F (1,4) = 3.72, $p > 0.1$, produced a difference in height (F1) between stressed and unstressed /e/, the intermediate learners, as a group, did realize a difference between the two /e/ types, F (1,4) = 17.01, $p = 0.01$. In particular, stressed /e/ was higher in the vowel space (had a lower F1) than unstressed /e/ for the intermediate learners.

We then investigated variation in F2. An ANOVA found main effects of 'stress,' F (1,12) = 112.67, $p < 0.001$, but not of 'speaker group.' Importantly, a significant two-way interaction was found, F (2,12) = 9.49, $p < 0.01$. Overall, stressed and unstressed /e/ tokens differ from each other, but this effect is modulated by speaker group. The interaction was arguably due to the fact that most speakers produced a difference in fronting between stressed and unstressed /e/, but this difference varied in magnitude across groups. In particular, both native controls, F (1,4) = 202.52, $p < 0.001$, and advanced learners, F (1,4) = 28.98, $p < 0.01$, produced a significant difference in fronting between the two /e/ types. On the other hand, stress effects for the intermediate learners were marginally not significant, F (1,4) = 7.32, $p = 0.05$. While, on average, the mean difference between stressed and unstressed /e/ in fronting was in the order of 0.82 Bark for the native speakers, the advanced learners' mean difference was 0.52 Bark. The mean difference for the intermediate learners was of 0.28 Bark. In other words, stressed /e/ vowels were significantly more fronted than unstressed /e/ tokens, but this difference was greater for the native Spanish speakers, smaller for the advanced learners and smallest for the intermediate learners. This is an important, unexpected finding.

In sum, all three groups of speakers implement stressed and unstressed /e/ tokens differently. While native Spanish speakers seem to simply front their stressed /e/ (with respect to their own unstressed /e/) and do not produce a difference in height between the two vowels, intermediate learners of Spanish lower their unstressed /e/ (with respect to their own stressed /e/) and also centralize it slightly (though significantly less than native speakers). Finally, advanced learners of Spanish stand somewhere in the middle of these two patterns: although they do centralize (horizontally) their unstressed /e/ (with respect to their own stressed /e/), this difference is smaller than for the natives and larger than for the intermediate learners; on the other hand, they pattern with native speakers in not lowering unstressed /e/ (relative to their own /e/).

Vowel /a/

The next step was to analyze stress effects for /a/. Regarding F1, an ANOVA yielded significant effects of 'stress,' F (1,12) = 100.59, $p < 0.001$, but no 'speaker group' effects and no significant interaction. Stress effects were due to the fact that unstressed /a/ is significantly higher in the vowel space (has a lower F1) than stressed /a/. Apparently, this pattern is robust across speaker groups.

The analysis of F2 values yielded identical results: significant effects of 'stress,' F (1,4) = 48.11, $p < 0.001$, but no effects of 'speaker group' and no interaction. Once again, the unstressed vowel reduction process seems to be similar across speaker groups. In particular, unstressed /a/ is significantly more fronted in the vowel space (has a higher F2) than stressed /a/. Taken together, the findings regarding /a/ suggest that unstressed vowel reduction of the Spanish lower vowels consists of slightly but significantly raising and fronting unstressed /a/ relative to stressed /a/. Interestingly, this is a pattern all three groups show in their speech.

Vowel /o/

Regarding Spanish /o/, the corresponding test revealed that 'stress' was a significant predictor of F1, F (1,12) = 20.69, $p < 0.001$, while 'speaker group' was not. There was no significant interaction between the two factors, but it is perhaps possible to point out a trend: $F\,(2{,}12) = 3.22$, $p = 0.07$. The analysis of F2 yielded very similar results: significant effects of 'stress,' $F\,(1{,}12) = 58.36$, $p < 0.001$, but not of 'speaker group.' In the case of F2 there was no interaction between the two main factors, not even a trend.

In sum, the tests suggest that Spanish unstressed /o/ differs from stressed /o/ in that it is significantly lower in the vowel space (it has a higher F1) and it is significantly more fronted (it has a higher F2) or closer to the vowel centroid. This is a pattern that the three groups share.

Vowel /u/

The case of the vowel /u/ was found to be clearly different from the other Spanish vowels. First, an ANOVA examining F1 found significant 'stress' effects, $F\,(1{,}12) = 16.9$, $p < 0.01$, as well as a marginally non-significant effect of 'speaker group,' $F\,(2{,}12) = 3.41$, $p = 0.06$. Importantly, there was no interaction between the two predictors. The effects of stress on F1 were arguably due to the fact that unstressed /u/ is significantly lower in the vowel space (it has a higher F1) than stressed /u/. Since no interaction was revealed by the ANOVA, it appears that this pattern is robust across speaker groups.

On the other hand, the analysis of F2 values resulted in significant effects of 'stress,' $F\,(1{,}12) = 41.53$, $p < 0.001$, as well as 'speaker group,' $F\,(2{,}12) = 5.46$, $p = 0.02$. Once again, no interaction was revealed by the statistical test. Stress effects were caused by the following pattern: unstressed /u/ is robustly more fronted than stressed /u/. This is so for all three speaker groups, as suggested by the lack of an interaction. A by-items ANOVA in which stress was not considered a factor showed the effects of 'speaker group' to be robust: $F\,(2{,}18) = 38.22$, $p < 0.001$. These effects were arguably due to the fact that the intermediate learners produced a very fronted /u/, native speakers produced a back /u/ and, interestingly, advanced learners also produced a back /u/, one that was even further back than the one produced by the natives; in other words, /u/ vowels were ordered on the following progression, from back to front in the vowel space: advanced learners > native speakers > intermediate learners. A Tukey HSD test on the by-items ANOVA revealed that all pair-wise comparisons reached significance.

Discussion

It appears that vowels /i/, /a/, and /o/ do not cause any major difficulties to adult second language learners of Spanish whose native language is English. The case of /e/ is special: when stressed and unstressed vowels are pooled, no group differences are found. When the effects of stress are taken into account, however, differences between the learners and the native speakers are revealed. Finally, when stressed and unstressed vowels are taken together, only vowel /u/ differs as a function of speaker group. Specifically, /u/ is more fronted when produced by intermediate learners of Spanish than when produced by a control group of native speakers. Interestingly, advanced learners produce an /u/ vowel located very far back in the vowel space. This suggests that the initial state of English learners of Spanish has a fronted /u/, most likely a transfer effect from English.

The horizontal position of English /u/ is subject to enormous dialectal variation. While many American English speakers possess a back /u/, /u/-fronting has been reported in a number of locations in North America (Labov et al. 2006). Most relevant for the present study are reports of /u/-fronting in the West, where the data in the present study were collected. Hall-Lew (2011) explains that /u/-fronting is a recent sound change that is nearing completion in California. The data in Díaz (2010) show that /u/-fronting exists in Arizona and that this pattern has consequences for the production of Portuguese vowels by native speakers of English from this state who are in the process of learning Portuguese as a second language. Although we do not have acoustic data for the English vowels of our participants, it seems reasonable to believe, given their age and origin, that extreme /u/-fronting in their Spanish is a transfer feature from their English. The differences between intermediate and advanced learners are, most likely, proof that this transfer feature is overcome with sufficient experience with Spanish.

These findings come to support Flege's (1995) Speech Learning Model or SLM. According to the SLM, the larger the difference between an L2 sound and its equivalent L1 sound the easier it will be for learners to develop a new category for the L2 sound—one separate from the closest L1 sound. This is apparently what is occurring with /u/ in the present study. The large difference in the position of /u/ between the native controls and the intermediate learners suggests that the initial stage of acquisition of this Spanish vowel consists of full transfer (or robust interlingual assimilation) from English fronted /u/. The fact that advanced learners produce a much less fronted /u/ suggests that, with further experience, learners succeed in producing a backer Spanish /u/. This result might be caused by the fact that the acoustic-perceptual difference between English fronted /u/ and Spanish /u/ is large.

The specific finding regarding /e/ was that, when stressed and unstressed vowels were pooled, no differences arose among the three groups. This was obviously due to a large range of variation for this vowel. When stressed and unstressed vowels were considered separately, however, differences among the three groups were evident. This was not true for the other vowels. The difference between stressed and unstressed /e/ is one of fronting (centralization) in the case of native speakers—unstressed vowels are more centralized—but one of height (lowering) in the case of the intermediate learners of Spanish—unstressed vowels are raised. The advanced learners do not produce a difference in height (like native speakers and unlike intermediate learners) but they produce a difference in fronting (like native speakers and unlike intermediate learners) between the two types of /e/. However, the magnitude of the difference between stressed and unstressed /e/ is smaller for the advanced learners than it is for the native speaker controls. In other words, the advanced learners seemed to be close to native like in their acquisition of the acoustics of Spanish /e/.

The case of /e/ might also be explained with reference to Flege's SLM. No English vowel is identical in its acoustic envelope to Spanish /e/. English learners might assimilate Spanish /e/ to either English /ɛ/ or /eɪ/. Although both vowels are different, they are both acoustically more similar to Spanish /e/ than English fronted /u/ is to Spanish /u/. Spanish /e/ is perhaps slightly higher and more fronted than English /ɛ/, but they are both monophthongs. While English /eɪ/ is as high and fronted as Spanish /e/, it is slightly diphthongal. Perhaps the similarities between Spanish and English mid front vowels make it difficult, even for experienced learners, to overcome equivalence classification. Alternatively, Spanish /e/ may be assimilated to both English /ɛ/ or /eɪ/ in different lexical items, thus blocking category formation for these learners.

One important finding of the present study is that native speakers of Spanish realize systematic differences between their stressed and unstressed vowels: unstressed /i/ was lower than stressed /i/, unstressed /a/ was higher and more fronted than stressed /a/, unstressed /o/ was more fronted (centralized) than stressed /o/, unstressed /e/ was realized further back (centralized), unstressed /u/ was more fronted than stressed /u/. This is perhaps surprising. Nevertheless, the few studies that have examined unstressed vowel reduction in Spanish have led to inconclusive results (Menke and Face 2010; Nadeu 2012; Ortega-Llebaría and Prieto 2010; Torreira and Ernestus 2011; among others). While some find that stressed and unstressed vowels do not differ in timbre in Spanish, or that they barely do (Menke and Face 2010; Ortega-Llebaría and Prieto 2010), others find differences between vowels in these prosodic positions (Nadeu 2012; Torreira and Ernestus 2011). The studies that do not find differences between these vowels have in common that they consist of read aloud speech data. Nadeu (2012), after asking her speakers to read the materials, asked them to pronounce them as fast as possible in repetition within a short time slot. Torreira and Ernestus (2011), on the other hand, analyze vowels extracted from conversational, casual interactions between close friends and find robust stress effects in vowel height. It seems reasonable to propose that formal tasks affect vowel reduction in Spanish, being that in this language this is a phonetic, gradient process rather than a categorical phonological pattern. We would like to surmise that the delayed repetition task that we used in the present study triggers a speech style that does not eliminate the application of a phonetic process of unstressed vowel reduction, as reading aloud seems to do.

Conclusion

The present study has analyzed the acoustic characteristics of the five Spanish vowels /i e a o u/ as produced by three groups of speakers: a group of native speakers of Spanish (controls), a group of intermediate and a group of advanced second language learners of Spanish whose first language is English. The speakers pronounced lexical items that contained the five vowels in either stressed or unstressed position. This allowed us to explore the acoustic differences between stressed and unstressed vowels for the three groups.

We have found that English learners differ from native speakers mostly in their production of /u/, which is more fronted for the intermediate learners than for the advanced learners and for the native controls. It was hypothesized that this was due to interlingual assimilation of Spanish /u/ with their native English /u/, very fronted in the dialect of these subjects. It was suggested that the large acoustic difference between English fronted /u/ and Spanish back /u/ is the cause for the large difference between the intermediate learners and the other two groups, and also responsible for the fact that advanced learners show obvious signs of the development of a back /u/ for Spanish, a clear sign of phonetic learning.

We also found that Spanish /e/ posed an exceptional challenge for these learners. Differences between /e/ vowels produced by learners and native speakers arose when investigating the effects of stress on this vowel for the different groups of speakers. It was hypothesized that the strong similarities (and slight differences) between Spanish /e/ and English /eɪ/ were responsible for the developmental challenge posed by this specific vowel.

Finally, our data presented evidence for the existence of vowel-specific patterns of phonetic, gradient unstressed vowel reduction even for our group of native speaker controls. While Spanish tends to be discussed amongst the languages that do not display vowel reduction in their phonology, it was shown how prior studies had already pointed out that vowel reduction exists in the phonetics of this language.

WORKS CITED

Barnes, Jonathan. (2006). *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. Berlin: Gruyter. Print.

Best, Catherine. (1995). "A Direct Realist View of Cross-language Speech Perception." *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Ed. Winifred Strange. 171–206. Timonium: York. Print.

Best, Catherine, and Michael Tyler. (2007). "Nonnative and Second-language Speech Perception: Commonalities and Complementarities." *Language Experience in Second Language Speech Learning: In Honor of James E. Flege*. Ed. Ocke-Schwen Bohn and Murray Munro. 13–34. Amsterdam: Benjamins. Print.

Boersma, Paul. (2001)."Praat, a System for Doing Phonetics by Computer." *Glot International* 5:341–45. Print.

Bybee, Joan. (2001). *Phonology and Language Use*. Cambridge: Cambridge UP. Print.

Casillas, Joseph. (2012). "Percepción y procesamiento de contrastes vocálicos en hablantes bilingües de español e inglés." Unpublished MA Thesis, Arizona. Print.

Chládková, Kateřina, Paola Escudero, and Paul Boersma. (2011). "Context-specific Acoustic Differences Between Peruvian and Iberian Spanish Vowels." *Journal of the Acoustical Society of America*. 130: 416–28. Print.

Díaz, Miriam. (2010). *L2 and L3 Acquisition of the Portuguese Stressed Vowel Inventory*. Unpublished PhD Dissertation. Arizona. Print.

Escudero, Paola, Paul Boersma, Andreia Rauber, and Ricardo Bion. (2009). "A Cross-dialect Acoustic Description of Vowels: Brazilian and European Portuguese." *Journal of the Acoustical Society of America* 126: 1379–93. Print.

Flege, James E. (1995). "Second Language Speech Learning: Theory, Findings and Problems." *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Ed. Winifred Strange. 233–77. Timonium: York. Print.

Flege, James, and Ocke-Schwen Bohn. (1989). "An Instrumental Study of Vowel Reduction and Stress Placement in Spanish-accented English." *Studies in Second Language Acquisition* 11: 35–62. Print.

Fourakis, Marios. (1991). "Tempo, Stress and Vowel Reduction in American English." *Journal of the Acoustical Society of America* 90: 1816–27. Print.

Giegerich, Heinz J. (1992). *English Phonology: An Introduction*. Cambridge: Cambridge UP. Print.

Guion, Susan. (2003). "The Vowel Systems of Quichua-Spanish Bilinguals: An Investigation into Age of Acquisition Effects on the Mutual Influences of the First and Second Languages." *Phonetica* 60: 98–128. Print.

Hall-Lew, Lauren. (2011). "The Completion of a Sound change in California English." *Proceedings of the 17th International Congress of Phonetic Sciences,* 807–10. Adelaide: Causal. Print.

Hualde, José I. (2005). *The Sounds of Spanish*. Cambridge: Cambridge UP. Print.

Labov, William, Sharon Ash, and Charles Boberg. (2006). *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin: Gruyter. Print.

Lindblom, Bjorn. (1963). "Spectrographic Study of Vowel Reduction." *Journal of the Acoustical Society of America* 35: 1773–81. Print.

Menke, Mandy, and Timothy Face. (2010). "Second Language Spanish Vowel Production: An Acoustic Analysis." *Studies in Hispanic and Lusophone Linguistics* 3: 181–214. Print.

Morrison, Geoffrey. (2003). "Perception and Production of Spanish Vowels by English Speakers." Ed. Maria J. Solé, Daniel Recasens, and Joaquín Romero. *Proceedings of the 15th International Congress of Phonetic Sciences.* 1533–36. Adelaide: Causal. Print.

———. (2008). "L1-Spanish Speakers' Acquisition of the English /i/-/ɪ/ Contrast: Duration-based Perception is not the Initial Developmental State." *Language and Speech* 51: 285–315. Print.

———. (2009). "L1-Spanish Speakers' Acquisition of the English /i/-/ɪ/ Contrast: Perception of Vowel Inherent Spectral Change." *Language and Speech*. 52: 437–62. Print.

Nadeu, Marianna. (2012). "Effects of Stress and Speech Rate on Vowel Quality in Catalan and Spanish." *Proceedings of the 13th Conference of the International Speech Communication Association (Inter-Speech)*. Print.

Ortega-Llebaría, Marta, and Pilar Prieto. (2010). "Acoustic Correlates of Stress in Central Catalan and Castilian Spanish." *Language and Speech* 54: 73–97. Print.

Simonet, Miquel. (2012). "The L2 Acquisition of Spanish Phonetics and Phonology." Ed. José I. Hualde, Antxon Olarrea, and Erin O'Rourke. *The Handbook of Hispanic Linguistics*. 729–46. Oxford: Wiley-Blackwell. Print.

Stockwell, Robert P., and J. Donald Bowen. (1965). *The Sounds of English and Spanish*. Chicago: U of Chicago P. Print.

Torreira, Francisco, and Mirjam Ernestus. (2011). "Realization of Voiceless Stops and Vowels in Conversational French and Spanish." *Journal of Laboratory Phonology* 2: 331–53. Print.

Zampini, Mary, and Kerry Green. (2001). "The Voicing Contrast in English and Spanish: The Relationship between Perception and Production." *One Mind, Two Languages: Bilingual Language Processing.* Ed. Janet Nicol. 23–48. London: Wiley-Blackwell. Print.