The role of phonological short-term memory in second language phonology: Exploring vowel quality among English-speaking learners of Spanish

Abstract

This study reports on the production of the five Spanish vowels in stressed and unstressed contexts by English-speaking advanced learners of Spanish at two experience levels who differed according to phonological short-term memory (PSTM) capacity. Our findings indicate that learners with higher PSTM capacity were less likely to reduce both stressed and unstressed vowels in Spanish. Consequently, learners with higher PSTM in fourth-year Spanish courses produced vowels that were more native-like than lower PSTM fourth-year learners. On the other hand, higher PSTM graduate students of Spanish produced a Spanish vowel space that was more expanded than that of native Spanish, particularly for /u/, compared to lower PSTM graduate learners, who were more native-like. These findings suggest that PSTM does play a role in second language phonology, even at advanced levels, and may help to explain why individual differences in pronunciation exist even at very high levels of second language proficiency.

Keywords: *Second language acquisition, phonetics/phonology, individual differences, phonological short-term memory, vowels*

1. Introduction

The Spanish and English vowel systems differ in several ways, particularly regarding height, positioning on the horizontal plane, and the effect of lexical stress on reduction (e.g., Bradlow, 1995; Ladefoged, 2006). Consequently, it is unsurprising that research into the second language (L2) acquisition of Spanish vowels by native speakers of English has found that even advanced L2 speakers often produce Spanish vowels with formant values that differ from those of native speakers (e.g., Cobb & Simonet, 2015; Long, Solon & Bongiovanni, 2018; Menke & Face, 2010; Moorman, 2017). At the same time, recent research on the relationship between phonological short-term memory (PSTM) and L2 pronunciation has found a largely facilitative role for PSTM: learners with higher PSTM are more likely to be rated as having more accurate L2 pronunciation, at least in beginning and intermediate levels (Kondo, 2012; Nagle, 2013; Reiterer et al., 2011; Slevc & Miyake, 2006). Although this relationship dissipates at higher levels of proficiency (Hu et al., 2013, Reiterer et al., 2011; Venkatagiri & Levis, 2007), advanced learners with higher PSTM have been found to more accurately produce L2 vowels in French and English (Inceoglu, 2019; Mora & Darcy, 2017) when assessed via acoustic correlates. Additionally, one study found a relationship between PSTM and L2 Spanish production in beginning and intermediate levels (Moorman, 2017). Thus, it stands to reason that although PSTM abilities may not correlate with global ratings of L2 pronunciation in advanced learners, it may facilitate the pronunciation of specific L2 sounds. With that in mind, this study analyzes the production of the five Spanish vowels in stressed and unstressed contexts by fourth-year Spanish majors and minors and graduate students of Spanish who differed according to PSTM. The goal is to examine whether advanced learners of L2 Spanish with higher PSTM abilities more accurately produced these sounds than their cohorts with lower PSTM.

2. Background

2.1 Spanish and English Vowel Systems

Advanced English-speaking L2 learners of Spanish have difficulty achieving target-like production of Spanish vowels, due to the differences in vowel height (F1) and vowel frontness/backness (F2) between the two languages. English has from eleven to fourteen contrastive vowels, depending on dialect (Bradlow, 1995; Ladefoged, 2006; Stockwell & Bowen, 1965). Spanish, on the other hand, has five contrastive monophthongal vowels, /a e i o u/ (Hualde, 2005) and is generally believed to have less dialectal variation (see Willis [2005] for an exception).

Bradlow (1995) conducted a comparative study of the four vowel phonemes that General American English and Madrid Spanish have in common, /e i o u/, and found that Spanish native speakers generally produced these phonemes in a more anterior position in the vowel space relative to native English vowels. In terms of vowel height, /i/ and /e/ tend to be produced higher in the vowel space in English than in Spanish. At the same time, /o/ is often lower in English than in Spanish. Regarding /a/, Spanish has only one low central vowel, while General American English has two, a low more fronted /æ/, as in *matter*, and a low more backed /ɑ/, as in *father* (Solon, Long, & Gurzynski-Weiss, 2017).

Additionally, English has a process of unstressed vowel reduction via centralization and shortened duration, and frequently deletion, while Spanish vowels are largely considered to be stable. Delattre (1969) compared vowel reduction in English, French, German and Spanish, and found English to have the most vowel reduction in unstressed syllables, followed by French and German;. Spanish was found to have the least. Recent research has found that Spanish does exhibit some degree of centralization (Cobb & Simonet, 2015; Menke & Face, 2010), but still considerably less than that of English-speaking L2 Spanish learners. This conclusion is supported by research findings that Spanish learners of English evidence smaller acoustic differences between stressed and unstressed vowels in English than English native speakers do (Flege & Bohn, 1989), likely due to a transfer of vowel reduction patterns from native Spanish.

2.2. L2 Acquisition of Spanish Vowels

Research on the L2 acquisition of Spanish vowels by English-speaking learners has generally found that as L2 learners' experience with Spanish increases, there is a gradual progression towards more native-like formant structures. Nonetheless, phonetic differences between learner and native production persist even in very advanced L2 speakers (Cobb & Simonet, 2015; Long, Solon, & Bongiovanni, 2018; Menke & Face, 2010; Moorman, 2017).

Menke and Face (2010) conducted a cross-sectional analysis of vowel production among university students at different level, plus a control group of native speakers, using a read aloud task containing 20 tokens per vowel, divided evenly between stressed and unstressed syllables. The results of their acoustic analysis revealed that the intermediate-level learner group pronounced vowels in a reduced vowel space compared to more advanced learners and native speakers. In particular, /u/ fronting was more prevalent the lower the learner proficiency level. On the other hand, the vowel spaces of graduating majors and PhD students were largely identical to that of native Spanish speakers.

Long, Solon and Bongiovanni's (2018) results corroborated these findings for intermediate Spanish university and study abroad learners. Learners that were enrolled in a third-year course both at-home and during a short-term study abroad produced L2 Spanish vowels in a reduced vowel space in comparison to native Spanish vowels as in previous research, although no statistical comparison to native speaker data was performed. In particular, intermediate learners evidenced a high degree of /u/ fronting. Similarly, Moorman (2017) found mean F1 and F2 Bark[1] values for English-speaking learners of L2 Spanish enrolled in second and sixth semester university Spanish courses to differ from those of native speakers. However, these findings were not statistically significant.

Regarding the effect of lexical stress, previous research indicates that English-speaking L2 Spanish learners centralize their vowels in unstressed syllables more than native speakers. Additionally, Menke and Face (2010) found that native speakers showed centralization in F2 values for /e/ and /u/. Unstressed /e/ was backed while unstressed /u/ was fronted relative to their stressed counterparts. Their learner groups demonstrated this same reduction for /e/ and /u, although the learners also showed stress effects on the other vowels as well.

Cobb and Simonet (2015) also examined the effect of lexical stress on Spanish vowel formants produced by native Spanish speakers, English-speaking intermediate and advanced L2 Spanish learners. The authors observed much more vowel reduction in native speaker speech than Menke and Face (2010), in addition to the vowel reduction previously attested for L2 learners. Specifically, they found that native speakers evidenced differences between stressed and unstressed contexts in F1 and/or F2 for all vowels. All learner groups demonstrated the same patterns of reduction, while advanced learners produced a larger difference than intermediate learners did. Overall, intermediate learners evidenced the most vowel reduction in unstressed syllables, with advanced learners reducing less, although still more than native speakers.

---

[1] Bark values are a normalized value for formant measurements created using the Bark Difference Method (Syrdal & Gopal, 1986).

Thus, it seems that L2 learners of Spanish progress toward native speaker formant values, but that differences persist at even highly advanced L2 levels. Given that some difficulties in L2 Spanish vowel production persist in advanced speakers, it appears that the English-Spanish vowel contrast is a particularly difficult feature for L2 learners to acquire. One possible explanation for the difficulties learners still experience even at advanced levels, as well as the somewhat inconsistent findings related to L2 production abilities, may be found in individual differences between learners. Therefore, in this study we explore the possibility that PSTM is related to learners' success in acquiring native-like vowel patterns. We hypothesize that vowel production and centralization may be a phonological structure for which PSTM provides an advantage even into high levels of L2 proficiency. This hypothesis is supported by research on PSTM and vowel production in L2 English, Spanish and French, as we discuss in the following section.

2.3 Phonological Short-Term Memory and L2 Pronunciation

PSTM corresponds to the phonological loop of working memory, and is responsible for storing verbal information over short periods of time (Baddeley, 1986). It is comprised of two parts: a storage system that holds phonological information for approximately two seconds, and a sub-vocal articulatory rehearsal that can preserve phonological information in storage beyond two seconds (Baddeley & Hitch, 1974). Research suggests that PSTM is implicated in the learning of unfamiliar phonological forms.

Correlations have been documented between PSTM and a range of L2 abilities, such as vocabulary knowledge (Masoura & Gathercole, 2005; Speciale, Ellis, & Bywater, 2004), collocations (Skrzypek & Singleton, 2013), listening comprehension (Kormos & Safar, 2008; Tsuchihira, 2008) and sound perception (e.g., Aliaga-Garcia, Mora, & Cerviño-Povedano, 2011; Cerviño-Povedano & Mora, 2011, 2015). Taken together, these findings indicate that learners with higher PSTM may be better able to parse phonological input, establish accurate long-term representations of L2 sounds and develop new phonetic categories (Nagle, 2013). These enhanced capabilities may be due to an improved capacity to sub-vocally rehearse L2 sounds during processing, which in turn augments the ability to notice differences between L1 and L2 sounds, allowing for more accurate representations of L2 sounds (Mora & Darcy, 2017). Thus, it stands to reason that learners with higher PSTM abilities may also produce the L2 in a more native-like manner due to these more accurate representations. These hypotheses are supported by findings that similar cerebral networks are engaged in speech, perception, and phonological working memory tasks (e.g., Acheson, Hamidi, Binder, & Postle, 2011; Hickok, Buchsbaum, Humpries, & Muftuler, 2003).

However, research investigating the relationship between PSTM and L2 pronunciation skills has yielded mixed results. Some studies have found a relationship between PSTM and aspects of L2 pronunciation (Kondo, 2012; Inceoglu, 2019; Moorman, 2017; Mora & Darcy, 2017; Nagle, 2013; Reiterer et al., 2011), while others have found none (Hu et al., 2013; Kissling, 2014; Reiterer et al., 2011; Slevc & Miyake, 2006; Venkatagiri & Levis, 2007). However, these studies differ drastically in their participant demographics and their methodologies, which may account for divergent findings. When comparing across studies, two factors in particular appear to contribute to these differing results: learner experience with the L2 and type of pronunciation assessment. Overall, studies that have examined novice or beginning learners of an L2 generally found a relationship between PSTM and L2 pronunciation, whether examined globally or via acoustic analysis of specific segments (Inceoglu, 2019; Kondo, 2012; Moorman, 2017; Nagle, 2013; Reiterer et al., 2011; Slevc & Miyake, 2006).

On the other hand, studies that assessed more advanced learners have reported more mixed results. Research examining global pronunciation has generally found that PSTM does not correlate with L2 pronunciation at advanced levels (Hu et al., 2013; Reiterer et al., 2011; Venkatagiri & Levis, 2007). Conversely, studies that examined advanced learners and specific sounds or classes of sounds have found divergent results depending on the structure. For vowels, research on L2 English and French has found a positive correlation between PSTM and vowel pronunciation. For example, Mora and Darcy (2017) found that among advanced native Spanish-speaking learners of L2 English, with an average of over 11.5 years of English study, PSTM correlated with more target-like duration difference between long /i/ and short /ɪ/ in English. Similarly, Inceoglu (2019), who explored English-speaking L2 learners of French at different levels, found a relationship between PSTM and nasalized vowel production. Lastly, although Moorman (2017) only examined beginner and intermediate learners of L2 Spanish, she found that PSTM negatively correlated with vowel production in beginning learners but positively correlated in intermediate learners. On the other hand, Kissling (2014) analyzed English-speaking learners of L2 Spanish in introductory to advanced courses and found no correlation between PSTM and production of rhotics, voiced stop lenition, or voice onset time for voiceless stops.

This body of research suggests that PSTM may have a larger impact on global L2 pronunciation in early-stage learners, but not more advanced learners. However, while PSTM may not have a global effect on L2 pronunciation in more advanced levels, it might still have a role in the pronunciation of particular L1-L2 contrasts or classes of sounds. It may be that acquiring the acoustics of L2 vowel production is particularly challenging and as such higher PSTM could aid in the pronunciation those challenging contrasts, as the studies just reviewed have suggested. Additionally, research has found correlations between higher PSTM and better perception of L2 vowels, even in advanced learners, further supporting this hypothesis (e.g., Aliaga-Garcia et al., 2011; Cerviño-Povedano & Mora, 2011, 2015).

3. This Study

The goal of the current study is to assess the relationship between PSTM and the production of L2 Spanish vowels in stressed and unstressed contexts in two groups of English-speaking learners of Spanish and native speakers of Spanish. The research questions guiding the study are as follows:

1. What is the effect of PSTM on the production of the five Spanish monophthongs among English-speaking L2 Spanish learners?
2. What is the effect of stress on the production of the five Spanish monophthongs among English-speaking L2 Spanish learners and native speakers of Spanish?
3. Does the effect of lexical stress on vowel production differ across PSTM levels in learners?

Given previous research, we hypothesize that L2 Spanish learners with higher PSTM will produce the five monophthong vowels with acoustic properties that more closely approximate those of natives. Additionally, while we expect the native speakers and all learners alike to reduce their Spanish vowels in unstressed contexts compared to stressed contexts, we hypothesize that this reduction will be greater for the learners relative to the native speaker group. Lastly, we expect this reduction effect to be greater in lower PSTM learners than higher PSTM learners, given the previous research findings outlined that indicate that PSTM has a larger effect in global pronunciation at earlier levels.

4. Method

4.1 Participants

Data were collected from twenty English-speaking L2 learners of Spanish and five native Spanish speakers. The twenty learners were equally divided among two proficiency levels: Spanish majors and minors enrolled in fourth-year university Spanish courses, and graduate students of Spanish. Additionally, within each proficiency level, five learners were categorized as having lower PSTM while five learners were classified as having higher PSTM. These twenty learners came from the free-speech portion of a larger study that assessed PSTM, working memory, grammatical proficiency, and vocabulary proficiency, as well as a linguistic background questionnaire. Since PSTM is continuous, the learners in each proficiency level of that study were divided into low, middle and high PSTM groups. From those groups, five learners were selected for the current study, with the intent of matching grammatical proficiency, vocabulary proficiency, study abroad experience and working memory as closely as possible. We initially attempted to have similar distributions of participant sex in each participant group, although this was not possible while maintaining the other distributions. Therefore, we normalized vowel formant measures to account for physiological differences in vocal tract length due to sex. Participant characteristics are provided in Table 1.

Table 1

Participant characteristics

Level	Fourth-year	Graduate	Native

	Low PSTM (N = 5)	High PSTM (N = 5)	Low PSTM (N = 5)	High PSTM (N = 5)	(N = 5)
PSTM (/144)	65	116.4	51.4	120.2	74
Mean working memory (/75)	56.4	65.8	59.4	61	65
Mean grammar score (/20)	10.6	11.2	15	14	18
Mean vocabulary score (-60 to 60)	15.6	20.4	34.2	28.6	60
Sex	F = 1	F = 4	F = 3	F = 5	F = 4
Study abroad experience / Country of Origin	3/5	3/5	5	5	Spain = 2 Mexico = 1 PR = 1 Colombia = 1

When comparing across low and high PSTM groups within each level, the groups are similarly matched regarding their working memory, grammatical proficiency, and vocabulary proficiency scores. Additionally, in the low and high PSTM groups within the fourth-year level, three of the five participants had participated in a study abroad of six months or less, with an average length abroad of 13 weeks. On the other hand, in both PSTM groups within the graduate level, all five participants had studied or worked abroad, with an average length of time abroad of 53 weeks. Five native speakers were chosen in order to represent a range of regions (Spain, Argentina, Colombia, Mexico, Puerto Rico), similar to previous research on the L2 acquisition of Spanish vowels (e.g.,Cobb & Simonet, 2015; Menke & Face, 2010), both for dialect representation and because learners in the U.S. university context are likely to have contact with native speakers from a variety of Spanish-speaking regions.

4.2 Tasks

4.2.1 *Measure of PSTM*

Participants were administered a serial non-word recognition (SNWR) task to assess PSTM. SNWR tasks test participants' ability to discriminate between sequences of non-words that follow the phonotactic constraints of a given language. The task contained 24 pairs of CVC Russian (a language which none of the participants knew) words and non-words in sequences of varying lengths produced by a female native speaker of Russian. Eight sequences contained five, six, and seven items each. Within these sequences, the (non)words were separated by 300 ms of silence. The pairs of sequences were presented to learners either identically, such that the first sequence (i.e., A,B,C,D,E) was presented in the same order during the second repetition (i.e., a,b,c,d,e), or so that two of the (non)words in the second repetition were switched (i.e., A,**B,C**,D,E; a,**c,b**,d,e). The first and last (non)words of each sequence were never switched.

Participants were asked to respond whether the pair was the same or different. Their answers were coded as correct or incorrect and weighted according to the length of the sequence. The pairs were presented in a randomized order in OpenSesame (Mathôt, Schreij, & Theeuwes, 2012) and participants had 1000 ms to indicate their response by hitting one of two keys on the keyboard marked for *yes* and *no*; the response time was set at 1000 ms so participants they would not have time to (sub)vocally rehearse the sequences, as determined via a prior pilot study. After 1000 ms, a non-answer was recorded and coded as incorrect. The task was preceded by a practice block consisting of four sequences, of which two were identical and two were different.

4.2.2 *Production task*

Participants' production of the five Spanish monophthong vowels was obtained via a self-paced oral response task. Participants were presented with ten prompts via PowerPoint, such as *Cuéntame tus planes para este fin de semana* 'Tell me your plans for this weekend'. They were asked to respond to each of these questions for one to two minutes each. Their responses were recorded on a TASCAM DR-40 4-Track portable digital recorder with a Shure WH20XLR dynamic headset microphone in a quiet room. All participants produced a total recording of at least ten minutes in length, from which five minutes (minutes two through seven for all participants) were used for the vowel analysis.

4.2.3 *Other tasks*

Each participant also completed a grammar test, a vocabulary proficiency task and a background questionnaire. The grammar measure was the 20-item grammar cloze test from the Diploma de Español como Lengua Extranjera (DELE) (Embajada de España, Washington, DC; cf. Duffield & White [1999]). The vocabulary task was the Lextale-ESP vocabulary test containing a randomized list of 60 words and 30 non-words in Spanish (Izura, Cuetos & Brysbaert, 2014). The grammar and vocabulary scores were included to ensure that the high and low PSTM groups did not differ in proficiency, operationalized by these measures. Lastly, the background questionnaire asked participants about their linguistic history and demographic information.

4.3 Vowel analysis

All monophthongal vowels produced by participants during minutes two through seven of the recording were marked in Praat (Boersma & Weenink, 2019). Some examples are provided in Figures 1-3, where the transition between vowels and stop, nasal and fricative consonants are visible. In Figure 1, a lightening of the formant structure for nasals marks the transition from vowel to nasal. The transition between stop and vowel is marked by a lack of formant structure for the former and a clear presence of formant structure for the latter. Similarly, fricatives lack formant structure and display frication in the upper portion of the spectrograph, as in Figure 2.

[INSERT FIGURES 1,2,3 HERE]

However, the transition between vowel and lateral was frequently more subtle. Overall, laterals were lighter than vowels in the spectrograph, had a more constricted waveform, and the formant transition was generally rapid. Thus, boundaries were marked where the lightening of the spectrograph, constriction of the waveform and boundary transition occurred, as in Figure 2. However, some laterals were produced with a more vocalic quality and the transition was more gradual. Similarly, the formant transition between vowels and intervocalic approximants, [ß,δ,ɣ], was more gradual, as in Figure 3. In these cases, the vowel was marked at the halfway point in the change of constriction in the waveform. For example, in Figure 3, the end boundary of /i/ in *conmigo* ‘with me’, was located at the approximate halfway point between the start of the increase in constriction in the waveform and the most constricted point of the approximant [ɣ]. This halfway point was determined visually.

All monophthongal boundaries that could be clearly identified were included. A number of vowels (N=547) were not included because: they occurred at a word boundary adjacent to another vowel (N = 300); they formed a diphthong or hiatus due to consonant lenition (N = 79); they were reduced through devoicing or being dropped altogether (N = 116);l creaky voice was produced (N = 15); they could not be clearly identified (N = 6); they were lengthened (N = 21); or there was background noise (N = 10). This resulted in a total of 12,452 vowels that were included in the initial acoustic analysis. A little more than half were analyzed by the first author (N=6813) and the remaining were analyzed by a graduate research assistant trained by the first author (N=5639). Additionally, the first author verified all the boundary markings performed by the research assistant. Of the 5639 vowels marked by the research assistant, the first author adjusted 132 (2.34%). Although 12,452 vowels were initially included in the acoustic analysis, 2,959 of these tokens were produced in monosyllabic words and thus could not be marked for stress, so were excluded from further analysis. This resulted in 9,493 tokens used in the final analysis, distributed across participant groups as illustrated in Table 2.

Table 2

Number of tokens produced per vowel in each speaker group

Vowel	Fourth-year		Graduate		Native
	Low PSTM	High PSTM	Low PSTM	High PSTM	
/a/	594	302	773	652	755
/e/	463	282	597	533	606
/i/	226	134	302	195	303
/o/	463	245	542	476	557
/u/	81	48	122	96	146
Total	1,827	1,011	2,336	1952	2,367

Acoustic dimensions were assessed by measuring the F1 and F2 frequencies at the midpoint of each vowel segment. The Praat script used to extract the frequencies was set to include a window length of 25 ms for all tokens, with five formants estimated under 5500 Hz for females and 5000 Hz for males. However, all /u/ tokens, as well as any token with an F1 or F2 outside two SDs from the mean for that vowel, were verified for accuracy. All /u/ occurrences were verified since F2 is lowest for /u/ and it seemed that the script sometimes was unable to reliably distinguishbetween F1 and F2. The remaining tokens were verified to ensure that any outliers were not due to errors in the script or in Praat's identification of formants. This lead to the verification of formants for 1,470 tokens, of which 1262 remained unchanged (85.9%) and 208 were recoded (14.1%). Most errors occurred with /u/ and /o/, due to F2 lowering.

Since the participant groups contained both male and female speakers, and vocal tract sizes differ individually and according to sex, the F1 and F2 values were normalized using the Nearey1 method in Norm Suite (Nearey, 1977; Thomas & Tyler, 2007). The Nearey1 formula was chosen since it has been found to perform well in reducing physiological variation and preserving sociolinguistic variation (Adank, Smits, & van Hout, 2004), while not using F3 values. Not incorporating F3 was important for this dataset, as these learners produced r-coloring of vowels, which lowers F3. The Nearey1 values were subsequently scaled to Hertz-like values. All statistical analyses were then conducted using these normalized measurements.

4.4 Statistical analysis

The median F1 and F2 of each vowel in both stressed and unstressed contexts were calculated for each speaker. We opted for the median, rather than the mean, following Cobb and Simonet (2015), because the median is more resistant to skewing by a handful of outliers. This resulted in 500 observations (five speakers X five speaker groups X two formant measurements X five vowels X two stress conditions): 250 each for F1 and F2. The data were prepared this way to prevent one speaker from skewing the median for a particular vowel X stress context in their group level by having many more observations than their peers simply due to the specific words they used in their oral response task.

Once the data were prepared, two mixed-model ANOVAs were performed in SPSS, one for F1 and one for F2. For each mixed ANOVA, the within-subject factors were 'vowel' with five categories (a, e, i, o, u) and 'stress' with two categories (stressed, unstressed). 'Speaker group' was set as the between-subjects factor with five levels (fourth-year low PSTM, fourth-year high PSTM, graduate low PSTM, graduate high PSTM, native). This type of ANOVA is appropriate for our statistical analysis since our dependent variable is continuous, and our within-subject factors and our between-subjects factor consist of at least two related categorical groups. Additionally, we have independence of observations for each combination of the within-subject factors, our dependent variable was normally distributed according to a Shapiro-Wilk test of normality, and a Levene's test for homogeneity indicated that our data displayed homogeneity of variances. Our data did, however, violate one assumption of the mixed ANOVA statistical test, which is that the data should display sphericity, according to Mauchly's Test of Sphericity (SPSS), which indicated that for measurements across vowels, sphericity was violated for F1 ($p = .001$) and F2 ($p =.001$). To correct this lack of sphericity, we ran the mixed ANOVAs using a Greenhouse-Geisser correction, in which sphericity was not assumed. This correction lowers the likelihood of a Type I error, in which the model erroneously rejects the null hypothesis.

5. Results

In Figure 4 , the position of the five vowels in the vowel space for each participant group is illustrated. The mean F1 and F2 for each vowel and for each group are provided in Tables 3 and 4..

[INSERT FIGURE 4 HERE]

Table 3

Mean F1 for /a, e, i, o, u/ for each speaker group

Measurement	/a/	/e/	/i/	/o/	/u/
			Fourth-year low PSTM		
Mean	513.172	421.367	349.823	430.272	363.024
SD	35.36	13.24	21.94	19.64	19.60

Lower bound	491.350	414.634	337.729	418.423	246.145
Upper bound	534.993	428.101	361.917	442.120	379.902
			Fourth-year high PSTM		
Mean	517.026	424.985	358.624	417.498	357.959
SD	36.91	14.45	17.59	11.10	21.31
Lower bound	495.205	418.252	346.530	405.649	341.081
Upper bound	538.848	431.719	370.718	429.346	374.838
			Graduate low PSTM		
Mean	580.989	463.868	372.155	471.797	392.615
SD	31.43	11.64	14.13	6.51	15.39
Lower bound	559.167	457.135	360.061	459.949	375.736
Upper bound	602.810	470.602	384.249	483.646	409.494
			Graduate high PSTM		
Mean	568.485	453.887	376.943	482.432	394.869
SD	19.42	10.82	9.41	16.50	22.13
Lower bound	546.663	447.153	364.849	470.583	377.991
Upper bound	590.306	460.620	389.037	494.280	411.748
			Native speakers		
Mean	575.674	451.433	361.311	459.744	391.285
SD	30.74	14.34	14.22	10.68	21.99
Lower bound	553.853	444.700	349.218	447.895	374.406
Upper bound	597.496	458.166	373.405	471.592	408.163

Table 4.

Mean F2 for /a, e, i, o, u/ for each speaker group

Measurement	/a/	/e/	/i/	/o/	/u/
			Fourth-year low PSTM		
Mean	1447.74	1665.17	1770.58	1298.73	1275.16
SD	32.60	39.58	122.12	32.56	75.96
Lower bound	1422.97	1628.00	1707.05	1268.29	1225.60
Upper bound	1472.51	1702.33	1834.11	1329.16	1324.72
			Fourth-year high PSTM		
Mean	1446.81	1668.43	1842.02	1246.17	1268.73
SD	40.74	62.33	68.82	54.71	73.68
Lower bound	1422.04	1631.27	1778.50	1215.73	1219.17

Upper bound	1471.59	1705.60	1905.55	1276.60	1318.30
Graduate low PSTM					
Mean	1445.59	1654.13	1919.90	1191.34	1190.85
SD	32.13	29.98	72.34	41.07	91.75
Lower bound	1420.82	1616.97	1856.37	1160.90	1141.29
Upper bound	1470.36	1691.29	1983.42	1221.77	1240.41
Graduate high PSTM					
Mean	1416.92	1682.21	1943.67	1213.26	1107.83
SD	28.29	41.56	50.04	20.50	49.75
Lower bound	1392.15	1645.05	1880.14	1182.83	1058.27
Upper bound	1441.69	1719.37	2007.19	1243.70	1157.39
Native speakers					
Mean	1458.55	1690.45	1898.11	1223.88	1177.11
SD	19.84	42.75	86.02	42.35	88.57
Lower bound	1433.78	1653.29	1834.59	1193.45	1127.55
Upper bound	1483.32	1727.61	1961.64	1254.31	1226.68

The mixed ANOVA indicated that there was a main effect of speaker group for F1 ($F(4,20) = 68.651$, $p < 0.001$) but not for F2 ($F(4, 20) = 1.216$, $p = 0.335$). There was also a main effect of vowel for both F1 ($F(2.291, 45.825) = 499.387$, $p < 0.001$) and F2 ($F(2.325, 45.507) = 900.873$, $p < 0.001$). The mixed ANOVA revealed a main effect of stress for F1 ($F(1, 20) = 81.325$, $p < 0.001$) but not F2 ($F(1, 20) = 0.169$, $p = 0.685$). There was also an effect of vowel X stress for F1($F(3.034, 60.672) = 20.719$, $p < 0.001$) and F2 ($F(2.470, 49.405) = 25.317$, $p < 0.001$). However, for brevity, and based on our research questions, our discussion focuses on the effects of vowel X speaker group, and vowel X stress X speaker group.

The mixed ANOVA revealed that there was a significant vowel X speaker group interaction for both F1 ($F(9.165, 45.825) = 2.085$, $p = 0.050$) and F2 ($F(9.301, 46.507) = 6.207$, $p < 0.001$). For F1, fourth-year low and high PSTM learners had a significantly lower F1 value for /a/ and /e/, and thus produced these vowels significantly higher in the vowel space than the remaining groups. Graduate low and high PSTM learners, as well as native speakers, were not significantly different from each other for /a/; however, for /e/, the graduate low PSTM learners produced this sound lower in the vowel space than the graduate high PSTM learners and native speakers. Regarding /i/, none of the learner groups were significantly different from the native speakers, although the graduate high PSTM group approached significance ($p = 0.071$); however, some of the learner groups were significantly different from each other. The fourth-year low PSTM group had an F1 that was significantly lower than either graduate group, producing their /i/ higher in the vowel space. The fourth-year high PSTM group produced /i/ significantly higher than that of the graduate high PSTM group. The graduate groups were not significantly different from each other regarding /i/. For /o/, both fourth-year groups produced a higher /o/ with lower F1 values than the graduate student groups and native speakers. The graduate low PSTM learners were not significantly different from the graduate high PSTM learners nor the native speakers. However, the graduate high PSTM learners were significantly different from native speakers, producing a lower /o/ with higher F1 values.

Regarding F2, for /a/, only the graduate high PSTM learners differed from native speakers, while the remaining groups did not differ from each other. The graduate high PSTM learners produced an /a/ that was significantly backed with a lower F2 than native speakers. None of the speaker groups were significantly different from each other for /e/. However, there were significant differences between the groups for /i/. The fourth-year low PSTM learners produced an /i/ that was significantly backed and had a lower F2 relative to the native speakers and graduate students. The fourth-year high PSTM learners' /i/ was more backed than that of the graduate students, who had a more fronted /i/ relative to native speakers, although this difference was not significant between the graduate levels and the native speakers. For /o/, the fourth-year low PSTM learners differed from the other groups, producing an /o/ that was fronted with a higher F2 relative to remaining groups. Although the /o/ of the fourth-year high PSTM learners was backed relative to the fourth-year low PSTM learners, it was fronted compared to graduate low PSTM learners who displayed the furthest back /o/, although this difference was not significant between them and the graduate high PSTM learners or native speakers. Finally, /u/ was significantly fronted in both fourth-year groups relative to the remaining groups, while the graduate high learners produced a /u/ that was significantly backed compared to the remaining groups. Graduate low PSTM and native speakers were not significantly different from each other and produced a /u/ in the middle of the fourth-year learners and the graduate high PSTM learners.

In Figures 5and 6, the positions of each of the five vowels produced by the five speaker groups are provide in both stressed and unstressed contexts, respectively. The relevant means, SD and confidence intervals are provided in Tables 5 through 8, each located beneath its related figure. Lastly, and most pertinent to our research questions, there was an effect of vowel X stress X speaker group for F2 ($F(9.881, 49.405) = 2.555$, $p = 0.015$), but not F1 ($F(12.134, 60.672) = 1.053$, $p = .415$).

[INSERT FIGURE 5 HERE]

Table 5

Descriptive statistics for F1 by speaker group in stressed contexts

Measurement	/a/	/e/	/i/	/o/	/u/
			Fourth-year low PSTM		
Mean	538.63	432.93	346.93	434.79	366.24
SD	27.46	6.46	25.44	22.39	25.82
Lower bound	510.83	423.58	331.11	419.36	346.06
Upper bound	566.43	442.28	362.75	450.22	386.42
			Fourth-year high PSTM		
Mean	540.40	436.95	365.74	420.44	357.50
SD	30.60	8.79	21.03	14.90	8.64
Lower bound	512.60	427.60	349.92	405.01	337.32
Upper bound	568.21	446.30	381.56	435.87	377.68
			Graduate low PSTM		
Mean	599.38	469.84	366.33	474.60	389.52
SD	35.65	12.97	11.42	5.21	22.12
Lower bound	571.58	450.04	350.51	459.17	369.34
Upper bound	627.18	468.74	382.14	490.03	409.70
			Graduate high PSTM		
Mean	577.27	459.39	370.80	484.33	400.90
SD	23.31	10.86	1.09	22.54	23.24
Lower bound	549.47	450.04	354.98	468.90	380.72
Upper bound	605.07	468.74	386.62	499.76	421.08

Native speakers					
Mean	592.38	458.70	358.80	464.50	396.93
SD	35.65	9.85	14.72	10.45	23.85
Lower bound	564.58	449.35	342.98	449.07	376.75
Upper bound	620.18	368.05	374.62	479.93	417.11

Table 6

Descriptive statistics for F2 by speaker group in stressed contexts

Measurement	/a/	/e/	/i/	/o/	/u/
Fourth-year low PSTM					
Mean	1423.38	1667.53	1839.79	1296.15	1244.82
SD	25.94	46.31	112.47	33.23	49.23
Lower bound	1396.09	1623.26	1767.36	1255.96	1201.714
Upper bound	1450.67	1711.79	1912.22	1336.34	1287.92
Fourth-year high PSTM					
Mean	1427.31	1662.98	1878.62	1227.59	1246.13
SD	37.00	68.17	60.18	56.47	78.92
Lower bound	1400.02	1618.71	1806.19	1187.40	1203.03
Upper bound	1454.60	1707.24	1951.05	1267.78	1289.23
Graduate low PSTM					
Mean	1438.76	1661.44	1962.25	1170.57	1146.27
SD	41.18	22.79	70.11	40.78	17.09
Lower bound	1411.47	1617.17	1889.83	1130.38	1103.16
Upper bound	1466.05	1705.70	2034.68	1210.76	1189.37
Graduate high PSTM					
Mean	1405.64	1690.39	1960.32	1213.07	1121.38
SD	19.91	50.64	39.47	33.23	32.40
Lower bound	1378.35	1646.12	1887.89	1172.88	1078.28
Upper bound	1432.93	1734.65	2032.75	1253.26	1164.49
Native speakers					
Mean	1451.70	1714.33	1938.62	1225.12	1105.77
SD	12.06	37.17	86.00	50.78	26.07
Lower bound	1424.41	1670.06	1866.19	1184.93	1062.67
Upper bound	1479.00	1758.59	2011.05	1265.31	1148.88

[INSERT FIGURE 6 HERE]

Table 7

Descriptive statistics for F1 by speaker group in unstressed contexts

Measurement	/a/	/e/	/i/	/o/	/u/
			Fourth-year low PSTM		
Mean	487.71	409.81	352.72	425.75	359.80
SD	20.95	4.27	20.38	17.77	13.09
Lower bound	469.47	401.42	338.86	415.67	340.93
Upper bound	505.955	418.19	366.57	435.84	378.68
			Fourth-year high PSTM		
Mean	493.65	413.02	351.51	414.56	358.42
SD	27.59	5.89	11.27	5.79	30.76
Lower bound	475.41	404.64	337.65	404.47	339.54
Upper bound	511.891	421.41	365.36	424.64	377.30
			Graduate low PSTM		
Mean	562.60	457.90	377.99	468.99	395.71
SD	21.00	6.89	15.30	6.97	4.43
Lower bound	544.36	449.52	364.13	458.92	376.83
Upper bound	580.84	466.28	391.84	479.08	414.59
			Graduate high PSTM		
Mean	559.70	448.38	377.99	480.53	388.84
SD	10.61	8.35	15.30	9.78	21.70
Lower bound	541.46	440.00	369.23	470.45	369.96
Upper bound	577.94	456.76	396.94	490.62	407.71
			Native speakers		
Mean	558.97	444.17	363.82	454.99	385.64
SD	12.57	15.28	14.91	9.51	20.96
Lower bound	540.73	435.79	349.97	444.91	366.76
Upper bound	577.21	452.55	377.68	465.07	404.52

Table 8

Descriptive statistics for F2 by speaker group in unstressed contexts

Measurement	/a/	/e/	/i/	/o/	/u/
			Fourth-year low PSTM		
Mean	1472.11	1662.80	1701.37	1301.31	1305.50
SD	15.31	36.95	94.42	25.55	90.87
Lower bound	1446.16	1622.76	1637.38	1267.28	1227.51
Upper bound	1498.06	1702.85	1765.35	1335.34	1383.50
			Fourth-year high PSTM		
Mean	1466.31	1673.89	1805.43	1264.74	1291.34
SD	37.61	63.40	60.71	51.80	68.61
Lower bound	1440.36	1633.84	1741.44	1230.71	1213.34
Upper bound	1492.27	1713.93	1869.41	1298.77	1369.33
			Graduate low PSTM		
Mean	1452.42	1646.82	1877.54	1212.10	1235.44
SD	22.59	37.00	48.73	32.45	116.95
Lower bound	1426.47	1606.77	1813.55	1178.07	1157.44
Upper bound	1478.37	1686.86	1941.52	1246.13	1313.43
			Graduate high PSTM		
Mean	1428.20	1674.02	1927.01	1213.46	1094.27
SD	32.96	33.99	58.17	14.18	63.71
Lower bound	1402.25	1633.98	1863.03	1179.43	1016.28
Upper bound	1454.15	1714.07	1991.00	1247.49	1172.27
			Native speakers		
Mean	1465.40	1666.57	1857.60	1222.64	1248.45
SD	24.97	36.11	71.76	38.12	65.16
Lower bound	1439.45	1626.52	1793.62	1188.61	1170.46
Upper bound	1491.36	1706.61	1921.59	1256.67	1326.45

For clarity of presentation, these findings are discussed by vowel in the following subsections.

5.1 Vowel /a/

Unstressed /a/ had a significantly lower F1, and was thus higher, than stressed /a/ for all groups. Although the effect of stress was similar for all groups, pairwise comparisons indicate that the mean F1 values for stressed /a/ were significantly different across PSTM groups. Both fourth-year groups had a significantly lower F1 for stressed /a/ than the other groups. The same effect was observed for unstressed /a/ although the effect was larger than for stressed /a/, indicating that the fourth-year learners reduced unstressed /a/ more than the other speaker groups.

On the other hand, unstressed /a/ and stressed /a/ only had a significantly different F2 for both fourth-year groups and the graduate high PSTM group. In all three groups, unstressed /a/ was fronted relative to stressed /a/. The mean F2 for both stressed and unstressed /a/ was significantly different across participant groups. The graduate high PSTM group produced a stressed /a/ that was backed relative to the other groups, while their unstressed /a/ was backed compared to both fourth-year groups and native speakers.

5.2 Vowel /e/

Unstressed /e/ was significantly higher in the vowel space, and thus had a lower F1, than stressed /e/ for all groups except for the graduate high PSTM learners, for whom the effect only approached significance ($p = 0.061$). Despite the similar effect of stress, the mean F1 was significantly different across speaker groups for both stressed and unstressed /e/. The F1 of both stressed and unstressed /e/ was higher for the fourth-year groups relative to the other speaker groups.

Regarding F2, only native speakers demonstrated a significant effect of stress, who produced unstressed /e/ more backed compared to stressed /e/. The mean F2 of /e/ was not significantly different across speaker groups in stressed or unstressed contexts.

5.3 Vowel /i/

An effect of stress was not found on the height (F1) of /i/ in any speaker group. However, the mean F1 values of stressed and unstressed /i/ were significantly different across speaker groups. The fourth-year low PSTM learners had a higher stressed /i/ than the graduate high PSTM learners. For unstressed /i/, both fourth-year groups were significantly different than both graduate groups, with a higher /i/ and consequently lower F1 values. The native speakers produced an unstressed /i/ between that of the fourth-year learners and the graduate learners and were not significantly different from any learner group.

An effect of stress was observed on the F2 of /i/ in fourth-year low PSTM, fourth-year high PSTM, and graduate low PSTM learners, as well as for native speakers, but not graduate high PSTM learners. Unstressed /i/ was significantly backed compared to stressed /i/. Additionally, the mean F2 values were significantly different across speaker groups for stressed and unstressed /i/. Stressed /i/ was significantly backed in the fourth-year low PSTM group compared to both graduate groups and nearly significantly backed from that of native speakers ($p = 0.058$). Unstressed /i/ was significantly further back for the fourth-year low PSTM group compared to all other groups. Unstressed /i/ for fourth-year high PSTM learners was more fronted compared to the fourth-year low PSTM group, but backed compared to all other groups.

5.4 Vowel /o/

For F1 of /o/, no effect of stress was found in any of the speaker groups. However, there were significant differences in mean F1 across speaker groups, for both unstressed and stressed /o/. Both stressed and unstressed /o/ was significantly higher and with a lower F1 for both fourth-year groups relative to all other groups. Graduate low PSTM learners were not significantly different from native speakers in either stress context; however, graduate high PSTM learners produced stressed and unstressed /o/ significantly lower than native speakers.

Regarding F2, there was not a significant influence of stress in any of the speaker groups, although the effect approached significance for the graduate low PSTM learners ($p = 0.057$) and the fourth-year high PSTM group ($p = 0.086$). The mean F2 values were significantly different across speaker groups for both stressed /o/ and unstressed /o/. For stressed /o/, the fourth-year low PSTM learners produced a more fronted vowel compared to all other groups. The fourth-year high PSTM learners also produced stressed /o/ in a more fronted position compared to the graduate groups and the native speakers, but this difference was only significant relative to the graduate low PSTM group. The same pattern was found for unstressed /o/.

5.5 Vowel /u/

There was no effect of stress on F1 of /u/ in any of the speaker groups. However, both stressed and unstressed /u/ had significantly different mean F1s across speaker groups. The F1 values of /u/ were significantly lower in the fourth-year learner groups compared to the other groups. This difference was observed in both stressed and unstressed contexts.

Mean F2 values differed by stress only for the graduate low PSTM and native speakers. In these two groups, unstressed /u/ was fronted relative to stressed /u/. In both fourth-year groups and in the graduate high PSTM group, this effect was not observed. However, when comparing the productions of stressed and unstressed /u/ across groups, stressed /u/ was significantly fronted for both fourth-year learner groups relative to unstressed /u/ of the remaining groups, who were not significantly different from each other. The unstressed /u/ of the graduate high PSTM learners was significantly backed, relative to the other groups, who were not significantly different from each other.

In the following section, we synthesize these findings in light of the research questions that motivated the study.

6. Discussion

The first research question asked whether English-speaking Spanish learners with differing PSTM abilities produce the five Spanish monophthongs with different acoustic properties. The statistical analysis revealed that PSTM largely did not appear to affect F1 values (vowel height) for /a/, /i/, and /u/. For these vowels both high and low PSTM fourth-year learners patterned together, while both graduate groups patterned like native speakers. Only in the graduate level was there a difference in F1 values between PSTM groups for /e/ and /o/: The graduate low PSTM learners produced /e/ lower in the vowel space relative to the graduate high PSTM group and native speakers, and thus were arguably less native-like. Conversely, the graduate high PSTM learners produced /o/ lower in the vowel space relative to the native speaker group, and thus were less native-like in their production of this vowel. Consequently, for F1, the effect of PSTM seems unclear.

On the other hand, there appears to be a larger effect of PSTM on F2 values (frontness/backness) of the five Spanish vowels, given that differences were observed between groups for /a/, /i/, /o/ and /u/. The graduate high PSTM learners produced an /a/ that was significantly more backed, with a lower F2, than native speakers. There were significant differences between the groups for F2 of /i/. The fourth-year low PSTM learners backed their /i/ the most, followed by fourth-year high PSTM. Both graduate groups fronted /i/ relative to native speakers, but this difference was not significant. The fourth-year low PSTM learners fronted /o/ the most followed by fourth-year high PSTM learners. Finally, /u/ was significantly fronted in both fourth-year groups relative to the remaining groups, while the graduate high PSTM learners produced a /u/ that was significantly backed compared to the remaining groups. Overall, it appears that fourth-year learners reduced their vowel space relative to native speakers regarding F2, with the fourth-year low PSTM learners reducing theirs the most, and, thus, were less native-like. Conversely, the graduate students generally produced vowels that were closer to those of the native speaker group, with the graduate high PSTM learners expanding their vowel space past that of the graduate low PSTM and native speaker groups, particularly for /u/.

The finding that there is a larger effect of PSTM on F2 values rather than F1 values is not surprising given that vocalic reduction according to stress affects the F2 dimension more than the F1 dimension in Spanish, for both learners and native speakers (Menke & Face, 2010). Thus, learners who differ in amount of vocalic reduction would differ more on the F2 dimension. Additionally, vowels in general are subject to considerable dialectal variation in the United States, particularly on the F2 dimension (Labov, Ash, & Boberg, 2006). In particular, the position of the back vowels in certain varieties of American English are frequently fronted (e.g. Cheng, Faytak, & Cychosz, 2016), which may influence L2 production as well. Furthermore, the finding that /u/-fronting was particularly variable among groups is in line with Cobb and Simonet (2015), who found that the /u/ was the most variable vowel between experience levels in their study.

The second research question asked whether learners and native speakers produce the five Spanish monophthongs in stressed and unstressed contexts with differing acoustic properties. This question is related to research question 3, which asked whether the effect of lexical stress on vowel production differs across PSTM levels. Findings indicate that all five speaker groups demonstrated an effect of stress on vowel production, although this effect depended on the particular vowel, speaker group and formant. For F1, the effect of stress on each vowel was similar across all groups, who demonstrated an effect of stress on /a/ and /e/ that was either significant or approached significance. Both vowels were higher in the vowel space than their stressed counterparts. There was no effect of stress on F1 values for /i/, /o/, and /u/ for any speaker group. There were more differences between groups regarding the effect of stress on the vowels' F2 values. Only in the fourth-year groups and the graduate high PSTM group was there an effect of stress on /a/: unstressed /a/ was fronted relative to stressed /a/. For /e/, only for native speakers was unstressed /e/ significantly backed relative to stressed /e/. Regarding /i/, stress significantly affected F2 values for the fourth-year low PSTM, fourth-year high PSTM, graduate low PSTM and native speaker groups, but not for graduate high PSTM learners. For the first four groups, unstressed /i/ was significantly backed compared to stressed /i/, with the fourth-year low PSTM learners showing the greatest degree of centralization. The groups were similar in the lack of effect of stress on the F2 values of /o/.

Taken together, these findings indicate that PSTM can affect advanced learners' production of L2 vowels, as well as mediate the effect of stress on L2 vowel production. However, this effect is not always facilitative. At the fourth-year level, it appears that while both groups reduce their vowel space relative to the native speakers and graduate learners, the low PSTM group does so more, particularly regarding /o/ and /i/ production, in both stressed and unstressed contexts. Thus, high PSTM fourth-year learners, although they still differ from native speakers, do so less than low PSTM fourth-year learners. On the other hand, in the graduate group, it seems that the high PSTM learners are those that differ more from native speakers due to an expansion of the vowel space relative to native speakers. This finding is clearest for /u/ overall, as well as /u/ and /i/ in unstressed contexts.

Why the high PSTM graduate learners would overshoot the native speaker norm is unclear, when PSTM appears to have a facilitative effect in the lower fourth-year level, based both on our findings as well as on previous research (Inceoglu, 2019; Mora & Darcy, 2017; Moorman, 2017). It is a documented phenomenon that learners can sometimes exceed native speaker norms, even if they had previously approximated it at lower levels. For example, in morphosyntax, the overreaching of native speaker norms has been found for the use of overt subject pronouns in Spanish (Geeslin, Linford & Fafulas, 2015; Geeslin, Linford, Fafulas, Long, & Díaz-Campos, 2013; Linford, 2016). For phonology, while native speakers of Spanish produce other variants more often than the trill in prescriptive trill contexts (e.g., Bradley, 2006; Díaz-Campos, 2008; Henriksen & Willis, 2010; Willis, 2007), advanced learners of Spanish produce the trill in anywhere from two-thirds of trill contexts (graduate students; Rose, 2010) to over 80% of trill contexts (faculty members; Reeder, 1998), far surpassing actual native speaker use. It may be that, as English speakers develop in their L2 acquisition of Spanish, they progress from over-reducing both stressed and unstressed vowels to under-reducing them. The high PSTM learners in both groups may be better able to perceive the differences in vowel production between English and Spanish, and thus lessen their vocalic reduction. For the graduate learners, this may have led to under-reducing the vocalic space in the high PSTM group. Future research across a variety of class-levels and with a larger group of learners is needed to better understand this possibility, and the contexts in which it may occur.

Another possible explanation for the fact that graduate students appear to overshoot the native speaker norm, and that the high PSTM graduate learners do so in particular, are the speech rate differences between the native speakers and all learner groups. Although highly proficient, the graduate level learners nevertheless conversed at a slower speech rate than native speakers, who spoke faster and produced more phonetic reduction overall, such as intervocalic stop lenition. It may be that learners were simply more conscious of their speech, and both spoke more slowly and produced each sound more carefully, due to a desire to sound more native-like. Perhaps if data had been collected via a read aloud task, all groups would have paid more attention to their speech, and reduced less. Thus, it may be that with a different data elicitation task, the native speakers would reduce less and have a more expanded vowel space, similar to the high PSTM graduate learners. That is, if the graduate learners and native speakers were speaking at similar rates, they may have demonstrated similar amounts of vocalic reduction. This suggestion is supported by the findings of Menke and Face (2010), who used a paragraph reading task and found less reduction due to stress in their group of native speakers and fourth-year learners (graduating Spanish majors) than in our study. On the other hand, Cobb and Simonet (2015) used a word-reading task and found similar reduction in unstressed contexts to our study. Thus, this difference between learners and native speakers may not, in fact, be an artefact of speech rate or attention to speech. Future research should examine this avenue to explore more thoroughly the differences between advanced learners and native speakers.

Lastly, the conflicting findings for the effect of PSTM on L2 vowel production in our graduate learners may be due to other differences between the high and low PSTM groups within each level that were not accounted for. Although the data were normalized to account for variation in vocal tracts between male and female speakers, it is possible that the different number of male and female participants in each group otherwise skewed the results. Additionally, although the same number of participants had studied abroad in each PSTM group within each level, their study abroad experiences differed by country and likely by study abroad program type. Again, additional studies would be needed to determine the potential effects of these other variables.

7. Conclusion

This study has expanded the study of L2 vowel production by examining the relationship between PSTM, experience with the L2, and word stress. Each of these factors had previously been studied and found to influence L2 phonology, and often vowel production. Our findings confirm what previous studies had shown: that these factors do in fact relate to vowel production, with learners with more L2 experience, learners with higher PSTM, and the presence of lexical stress, on the whole, leading to more native-like vowels. The goal of this study was to shed further light on the development of L2 vowel systems, however, by exploring potential interactions between those variables. Understanding how different factors interact in affecting the production of L2 vowels can be particularly helpful when trying to explain why vowels seem more difficult to acquire than other sounds, and why even advanced learners often continue to exhibit non-native-like production.

Our findings have confirmed that learners' PSTM plays a role in accurate vowel production, both at the fourth year and graduate levels. However, it also seems that experience plays an overall larger role in the acquisition of L2 vowels: for most vowels, and in stressed and unstressed contexts, graduate learner groups patterned more like native speakers than fourth-year learner groups, regardless of PSTM in those groups. Therefore, while both PSTM and stress conditions play a part in L2 vowel production, our findings suggest that experience with the target language can largely override any of those effects. Further research should endeavor to understand the specific nature of those experiences – classroom, study abroad, learning strategies, etc. – and how they might be involved in L2 vowel production. Additionally, future studies would benefit by looking at multiple variables simultaneously, as we have here, to better understand the complex nature of second language phonology.

References

Acheson, D.J., Hamidi, M., Binder, J.R., & Postle, B.R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience, 23*, 1358-1367.

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustic Society of America, 116*, 3099-107.

Aliaga-García, C., Mora, J.C., & Cerviño-Povedano, E. (2011). L2 speech learning in adulthood and phonological short-term memory. *Poznán Studies in Contemporary Linguistics, 47*, 1-14.

Baddeley, A.D. (1986). *Oxford psychology series, No. 11. Working memory*. New York, NY: Clarendon Press/Oxford University Press.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation*, *Vol. 8,* (pp. 47-89). New York: Academic Press.

Bradley, T. G. (2006). Phonetic realizations of /sr/ clusters in Latin American Spanish. In M. Díaz-Campos (Ed.), *Selected proceedings of the 2nd conference on laboratory* *approaches to Spanish phonetics and phonology*(pp. 1–13). Somerville, MA: Cascadilla Proceedings Project.

Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.01, retrieved 14 August 2019 from http://www.praat.org/.

Bradlow, A.R. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America, 97*, 1916-1924.

Cardinal, R.N., & Aitken, M.R.F. (2006). *ANOVA for the behavioural sciences researcher.* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Cerviño-Povedano, E.,& Mora, J.C. (2011). Investigating Catalan learners of English over-reliance on duration: vowel cue weighting and phonological short-term memory. In K. Dziubalska-Kolaczyk, M. Wrenbel, &M. Kul (Eds.), *Achievements and perspectives in SLA of speech: New Sounds 2010. Volume 1* (pp. 56-64). Frankfurt am Main: Peter Lang.

Cerviño-Povedano, E., & Mora, J.C. (2015). Spanish EFL learners' categorization of /i:-I/ and phonological short-term memory. *Procedia- Social and Behavioral Sciences, 173,* 18-23.

Cheng, A., Faytak, M., & Cychosz, M. (2016). Language, race, and vowel space: Contemporary Californian English. In E. Clem, V. Dawson, A. Shen, A.H. Skilton, G. Bacon, A. Cheng & E.H. Maier (Eds.), *Proceedings of the 42nd Annual Meeting of the Berkeley Linguistics Society* (pp. 62-78). Berkeley, CA: Berkeley Linguistics Society.

Cobb, K., & Simonet, M. (2015). Adult second language learning of Spanish vowels. *Hispania, 98*, 47-60.

Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics in Language Teaching, 7*, 295-325.

Díaz-Campos, M. (2008). Variable production of the trill in spontaneous speech: Sociolinguistic implications. In L. Colantoni & J. Steele (Eds.), *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 47-58). Somerville, MA: Cascadilla Proceedings Project.

Duffield, N., & White, L. (1999). Assessing L2 knowledge of Spanish clitic placement: Converging methodologies. *Second Language Research, 15,* 133-160.

Farvardin, M.T., Afghari, A., & Koosha, M. (2014). The effect of dual n-back task training on phonological memory expansion in adult EFL learners at the beginner level. *English Language Teaching, 7,* 137-143.

Flege, J., & Bohn, O.S. (1989). An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition, 11,* 35-62.

Geeslin, K., Linford, B., & Fafulas, S. (2015). Variable subject expression in second language Spanish: Uncovering the developmental sequence and predictive linguistic factors. In A. M. Carvalho, R. Orozco, & N.L. Shin (Eds.), *Subject Pronoun Expression in Spanish: A Cross-dialectal perspective.* Washington, DC: Georgetown University Press.

Geeslin, K., Linford, B., Fafulas, S., Long, A., & Díaz-Campos, M. (2013). The group vs. the individual: Subject expression in L2 Spanish. In J. Aaron, J.C. Cabrelli Amaro, G. Lord,& A. de Prada Pérez (Eds.), *Selected Proceedings of the 16th Hispanic Linguistics Symposium* (pp. 156-174). Somerville, MA: Cascadilla Proceedings.

Henriksen, N., & E. W. Willis. (2010). Acoustic characterization of phonemic trill production in Jerezano Andalusian Spanish. In M. Ortega-Llebaria (Ed.), *Selected proceedings of the fourth conference on laboratory approaches to Spanish phonology*(pp. 115–127). Somerville, MA: Cascadilla Proceedings Project.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience, 15*, 673-682.

Hu, X., Ackermann, H., Martin, J.A., Erb, M., Winkler, S., & Reiterer, S.M. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: behavioral predictors and neural substrates. *Brain and Language, 127*, 366-376.

Hualde, J.I. (2005). *The sounds of Spanish*. New York, NY: Cambridge University Press.

Inceoglu, S. (2019). L2 French vowel production: The relationship with speech perception and phonological memory. In J. Levis, C. Nagle, & E. F. Todey (Eds.), *Proceedings* *of the 10th Pronunciation in Second Language Learning and Teaching Conference* (pp. 147-157). Ames, IA: Iowa State University.

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to (rapidly and efficiently assess the Spanish vocabulary size. *Psicologica*, *35*, 49-66.

Kane, M.J., Hambrick, D.Z., Tuholski, S.W., Wilhelm, O., Payne, T.W., & Engle, R.W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of experimental psychology: General, 133,* 189-217.

Kissling, E. M. (2014). What predicts the effectiveness of foreign-language pronunciation instruction? Investigating the role of perception and other individual differences. *Canadian Modern Language Review, 70*, 532-558.

Kondo, A. (2012). Phonological memory and L2 pronunciation skills. In A. Stewart & N. Sonda (Eds.), *JALT 2011 Conference Proceedings.* Tokyo: JALT.

Kormos, J. & Safar, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11,* 261-271.

Ladefoged, P. (2006). *A Course in Phonetics.* Boston: Thomson Wadsworth.

Linford, B. (2016). *The second-language development of dialect-specific morphosyntactic variation in Spanish during study abroad* (Unpublished doctoral dissertation). Indiana University, Bloomington.

Long, A.Y., Solon, M., & Bongiovanni, S. (2018). Context of learning of second language development of Spanish vowels. *Studies in Hispanic and Lusophone Linguistics, 11*, 59-87.

Maridaki-Kassotaki, K. (2002). The relation between phonological memory skills and reading ability in Greek-speaking children: Practical applications. *European Journal of the Psychology of Education, 1,* 63-73.

Masoura, E.V., & Gathercole, S.E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory, 13*, 422-429.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*, 314-324.

Menke, M.R. & Face, T.L. (2010). Second language Spanish vowel production: An acoustic analysis. *Studies in Hispanic and Lusophone Linguistics, 3,* 181-214.

Moorman, C.M. (2017). *Individual differences and linguistic factors in the development of mid vowels in L2 Spanish learners: A longitudinal study* (Unpublished doctoral dissertation). Georgetown University, Washington DC.

Mora, J.C. & Darcy, I. (2017). The relationship between cognitive control and pronunciation in a second language. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 95-120). Bristol, UK: Multilingual Matters.

Nagle, C. (2013). A reexamination of ultimate attainment in L2 phonology: Length of immersion, motivation, and phonological short-term memory. In E. Voss, S.-J.D. Tai, &Z. Li. (Eds.), *Selected proceedings of the 2011 Second Language Research Forum: Converging theory and practice* (pp. 148-161). Somerville, MA: Cascadilla Proceedings Project.

Nearey, T.M. (1977). *Phonetic feature systems for vowels* (Unpublished doctoral dissertation). University of Alberta, Alberta, CA. Reprinted 1978 by the Indiana University Linguistics Club.

O'Brien, I. (2006). *Phonological memory and second language speech production: A longitudinal study of English-speaking adults learning Spanish* (Unpublished doctoral dissertation). Université du Québec, Montréal.

Redick, T.S., Broadway, J.M., Meier, M.E., Kuriakose, P.S., Unsworth, N., Kane, M.J., & Engle, R.W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*, 164-171.

Reeder, J. T. (1998). English speakers' acquisition of voiceless stops and trills in L2 Spanish. *Texas Papers in Foreign Language Education, 3*, 101–108.

Reiterer, S.M., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., Winkler, S., & Ackermann, H. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: functional neuro-imaging and brain morphology. *Frontiers in Psychology, 2*, 271.

Rose, M. (2010). Intervocalic tap and trill production in the acquisition of Spanish as a second language. *Studies in Hispanic and Lusophone Linguistics, 3*, 379–419.

Skrzypek, A., & Singleton, D. (2013). *Vigo International Journal of Applied Linguistics, 10,* 105-129.

Slevc, L.R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science, 17*, 675-681.

Solon, M., Long, A., & Gurzynski-Weiss, L. (2017). Task complexity, language-related episodes, and production of L2 Spanish vowels. *Studies in Second Language Acquisition, 39*, 347-380.

Speciale, G., Ellis, N.C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics, 25*, 293-321.

Syrdal, A.K., & Gopal, H.S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustic Society of America, 79,* 1086-1100.

Thomas, E.R. & Tyler, K. (2007). NORM: The vowel normalization and plotting suite. [Online resource: http://lingtools.uoregon.edu/norm/norm1.php].

Tsuchihira, T. (2007). L2 working memory capacity and L2 listening test scores of Japanese junior college students. *Journal of Bunkyo Gakuin University, Department of Foreign Languages and Bunkyo Gokuin College, 7,* 159-175.

Tukey, J.W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Venkatagiri, H. S. & Levis, J.M. (2007). Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness, 16*, 263-277.

Willis, E. W. (2005). An initial examination of southwest Spanish vowels. *Southwest Journal of Linguistics, 24*, 185-198.

Willis, Erik W. W. (2007). An acoustic study of the “pre-aspirated trill” in narrative Cibaeño Dominican Spanish. *Journal of the International Phonetic Association, 37*, 33–49.

Resumen

En este estudio, se analiza la producción de las cinco vocales españolas en sílabas tónicas y átonas por aprendices angloparlantes de español divididos en dos niveles de competencia y que se diferencian por su capacidad de memoria fonológica a corto plazo. Nuestros resultados muestran una menor tendencia por parte de los aprendices con mayor capacidad de memoria fonológica a reducirlas vocales en español tanto en contextos tónicos como átonos. Por consiguiente, entre los estudiantes de cuarto curso, los aprendices con mayor memoria fonológica produjeron vocales que se parecían más a las vocales de los hablantes nativos que los aprendices con menor memoria fonológica. Por otro lado, el espacio vocálico producido por los aprendices graduados con mayor memoria fonológica resultó ser más extenso que el de los nativos, particularmente para la /u/, en comparación con los aprendices graduados con menor memoria fonológica, cuyas vocales se parecían más a las de los hablantes nativos. Estos resultados apuntan a que la memoria fonológica juegaun papel importante en la fonología de segundas lenguas incluso en los niveles más avanzados, lo que podría explicar por qué hay diferencias individuales en la pronunciación de segundas lenguas incluso en los hablantes con competencia muy avanzada.

Palabras claves: *Adquisición de segundas lenguas, fonética/fonología, diferencias individuales, memoria fonológica, vocales*

.