RESEARCH

Relative difficulty in the L2 acquisition of the Spanish dorsal fricative

Matthew Patience

Research on relative difficulty in L2 production has revealed that learners target the most salient parameter when acquiring new sounds (Colantoni & Steele, 2008). For example, L1 English-L2 French learners acquire the more salient fricative manner of the French /ʁ/ before the voicing and duration parameters (Colantoni & Steele, 2007, 2008). Previous work in this framework has not compared the acquisition of place and manner parameters. If the more salient parameter is targeted first, we should expect L2 learners to acquire the manner of articulation before the place of articulation, given that manner is a more salient feature than place (Miller & Nicely, 1955; Bedoin et al., 2013). This hypothesis was tested by investigating the L2 production of the Spanish voiceless dorsal fricative by L1 English speakers living in Madrid, a region in which the fricative has a strident realization (Hualde, 2014) and a uvular place of articulation (Ibabe et al., 2016). Fourteen L1 English-L2 Spanish speakers and 14 native Spanish controls performed a picture description task that elicited the target in two vocalic contexts: [aχe, eχa]. An acoustic analysis revealed that the L2 speakers produced fricatives with a similar amplitude compared to controls. However, in the [eχa] context, the learners produced fricatives with a more anterior place of articulation and less frication. The results are consistent with the finding from previous work that learners focus on the most salient property when learning new segments, and provide further evidence that vocalic context is an important factor in production difficulty.

Keywords: L2 phonetics; L2 acquisition; fricative; articulatory complexity; perceptual salience

1. Introduction

One of the primary goals in research on the acquisition of non-native sounds is to determine sources of difficulty. Scholars investigating L2 production have found that a segment's parameters (e.g., voicing, manner) are not acquired simultaneously (Colantoni & Steele, 2007, 2008), indicating that some parameters are more difficult to acquire than others. Previous work has also found that the individual parameters are acquired in some phonetic contexts (e.g., intervocalic) before others, and that neighboring vowels can hinder or facilitate acquisition (Yavas, 1997; Waltmunson, 2005). Research on relative difficulty is particularly important in the field of L2 speech, because it has been demonstrated that current theories of L2 acquisition are often unable to predict and/or explain the sources of difficulty in L2 production. More specifically, existing theories cannot account for the successive acquisition of parameters and/or they cannot account for difficulty that arises due to phonetic context or neighboring segments (see Colantoni & Steele, 2008 for discussion). For example, the Markedness Differential Hypothesis (Eckman, 1977, 2008) predicts that segments

that do not exist in a learner's L1, and are marked, will be difficult to acquire. It makes no claims, however, regarding which property of a segment may be difficult (e.g., manner, voicing). Previous work on the acquisition of the French uvular rhotic /ʁ/ has revealed that the most salient parameter (manner) is acquired in the most salient positions (onset, intervocalic) first (Colantoni & Steele, 2007, 2008). However, these studies did not investigate place of articulation, therefore it is unknown whether place is likely to be acquired before manner, or vice versa. In order to develop a theory of relative difficulty, it is necessary to understand how and when learners acquire the different parameters of non-native sounds, and which factors influence the acquisition of such parameters. Manner has consistently been shown to be a more salient parameter than place of articulation (Miller & Nicely, 1955; Peters, 1963; Bedoin & Krifi, 2009; Bedoin, Krzonowski, & Ferragne, 2013; Gonzalez-Poot, 2014; Schwartz, 2017). If learners target the more salient parameter in acquisition, we should expect manner to be acquired before place of articulation. In the present study, this hypothesis was examined by investigating the L1 English-L2 Spanish production of the voiceless dorsal fricative in two vocalic contexts. In the Madrid dialect, the fricative is generally produced with a uvular place of articulation (Ibabe, Petrirena, & Aguirrezabal, 2016) and a high degree of

University of Toronto, CA

matt.patience@utoronto.ca

stridency (Hualde, 2014). L1 English speakers learning the Madrid dialect must therefore acquire a segment with a place of articulation that does not exist in their L1, in addition to learning how to produce high amplitude frication at the new place of articulation. The two vocalic contexts in which the fricative was produced were expected to further influence the degree of difficulty, based on previous findings that adjacent vowels can contribute to the difficulty of articulating a neighboring segment (Yavas, 1997; Waltmunson, 2005). While the large majority of studies investigating L1 English-L2 Spanish production are conducted on L2 speakers learning the target language via classroom instruction, and with very little target language input, the speakers of the present study had been living in Madrid for on average 4.37 years (*SD* = 3.48). Accordingly, the productions were from speakers who had received a great deal of input from Spanish, and had a large amount of experience speaking Spanish. Moreover, while L2 production is often elicited through reading tasks or carrier sentences, the speakers of the present study performed a semi-spontaneous picture description task. The productions analyzed are therefore expected to be more representative of truly spontaneous speech than productions elicited in more controlled reading tasks.

In the remainder of this section, the phonetic and phonological characteristics of the Spanish dorsal fricative are discussed first, with a focus on the place and manner parameters. North American English segments that may influence the acquisition of /χ/ are also described. This discussion is followed by a summary of previous work on relative difficulty in L2 speech, and a summary of previous findings on the L2 acquisition of the Spanish dorsal fricative. The relative salience of manner compared to place parameters is then considered. The section concludes by detailing the research questions and predictions of the present experiment.

1.1. Relevant phonetic and phonological characteristics of Spanish and English

The Spanish dorsal fricative is typically described as a voiceless velar fricative /x/. However, it is generally considered to have a more posterior place of articulation in central Spain, ranging from either post-velar to uvular (Schwegler et al., 2010; Hualde, 2014). While no comprehensive acoustic studies have been performed on the Spanish fricative, an articulatory MRI study confirmed that the fricative is produced with primarily a uvular place of articulation (Ibabe et al., 2016). It has also been reported that the exact place of articulation is partially influenced by neighboring vowels, with more posterior articulations arising after central /a/ and posterior vowels /o, u/ (Hualde, 2014). An additional characteristic of the central Spanish fricative is that it is produced with a more strident articulation and greater degree of frication than the velar fricative produced in other dialects (Schwegler et al., 2010; Hualde, 2014). However, the amount of frication present is somewhat variable (Schwegler et al., 2010). Phonologically, /χ/ only occurs in onsets, either word-initial (e.g., /χa.mas/ 'never') or intervocalically (/a.χo/

'garlic'). It is represented orthographically by <j>, <x>, and also by <g> when it precedes <i> or <e>.

The uvular fricative is an articulatorily complex segment,[1] both in terms of manner and place of articulation. Regarding place of articulation, Lindblom and Maddieson (1988) argue that segments requiring a high degree of articulatory displacement (e.g., uvular, retroflex, pharyngeal) are more complex than segments whose place of articulation is in a near rest position (i.e., bilabial, dental/alveolar, and velar). Regarding manner, research on phonetic constraints demonstrates that fricatives are highly complex, due to the precise aerodynamic requirements required for their production (i.e., sufficient air pressure and a narrow constriction in the vocal tract). A small change to the air pressure or size of the constriction could result in a segment produced with no noise, thus resembling an approximant. Moreover, strident fricatives are more complex than less strident fricatives, because they are produced with a greater amplitude, which is achieved with a higher rate of air flow through the constriction (in the case of /χ/).

Given the inherent articulatory complexity of /χ/, we might expect it to be a difficult sound for non-native speakers to produce, especially because North American English does not have a dorsal fricative. The most similar English segment is (arguably) the glottal fricative /h/, given the similar manner and posterior articulation. Note, however, that in addition to having a different place of articulation, /h/ also has a much lower intensity than /χ/ (Martínez Celdrán & Fernandez Planas, 2007). The only other somewhat related segments in English are the velar stops [k, g], which share a similar place of articulation. While English speakers have experience producing velar segments (i.e., /k, g/), they do not have experience producing segments with uvular places of articulation, or producing dorsal segments with frication (velar or uvular). As a result, acquiring the articulation of /χ/ could be a challenging task for L1 English speakers.

Note that orthographic influence is not expected to play a role in the acquisition of /χ/ by L1 English-L2 Spanish speakers, based on the fact that the segments represented by <j>, <x>, and <g> in English share little similarity to /χ/.

1.2. Relative difficulty in L2 acquisition

Research on L2 production has revealed that the parameters of non-native segments (manner, place of articulation, voicing, duration) are not acquired simultaneously. Some parameters are more difficult than others, and the same parameter may be easier to acquire in certain word positions. This has been referred to as relative difficulty (see Colantoni et al., 2015: 61 for discussion). The results from studies using this approach suggest that the relative salience of the parameters can influence the order in which they are acquired. Colantoni and Steele (2007, 2008) investigated the acquisition of the French uvular rhotic in four word positions: initial, intervocalic, word-medial coda position, and word-final coda position. Colantoni and Steele (2008) also investigated the L2 acquisition of the Spanish tap. The French learners demonstrated nativelike production of the manner parameter, but only in the more

salient initial and intervocalic positions. In contrast, the duration and voicing parameters were not produced with nativelike values in these positions. Colantoni and Steele argue that learners target the more salient parameter in the most salient positions, which explains why manner was the first parameter acquired, and also why it was acquired first in the more salient contexts. While the L2 Spanish learners also acquired the manner of /ɾ/ in the more salient intervocalic position, this was likely due to the fact that English has a flap allophone that exists in the same context.

Other research on difficulty in L2 speech has revealed that the vocalic context in which consonants are produced can greatly influence the complexity of production, and should therefore be considered in research on relative difficulty. Yavas (1997) investigated the acquisition of voicing in the L2 English production of word final /b, d, g/, by L1 Mandarin, L1 Portuguese, and L1 Japanese speakers. The learners had more difficulty voicing /d, g/ when in the context of a high vowel, a pattern also observed in native English speakers. The results are explained by Yavas in terms of articulatory constraints. High vowels create a more narrow constriction in the vocal tract, which disfavors voicing. Similar effects of high vowels were found in Waltmunson (2005), who investigated the acquisition of the trill in L1 English-L2 Spanish speakers. His speakers produced trills less accurately when they were preceded by a high vowel. These results can also be explained by articulatory constraints. The more narrow constriction created by a preceding high vowel would make it more difficult to produce the air pressure required for trilling. Moreover, as discussed in Sole (2002), the tongue position of a high front vowel (tongue predorsum raising/fronting), competes with the tongue position required for an alveolar trill (predorsum lowering/backing). Consequently, producing a trill following a high vowel is particularly difficult. Both the Yavas and Waltmunson studies demonstrate that the complexity of a consonant's articulation varies depending on its neighboring segments. We should thus expect that the difficulty of articulating the uvular fricative may also vary according to its neighboring vowels. Detailed predictions are laid out in Section 1.5.

1.3. L2 acquisition of /χ/
To my knowledge, no previous work has investigated the acquisition of the Spanish dorsal fricative (from any dialect). However, some scholars have observed that L1 English speakers initially substitute [h] (Schwegler, 2010; Fernández, 2012), which is likely due to perceptual assimilation. Note that the dialect the L2 speakers were learning was not specified, thus it is not clear whether L2 speakers learning /χ/ might also initially produce [h]. However, [h] substitutions should be considered a possibility.

1.4. Relative salience of manner versus place
A significant amount of work has investigated the saliency of place compared to manner, and these studies have consistently identified manner as the more salient

parameter (Miller & Nicely, 1955; Peters, 1963; Bedoin & Krifi, 2009; Bedoin et al., 2013; Gonzalez-Poot, 2014; Schwartz, 2017). For example, in Miller and Nicely (1955), listeners heard consonants that were masked with noise and spoken over a voice communication system. They had to identify which consonant they heard. The results revealed that speakers had the most difficulty correctly distinguishing place of articulation contrasts. The authors suggest that place features are easily seen on a speaker's lips, which could explain why place features are generally less salient than other features such as manner and duration. Similar results were found in Peters (1963). In the author's study, listeners rated the similarity of numerous English contrasts. The similarity scores were highest in the segments that contrasted according to place, and lowest in segments that contrasted according to manner. Data from cognitive research supports the finding that manner is a more salient parameter. Using ERPs to analyze sensitivity to segmental properties, Bedoin et al. (2013) found that manner was more easily detected than place in voiceless consonants. These studies all demonstrate that manner is generally considered to be highly salient, especially when compared to place of articulation. Consequently, we can be relatively certain that the manner of /χ/ is more salient than place, especially because /χ/ is not just a fricative but a strident (and therefore louder) fricative.

1.5. Current study
The present paper had two objectives: (1) Given that previous research on relative difficulty has not yet examined the acquisition of the place of articulation parameter, the goal here was to determine whether the more salient manner of /χ/ is acquired before the place of articulation. (2) We know from previous work that the vocalic context can influence the articulatory complexity of neighboring segments (Yavas, 1997; Waltmunson, 2005). Therefore the second goal was to determine whether the difficulty of acquiring the place and manner parameters of /χ/ varies with the vocalic contexts. To achieve these objectives, the study was designed to answer two questions, which are presented below with their respective hypotheses.

RQ1. Do L1 English-L2 Spanish speakers experience more difficulty acquiring the target manner of /χ/, or the target place of articulation?
Previous work on L2 fricative production has revealed that learners first acquire the most salient parameter (Colantoni & Steele, 2007, 2008). As discussed in Section 1.4, manner is a more salient feature than place of articulation. If learners target the most salient parameter first, we should expect learners to have less difficulty with the manner parameter than the place parameter.

RQ2. Do learners experience more difficulty producing /χ/ in the a_e as opposed to e_a context?
The neighboring vowels of a segment can influence the segment's production difficulty. Consequently, the ease of producing /χ/ is expected to vary according to the context in which it is realized. Regarding manner, learners should experience more difficulty producing the target

Table 1: Summary of L1 Spanish control and L1 English-L2 Spanish speaker participant profiles.

Language	N	Age	AoA	LoR	Instructed exposure	Naturalistic exposure	Oral proficiency
L1 Spanish	14	20.9 (3.1)	NA	NA	NA	NA	NA
L1 English	14	2.36 (0.45)	1.3 (0.69)	4.37 (3.48)	5.85 (3.77)	4.40 (3.52)	2.3 (0.6)

Notes: AoA = age of onset of acquisition of Spanish; LoR = length of residence in Madrid. All duration values represent years. Means are reported, with standard deviations in brackets.

manner of /χ/ following /e/, given the greater distance between the tongue and the target place of articulation (and therefore a longer distance that the tongue must travel to create a sufficiently narrow constriction that is required for producing a fricative). Regarding place of articulation, we might expect the a_e context to be easier. The more posterior position of the tongue when /χ/ follows /a/ should make the uvular place of articulation easier to attain by the L2 speakers.

2. Methodology

2.1. Participants

14 L1 English-L2 Spanish and 14 L1 Spanish controls participated in the study. Each group consisted of seven male and seven female speakers. All participants completed a detailed language background questionnaire, which was used to ensure they met the required criteria. All L1 Spanish control speakers were from Madrid and had lived in Madrid since birth. They were all native Spanish speakers with no or limited experience speaking other languages.

All of the L1 English-L2 Spanish speakers were native speakers of North American English, had been living for a minimum of six months in Madrid, and had spent the large majority of their Spanish learning experience in Madrid. One speaker had spent three weeks in Mexico, but had been living in Madrid for 11 years. Another speaker had spent four months travelling through Mexico, several years before moving to Madrid. However, this speaker had been living in Madrid for nearly five years at the time of testing. Therefore, while speakers may at some point have had exposure to other dialects, the large majority of their input was from their time living in Madrid. All L2 speakers spoke Spanish daily and claimed to speak and be most familiar with the Madrid dialect. Participants were asked to self-rate their Spanish ability, which they rated from advanced to near-native. Independent measures of oral proficiency were also established, via accentedness ratings performed by native Spanish speakers. Accent ratings are a more objective and comparable measure than years of study combined with non-standardized labels such as beginner/advanced. Moreover, accentedness ratings specifically target oral proficiency (as opposed to proficiency in other domains, such as the lexicon or morphosyntax), thus are arguably the most relevant measure for determining oral proficiency in a non-native language (Colantoni, Escudero, & Steele, 2015, p. 89). For this reason, they have often been used to determine oral proficiency in previous work (e.g., Colantoni & Steele, 2007; Colantoni & Steele, 2008; Kopečková, 2016; Lloyd-Smith et al., 2017).

To establish the participants' oral proficiency, all L2 Spanish participants were required to read "The North Wind and the Sun" in Spanish. The recordings were then presented in random order to native (n = 9) Spanish speakers, who listened to, and rated from one to five, how strong they felt each speaker's accent was.[2] The scores from all judges were averaged, and the resulting values were used as measures of each speaker's overall oral proficiency. A summary of the participant profiles, including oral proficiency, is displayed in **Table 1**. Note that while all L2 participants had a significant amount of experience speaking Spanish, and considered themselves to be advanced speakers, they had noticeable accents, with an average oral proficiency rating of 2.3/5.

2.2. Task and stimuli

The present paper's data comes from a larger project investigating the L2 acquisition of Spanish segments by native English speakers. The project required all segments to be produced in two vocalic contexts (e_a and a_e), and in similar sentential positions (sentence initial and sentence final). To meet these requirements, speakers produced simple SVO sentences describing pictures involving two characters with nonce names. The names contained the target segments. In the present paper, the two stimuli produced with /χ/ were *Cheja* /ˈtʃeχa/ and *Laje* /laχe/. The pictures were designed to elicit the production of each name in either subject position (N = 4) or object position (N = 4). Participants were also asked to briefly introduce the characters (e.g., *This is Cheja*). Therefore, each target was produced 10 times (four times in subject position, four times in object position, and twice when being introduced), in two contexts (a_e, e_a), resulting in 20 productions of /χ/ per speaker.

The motivation for using a picture description task was to analyze the speakers' ability to produce the target segments in a semi-spontaneous and complex linguistic context. A picture description task involves a relatively high level of difficulty in which the speakers are not expected to focus specifically on their articulation, thus the elicited production should represent the speakers' oral ability when using language in a real world context (i.e., spontaneous speech).

2.3. Procedure

Participants were presented with a series of drawings on a computer screen. Each drawing had two characters at the top, and four scenarios involving each of the two characters. Below each character at the top of the screen were images of speakers, which the participants had to click on to
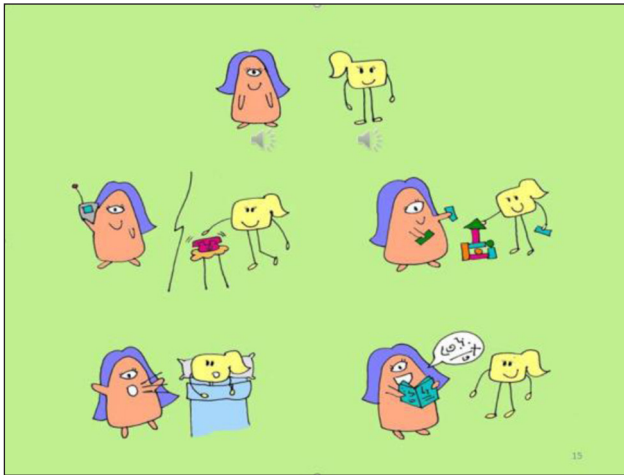
Figure 1: Example of a picture that elicited productions of the target stimuli.

hear the names of the characters. After listening to the names, the participants were required to first introduce the characters. They subsequently had to produce simple subject-verb-object utterances using the names of the characters. So, for example, in **Figure 1**, the participant was expected to produce the following (in Spanish):

This is Nague and this is Cheja. In the first picture, Nague calls Cheja. In the second picture, Nague plays with Cheja. In the third picture, Nague wakes up Cheja. In the last picture, Nague reads to Cheja.

Note that participants were allowed to listen to the names of the stimuli more than once if they forgot what they were.

2.4. Data Analysis

Recordings were extracted and examined acoustically in Praat (Boersma & Weenink, 2017). Segment boundaries were marked according to the onset and offset of noise that was visible in both the waveform and spectrogram. An acoustic analysis was subsequently conducted on the target productions, consisting of five acoustic measures which were calculated to determine place and manner of articulation: two of the spectral moments (center of gravity, kurtosis), F2 transition, intensity ratio, and zero-crossings ratio.

The spectral moments are one of the most frequently used measures to determine place of articulation in fricatives. The center of gravity indicates the frequency range of the frication noise, whereas the kurtosis indicates the concentration of noise. These values vary according to the size of the oral cavity, thus can be used to infer place of articulation. While a comprehensive acoustic analysis of the Spanish fricative has not been conducted, previous acoustic analyses on languages with a phonemic velar-uvular contrast (Tlingit) found that the uvular fricative is produced with a higher center of gravity, and a lower kurtosis (Denzer-King, 2013). The spectral moments were calculated using time-averaged Discrete Fourier Transform (DFT) windows, as described in Shadle, Cohn, Fougeron, and Huffman (2012).

The F2 transition is another measure that can determine place of articulation (Thomas, 2010). Gordon, Barthmaier, and Sands (2002) investigated the production of /x/ and /χ/ in Aleut. They found that the F2 transition was a more reliable acoustic measure than the spectral moments for distinguishing between velar and uvular fricatives, with uvular fricatives displaying a lower F2 transition than velars. In the present study, the transition was calculated by measuring the F2 at the end-point of the vowel preceding /χ/. Values were subsequently converted to ERB, to make average values of the F2 transition more comparable across speakers.

To examine the manner of articulation, two measures were analyzed: a zero-crossings ratio, and an average intensity ratio. The number of zero-crossings indicates how frequently a waveform crosses over the time axis and can be used to indicate how periodic or aperiodic a signal is (Martinez Celdrán, 2015; Fuchs, 2016: 138). The higher the value, the more noisy the signal, and therefore the more fricative-like the production (i.e., less vowel-like). The zero-crossings ratio was calculated as the number of zero-crossings of the target segment divided by the number of zero-crossings of the preceding vowel. As a result, a value of 1 would indicate that the production had a similar level of periodicity as the vowel. As the number increases, it indicates a progressively less periodic and more fricative-like sound. If the L2 group experiences difficulty producing frication, we might expect more approximant-like productions and therefore a lower zero-crossings ratio.

A segment's intensity indicates how loudly the sound is realized, and thus can be used to determine whether a segment is produced with lesser or greater stridency. The intensity ratio was calculated by dividing the intensity of the segment over the intensity of the vowel. The L2 group may have difficulty producing nativelike levels of stridency. We would therefore expect them to produce segments with a lower intensity ratio than controls. This would also be the case if the L2 speakers do produce [h] substitutions. A summary of the five acoustic measures and their interpretation is summarized in **Table 2**.

After calculating the values of the target acoustic measures, five generalized linear mixed effects models were run, with the target acoustic measure as the dependent variable. Language and Context were included as predictors, as was the interaction Language*Context. A random intercept of Participant nested under Gender was also included, to control for standard inter-speaker and gender-based variation.[3] All statistics were run in SPSS v. 23, with a significance level of $p = .05$. In the presence of a significant interaction, post-hoc pairwise comparisons were run using Fisher's LSD, in order to examine specific group effects.

3. Results

The results of the experiment are presented in this section, beginning first with the place of articulation analysis, followed by the manner of articulation analysis.

Table 2: Summary of the five acoustic measures and their interpretation.

Acoustic measure	Interpretation
Place of articulation	
Center of gravity	higher value = more posterior articulation
Kurtosis	lower value = more posterior articulation
F2 transition	lower value = more posterior articulation
Manner of articulation	
Zero-crossings ratio	higher value = less periodic (i.e., more fricative-like) articulation
Intensity ratio	higher value = more strident articulation
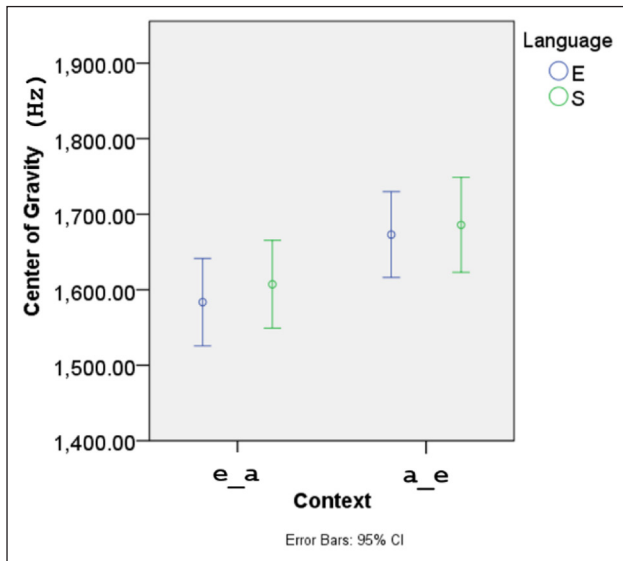
Figure 2: Center of gravity results by language group and vocalic context.
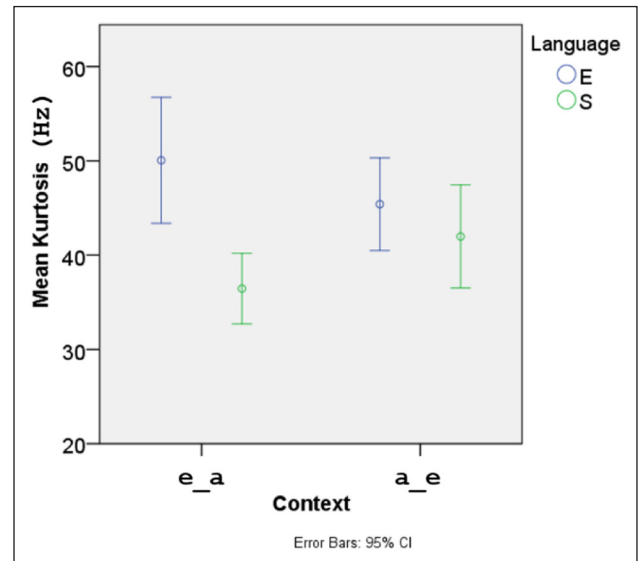

Figure 3: Kurtosis results by language group and vocalic context.

Table 3: Results of a generalized linear mixed-effects model examining center of gravity.

Fixed effects	β	SE	t	p-value
Language	–24.164	96.264	–0.251	.802
Context	–75.085	31.203	–2.406	.016
Language*Context	–19.960	45.288	–0.435	.664

Table 4: Results of a generalized linear mixed-effects model examining kurtosis.

Fixed effects	β	SE	t	p-value
Language	2.439	7.253	0.336	.737
Context	–5.996	3.145	–1.907	.057
Language*Context	11.421	4.564	2.502	.013

3.1. Place of articulation

Three measures were examined to analyze place of articulation, including two of the spectral moments and the F2 transition. **Figure 2** displays the center of gravity results. Overall, the values were similar for the two groups of speakers, and both groups displayed a tendency to produce a higher center of gravity in the a_e context.

The results of a generalized linear mixed effects model are displayed in **Table 3**. The model only revealed an effect for context, with productions in the a_e context being produced with a higher center of gravity than productions in the e_a context. Recall from Section 2.4 that a higher center of gravity indicates a more posterior place of articulation. These results therefore demonstrate that while the fricative was generally produced with a more posterior place of articulation in the a_e context, no differences were observed across language groups.

Figure 3 displays the kurtosis results. The L2 group showed a tendency to produce fricatives with a higher kurtosis than the control group, especially in the e_a context.

A generalized linear mixed effects model (**Table 4**) revealed no effect of language; however, an interaction between language and context was observed, indicating that the effect of context differed across language groups. The interaction was examined in greater detail via a post-hoc pairwise comparison. The comparison revealed no difference between groups in the a_e context ($\beta = -2.439$, $SE = 7.253$, $t = -0.336$, $p = .737$). However, the difference between groups in the e_a context approached significant ($\beta = -13.860$, $SE = 7.231$, $t = -1.917$, $p = .056$). These results indicate that the language*context interaction was primarily driven by a lower kurtosis value in the control compared to L2 speakers, in the e_a
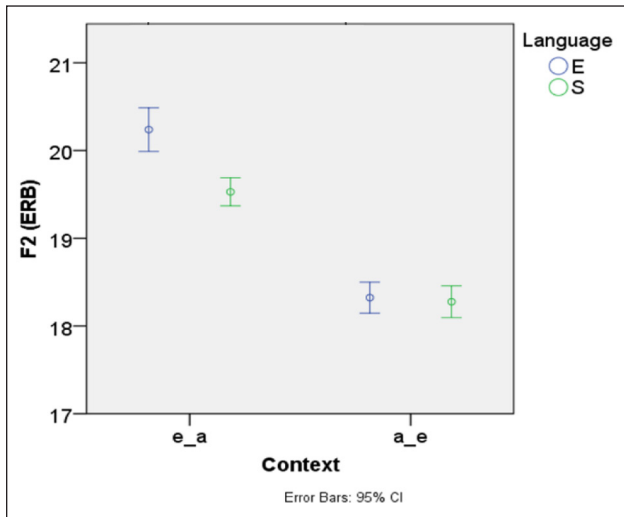
Figure 4: F2 transition results by language group and vocalic context.

Table 5: Results of a generalized linear mixed-effects model examining F2 transition.

Fixed effects	β	SE	t	p-value
Language	35.191	41.457	0.849	.396
Context	227.374	18.561	12.25	.000
Language*Context	172.664	26.881	6.423	.000

Figure 5: Mean relative number of zero crossings by both speaker groups in two vocalic contexts.

Table 6: Results of a generalized linear mixed-effects model examining the relative number of zero crossings.

Fixed effects	β	SE	t	p-value
Language	–0.146	0.211	–0.689	.491
Context	0.616	0.073	8.428	.000
Language*Context	–0.342	0.104	–3.287	.001

context. A lower kurtosis indicates a more posterior place of articulation, thus, the kurtosis results indicate that the L2 group produced a more anterior fricative, when following /e/.

In addition to the spectral moment measures, the F2 transition between the target fricative and the preceding vowel was examined, given that it should be a reliable measure for distinguishing between a velar and uvular place of articulation (Gordon et al., 2002). The F2 transition results are displayed in **Figure 4**.

The results of a generalized linear mixed effects model (**Table 5**) revealed no effect of language. However, an effect for context and an interaction between language and context was found. A post-hoc pairwise comparison examining differences between language groups in both contexts revealed that the F2 transition was higher in the e_a context for the L2 group compared to the control group ($\beta = 207.855$, $SE = 43.621$, $t = 4.765$, $p < .000$). No difference was observed between groups in the a_e context ($\beta = 35.191$, $SE = 41.457$, $t = 0.849$, $p = .396$). A higher F2 transition denotes a more anterior place of articulation, thus the data show that both groups of speakers produced more posterior fricatives in the a_e context, and that the L2 group produced fricatives with a more anterior place of articulation in the e_a context than the control group.

To recapitulate, the place of articulation analysis revealed consistent results across the three acoustic measures. Overall, the L2 group produced /χ/ with a more anterior place of articulation than the control group, in the e_a context.
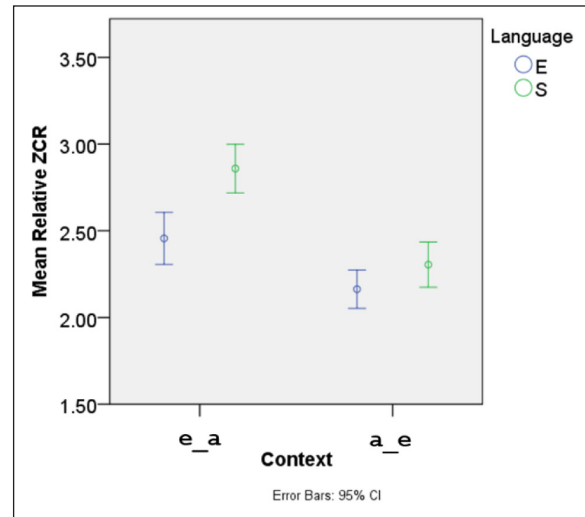
3.2. Manner of articulation

The manner of articulation analysis consisted of two measures: the relative zero-crossing rate, and the relative intensity. The relative zero-crossing rate was calculated by dividing the zero-crossing rate of the target fricative over the zero-crossing rate of the preceding vowel. **Figure 5** displays the results. We can see that the L2 group had a lower zero-crossing rate in both contexts, with a larger difference occurring in the e_a context. Both groups also produced fricatives with a lower zero-crossing rate in the e_a compared to the a_e context.

A generalized linear mixed effects model (**Table 6**) revealed an effect of context, indicating a higher relative zero-crossing rate in the e_a context. An interaction between language and context was also found. A post-hoc pairwise comparison examining differences between language groups in each context revealed that the L2 speakers produced a lower relative zero-crossing rate in the e_a context ($\beta = -0.488$, $SE = 0.211$, $t = -2.307$, $p = .021$) compared to the control speakers. No difference was observed in the a_e context ($\beta = -0.146$, $SE = 0.211$, $t = -0.689$, $p = .491$). Recall that a higher zero-crossing rate signifies a noisier (less approximant-like) fricative, thus the results demonstrate that the L2 speakers produced a more approximant-like fricative than the control speakers in the e_a context.

Figure 6 displays the results for the mean relative intensity, which was calculated as the average intensity of the target fricative divided by the average intensity of the preceding vowel. We can see that the L2 speakers produced fricatives with a slightly higher
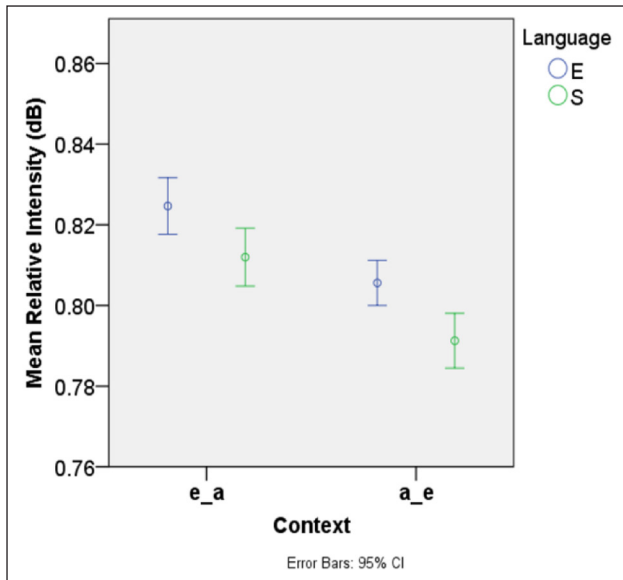
Figure 6: Amplitude results by language group and vocalic context.

Table 7: Results of a generalized linear mixed-effects model examining relative intensity.

Fixed effects	β	SE	t	p-value
Language	0.014	0.011	1.241	.215
Context	0.020	0.003	6.025	.000
Language*Context	–0.002	0.005	–0.400	.689

average relative intensity compared to the control group.

The results of a generalized linear mixed effects model (**Table 7**) revealed only an effect of context, indicating that the intensity was lower in the a_e compared to e_a context. No difference was observed between languages.

In sum, the results from the manner of articulation analysis revealed that the L2 group produced segments with less frication in the e_a context compared to the control group. In contrast, no differences were observed in the intensity of the fricatives produced by each group.

4. Discussion

In the present experiment, participants performed a picture description task that was designed to elicit production of /χ/ in intervocalic position. The objective was to determine whether L2 speakers have more difficulty acquiring the manner or place of articulation parameters, and whether the vocalic context influences the difficulty of acquiring the two parameters. In this section, the results of the research questions are summarized. This is followed by a discussion of how the results relate to research on relative difficulty. The section concludes by identifying the study's limitations, and proposing future avenues for research on relative difficulty in L2 speech.

4.1. Summary and discussion of RQs

The first research question sought to determine whether L1 English-L2 Spanish speakers experience more difficulty acquiring the place or manner parameter of /χ/. Learners

Table 8: Summary of results of L2 compared to L1 productions, overall and by vocalic context.

Acoustic parameter	Overall	Context	
		e_a	a_e
Center of gravity	✓	✓	✓
Kurtosis	✓	More anterior	✓
F2 transition	✓	More anterior	✓
Zero crossings	✓	Less frication	✓
Intensity	✓	✓	✓

Note: Results refer to L2 compared to L1 speakers; ✓ = no significant difference (i.e., native-like).

were expected to successfully acquire the fricative manner of /χ/, given its relatively higher degree of saliency. The results, which are summarized in **Table 8**, partially support the prediction. The L2 group did acquire the most salient property of the manner of articulation parameter, namely, the amplitude. However, they experienced some difficulty producing a fricative with the same degree of frication as the control group, in the e_a context. Therefore, we cannot conclude that manner was acquired with nativelike accuracy. However, the fact that the most salient property of /χ/ (i.e., it's amplitude) was acquired, suggests that the L2 group did indeed target the most salient aspect of the target fricative. Regarding place of articulation, the L2 group produced more anterior fricatives than the control group in one context (e_a); therefore, as predicted, the place of articulation was indeed difficult for the L2 speakers to acquire.

The second question of interest was whether L2 speakers have more difficulty producing /χ/ in the e_a or a_e context. Both manner and place of articulation were expected to be more difficult in the former. The results support this prediction. The L2 group produced /χ/ with less frication than the control group in the e_a context, which, as laid out in the predictions, can be explained by the longer distance that the tongue has to travel from the /e/ (front vowel) articulation to create frication than it does from the mid vowel /a/. The L2 group also produced /χ/ with a more anterior place of articulation than the control group in the e_a context. We can expect that the native speaker articulations of the present study were produced with a uvular place of articulation (or a post-velar articulation, given that /χ/ in Spanish is slightly fronted with front vowels). The results therefore suggest that the L2 speakers were producing segments with a velar place of articulation. Again, the difference in difficulty was likely due to the fronted tongue position required to produce the front vowel /e/. One could also propose that the more anterior productions of /χ/ were due to crosslinguistic influence from English. While English does not have a velar fricative, it does have a velar nasal and velar stops. Therefore, L1 English-L2 Spanish speakers may be more likely to produce fricatives at a velar as opposed to uvular place of articulation, due to L1 articulatory routines. However, the L2 group did not experience difficulty producing uvular fricatives in the a_e context. Therefore, it is unlikely that crosslinguistic influence from English was the sole reason for the velar articulations in the e_a context.

4.2. Comparison to previous studies

The results of the present study are consistent with previous work on relative difficulty. Colantoni & Steele (2007, 2008) found that L1 English-L2 French speakers first acquired the most salient property of the French /ʁ/ (manner) in the most salient positions (initial, intervocalic). As a result, the authors proposed that L2 speakers initially target the most salient parameter when acquiring non-native sounds. In the present study, the learners also acquired the most salient property first, namely the high amplitude of the Spanish /χ/. These results thus provide further evidence that learners target the properties of sounds that are most prominent. Nevertheless, learners are also restricted by articulatory constraints. In the current study, as well as in Yavas (1997) and Waltmunson (2005), the difficulty of producing the target was partially dependent on the vocalic context. These findings demonstrate the importance of considering neighboring vowels as a potential source of variability in L2 production, given that the tongue position required to articulate the vowels can create less favorable positions for realizing adjacent consonants.

4.3. Limitations

The primary limitation of the present paper is that the place of articulation analysis was conducted using acoustic measures. Consequently, while the results clearly indicated that the L2 speakers produced fricatives with a more anterior place of articulation, it is not possible to know what the exact place of articulation was for both groups of speakers. We can only infer and assume from the data that the L2 group was producing velar fricatives, while the control group was producing uvular ones. An additional limitation is that the present study only examined production of /χ/ by experienced L2 speakers. Future work should examine production by less and more advanced speakers, to achieve a better understanding of the order of acquisition of place and manner parameters. An articulatory study would be ideal, in order to determine in detail the extent to which the articulations of the L2 speakers differ from those of native speakers. Future research should also investigate how the acquisition of parameters varies with different types of segments, such as stops, trills, or laterals. While the results of the present study and those of Colantoni & Steele (2007;2008) reveal that the most salient properties are often acquired first, this may not always be the case. For example, research on the L2 acquisition of the Spanish trill has found that L1 English-L2 Spanish speakers experience a great deal of difficulty acquiring the trill manner (Waltmunson, 2005; Face, 2006; Johnson, 2008), despite the fact that the manner is highly salient. While these studies did not explicitly compare the acquisition of manner to that of place, voicing, or duration, we might expect the trill manner to be acquired after the other parameters, due to the trill's articulatory complexity.

5. Conclusion

The goal of the present paper was to contribute to our understanding of relative difficulty in L2 production. Previous work on relative difficulty has revealed that the parameters of non-native segments are not acquired simultaneously. The more salient parameter is targeted by learners, and generally acquired first (Colantoni & Steele, 2007, 2008). The present study investigated whether this was also the case in the acquisition of the place and manner parameters of the Spanish uvular fricative. The results revealed that learners did acquire the most salient property first, and thus provide further evidence that at least in the case of fricatives, learners initially acquire the salient manner property. However, the learners of the present study experienced difficulty producing fricatives with the same degree of noise as native speakers, and with the same place of articulation, when they produced /χ/ following /e/, but not following /a/. These findings demonstrate that learners do not only experience difficulty with individual segments, but also with sequences of articulatory gestures. L2 speech models should therefore be able to account for contextual difficulty, as well as the difficulty of each parameter.

Notes

[1] Ideally, two separate analyses would be conducted for the place of articulation measures, one for males, and one for females. Vocal tract sizes vary by gender, and the acoustic analyses are dependent on the size of the vocal tract. Nevertheless, due to the limited number of participants, separating the results into two groups would greatly diminish the power of the statistical analysis. For this reason, a single analysis was run, with a random intercept in which participant was nested under gender, to control for variation. Given that the groups were made up of equal amounts of males and females, the results should still be representative of the general population (to the extent that a small sample size can be representative).

[2] The following scale was provided to judges (in Spanish): 1 = 'clearly non-native, very strong foreign accent'; 2 = 'Strong foreign accent'; 3 = 'noticeable foreign accent, but not too strong'; 4 = 'almost no accent'; 5 = 'no accent (native speaker)'. Note that the scale included half points, and was therefore a nine point scale (1 – 1.5 – 2 – 2.5 – 3 – 3.5 – 4 – 4.5 – 5).

[3] In the present paper, "articulatory complexity" is used to refer to segments that consist of at least one articulatory gesture that is considered to be complex, either from an articulatory or phonetic perspective. Note, however, that complexity is a relative term. For example, voiced velar stops are considered more complex than voiced alveolar stops, because a more posterior place of articulation creates a smaller oral cavity. A smaller cavity leads to a decrease in supraglottal air pressure, which disfavors voicing (Ohala, 1997).

Acknowledgements

References

Bedoin, N., & **Krifi, S.** (2009). The complexity of phonetic features organisation in reading. In: Pellegrino, F., Marsico, E., Chitoran, I., & Coupé, C. (eds.), *Approaches to phonological complexity, Phonology & Phonetics,* 265–294. Berlin, DE: Mouton de Gruyter.

Bedoin, N., Krzonowski, J., & **Ferragne, E.** (2013). How voicing, place and manner of articulation differently modulate event-related potentials associated with response inhibition. In: *INTERSPEECH*, 906–910.

Boersma, P., & **Weenink, D.** (2017). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.29, retrieved 24 May 2017 from: http://www.praat.org/.

Colantoni, L., & **Steele, J.** (2007). Acquiring /ʁ/ in context. *Studies in Second Language Acquisition, 29*(3), 381–406.

Colantoni, L., & **Steele, J.** (2008). Integrating articulatory constraints into models of second language phonological acquisition. *Applied Psycholinguistics, 29*(3), 489–534. DOI: https://doi.org/10.1017/S0142716408080223

Colantoni, L., Steele, J., & **Escudero, P.** (2015). *Second language speech.* Cambridge, UK: Cambridge University Press.

Denzer-King, R. E. (2013). *The acoustics of uvulars in Tlingit* (Master's thesis). Retrieved from: https://rucore.libraries.rutgers.edu/rutgers-lib/39596/PDF/1/play/.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27,* 315–330. DOI: https://doi.org/10.1111/j.1467-1770.1977.tb00124.x

Eckman, F. (2008). Typological markedness and second language phonology. *Phonology and second language acquisition,* 95–115. Amsterdam, NL: John Benjamins.

Face, T. L. (2006). Intervocalic rhotic pronunciation by adult learners of Spanish as a second language. In: Klee, C. A., & Face, T. L. (eds.), *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, 47–58. Somerville, MA: Cascadilla Press.

Fernández, S. (2012). *El español como segunda lengua en anglohablantes [Spanish as a second language in English speakers]* (Master's thesis, University of Oviedo, Asturias, Spain). Retrieved from: http://digibuo.uniovi.es/dspace/bitstream/10651/4077/3/TFM_SergioFernandez.pdf.

Fuchs, R. (2016). *Speech Rhythm in Varieties of English.* Singapore: Springer. DOI: https://doi.org/10.1007/978-3-662-47818-9

Gonzalez-Poot, A. A. (2014). Conflict Resolution in the Spanish L2 Acquisition of Yucatec Ejectives: L1, L2, and Universal Constraints. In: Teddiman, L. (ed.), *Proceedings of the 2014 annual conference of the Canadian Linguistic Association,* 1–15.

Gordon, M., Barthmaier, P., & **Sands, K.** (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association, 32*(2), 141–174. DOI: https://doi.org/10.1017/S0025100302001020

Hualde, J. (2014). *Los sonidos del español.* Cambridge, UK: Cambridge.

Ibabe, A., Petrirena, R., & **Aguirrezabal, I.** (2016). ¿Son velares las consonantes velares del español? In: Planas, Ed. F., & Ma, A. (eds.), *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, 49–57.

Johnson, K. E. (2008). *Second language acquisition of the Spanish multiple vibrant consonant* (Doctoral dissertation). Retrieved from: ProQuest Dissertations and Theses. (Accession Order No. [3330747]).

Kopečková, R. (2016). The bilingual advantage in L3 learning: a developmental study of rhotic sounds. *International Journal of Multilingualism, 13*(4), 410–425. DOI: https://doi.org/10.1080/14790718.2016.1217605

Lindblom, B., & **Maddieson, I.** (1988). Phonetic universals in consonant systems. In: Hyman, L. M., & Li, C. N. (eds.), *Language, speech and mind. Studies in honour of Victoria A. Fromkin*, 62–78. London, UK: Routledge.

Lloyd-Smith, A., Gyllstad, H., & **Kupisch, T.** (2017). Transfer into L3 English. Global accent in German-dominant heritage speakers of Turkish. *Linguistic Approaches to Bilingualism, 7*(2), 131–162. DOI: https://doi.org/10.1075/lab.15013.llo

Martínez Celdrán, E. (2015). Naturaleza fonética de la consonante 'ye' en español. *Normas, 5*(1), 117–131.

Martínez Celdrán, E., & **Fernández Planas, A. M.** (2007). *Manual de fonética española: Articulaciones y sonidos de español* [Manual of Spanish phonetics: Articulations and sounds of Spanish]. Barcelona, ES: Editorial Ariel.

Ohala, J. J. (1997). Aerodynamics of phonology. In: *Proceedings of the 4th Seoul International Conference on Linguistics*, 92–97.

Peters, R. W. (1963). Dimensions of perception for consonants. *The Journal of the Acoustical Society of America, 35*(12), 1985–1989. DOI: https://doi.org/10.1121/1.1918876

Schwartz, G. (2017). Formalizing modulation and the emergence of phonological heads. *Glossa: a journal of general linguistics, 2*(1), 1–20. DOI: https://doi.org/10.5334/gjgl.465

Schwegler, A., Kempff, J., & **Ameal-Guerra, A.** (2010). Fonética y fonología españolas [Spanish phonetics and Phonology]. Hoboken, N.J.: John Wiley & Sons.

Shadle, C. H., Cohn, A., Fougeron, C., & **Huffman, M.** (2012). Acoustics and aerodynamics of fricatives. In: Cohn, A. C., Fougeron, C., & Huffman, M. K. (eds.), *The Oxford handbook of laboratory phonology,* 511–526. Oxford, UK: Oxford University Press.

Thomas, E. (2010). *Sociophonetics: an introduction.* Hampshire, UK: Palgrave Macmillan.

Yavas, M. (1997). The effects of vowel height and place of articulation in interlanguage final stop devoicing. *IRAL: International Review of Applied Linguistics in Language Teaching, 35*(2), 115.

WHITE R🌑SE
UNIVERSITY PRESS
Universities of Leeds, Sheffield & York

OPEN ACCESS 🔓