

L2 perception of Spanish palatal variants across different tasks*

CHRISTINE SHEA
JEFFREY RENAUD
*Department of Spanish and Portuguese,
University of Iowa*

(Received: September 11, 2011; final revision received: January 13, 2013; accepted: January 14, 2013; first published online 7 March 2013)

While considerable dialectal variation exists, almost all varieties of Spanish exhibit some sort of alternation in terms of the palatal obstruent segments. Typically, the palatal affricate [tʃ] tends to occur in word onset following a pause and in specific linear phonotactic environments. The palatal fricative [j] tends to occur in syllable onset in other contexts. We show that listeners' perceptual sensitivity to the palatal alternation depends upon the task and exposure to Spanish input. For native Spanish listeners, the palatal alternation boosts segmentation accuracy on an artificial speech segmentation task and also reduces latencies on a phonotactically-conditioned elision task. L2 Spanish listeners, on the other hand, only benefit from the palatal alternation in the second task. These results suggest that while Spanish L2 learners benefit from the presence of the alternation in linear phonotactic terms, this benefit does not carry over to a more abstract segmentation task.

Keywords: Spanish, allophones, L2 phonology, speech perception

Introduction

Second language (L2) learners are faced with many challenges when acquiring their target language. One of the most basic involves detecting and learning the words of the other language while managing their already established native (L1) linguistic system. This task is made even more difficult when we consider the high variability present in the speech signal. Variability can be indexical in nature (due to individual speaker characteristics) or phonotactically conditioned. Examples of phonotactically-conditioned variation include alternations that are more likely to occur in particular contexts than others, such as aspiration of voiceless stops in English. In the present study, we examine whether adult second language learners can make use of this type of information as well by examining their performance under different task demands.

The alternation tested involves the Spanish palatal obstruents. In many varieties of Spanish, the palatal affricate [tʃ] tends to occur in word onset following a pause and in specific linear phonotactic environments (following a nasal or a lateral). The palatal fricative [j] tends to occur in syllable onset in other contexts. While the actual phonetic realization may vary across dialects, almost all varieties of Spanish exhibit some sort of alternation in terms of these segments. Given this, we hypothesize that

native Spanish listeners will rely upon these statistical regularities of their native language when carrying out speech tasks. That is, they will tend to relate the palatal affricate to word onset position and the fricative to word medial. Since this knowledge depends upon extensive experience with Spanish, we also tested L2 Spanish listeners on the same set of tasks. We hypothesized that due to their more limited experience with Spanish, they will not have established the relationship between segments and positions that the native listeners are predicted to have. We also hypothesized that task demands will play a role in how these cross-linguistic effects play out. To test this, we conducted a series of four experiments to investigate how task demands modulate the way in which L2 listeners use distributional information in the input. In Experiments 1 and 2, we examined how English phonological relationships and phonetic cues play into the way L1 English/L2 Spanish listeners perceive the palatal alternation. In Experiment 3 participants carried out an artificial language segmentation task in which we varied the presence of the palatal variant. Finally, in Experiment 4, participants had to identify a word that followed an onset syllable. The final sound of the onset syllable either favored or disfavored, in phonotactic terms, the particular palatal onset that followed. These experiments combine to paint a picture of how L2 learners perceive and make use of the same information in the speech signal across a variety of different tasks.

* We would like to thank the anonymous reviewers for recommendations on how to improve this manuscript. We would also like to thank participants at the University of Iowa and the Centro de Enseñanza Extranjera at the Universidad Nacional Autónoma de México (CELE-UNAM) in Mexico City.

Palatals in Spanish and English

Spanish attests numerous palatal and palatoalveolar phonetic categories, graphically corresponding to *ll* (e.g.,

Address for correspondence:

Christine Shea, 412 Phillips Hall, Department of Spanish and Portuguese, University of Iowa, Iowa City, IA 52240, USA
christine-shea@uiowa.edu

llave “key”), *y* (*playa* “beach”) and *hi(e)* (*hierro* “iron”). Cross-dialectally the Spanish palatal obstruents exhibit a tremendous degree of variability. In Castilian Spanish, these targets can range from a glide [j] to fricative [ç] to affricate [tʃ] along a continuum of constriction degrees (Aguilar, 1997, p. 69, in Hualde, 2005, p. 165). In this variety, the voiced palatal continuant [j] usually occurs after a vowel or a continuant consonant (e.g., in *maya* [ˈmaja] “Mayan”, *la llave* [la ˈjaβe] “the key”, *la hierba* [la ˈjerβa] “the grass”; see Hualde, 2005) while a voiced affricate or plosive occurs after a lateral or nasal consonant or after a pause. In central Mexican and Caribbean Spanish, however, affricate productions are often found in intervocalic position (see Jimenez Sabater 1975, pp. 108–110; Lope Blanch 1989, 1996; Martin Butragueño, forthcoming, in Campos-Astorkiza, 2012, p. 98). In northern Mexico and parts of the American Southwest, the glide predominates (see Alvar 1996 in Campos-Astorkiza, 2012, p. 98; Lipski 1990). As discussed below, Standard Argentinian Spanish (including Uruguay), has a different pronunciation of these segments, exhibiting a prepalatal fricative [ç] where other dialects have a palatal target. Indeed, when studying these segments, Hualde (2004, 2005) states that it is best to conceptualize them as occurring along a continuum of dialect-dependent and position-dependent factors.

There have been few exhaustive studies carried out on the palatal obstruents and their phonetic realizations across dialects. Work by Aguilar (1997) on the Castilian dialect stands out, as does work by Martínez-Celdrán (2008). Aguilar (1997) examined glide [j] and high front vowel [i] productions by Castilian Spanish speakers. She found that the glide [j] is shorter than the vowel (an average of 82 ms vs. 110 ms, respectively) and with higher F1 and lower F2 values. In terms of the obstruents, Martínez Celdrán and Fernández Planas (2007) claim that there are only two obstruent palatal phones in Spanish, [ç] and [tʃ], and one palatal semivowel [j]. The first is an approximant, manifesting little if any frication in its articulation while for the second, the authors prefer to call it a double articulation, rather than an affricate, given that the second element does not exhibit the frication that accompanies the release of the stop.

Phonological analyses vary as to whether the fricative is a phoneme itself, due to its [consonantal] feature, or rather a consonantal variant of the high front vowel, conditioned by position in the word (Hualde, 2004). For the purpose of the present study, we remain agnostic as to precisely how the palatal obstruents are represented.

The palatal targets tend toward complementary distribution determined in part by three linguistic factors: prosody, the preceding segment and word position. The likelihood of either the approximant [j] or fricative [ç] is higher in unstressed syllables, while stressed syllables favor strengthening, resulting in either the fricative [ç]

or affricate [tʃ]. The surrounding phonetic context, specifically the preceding sound, also influences the alternation. The approximant and fricative are more likely in post-vocalic position, while post-pause, post-nasal and post-lateral positions favor the affricate or even the palatal stop [tʃ]. Word-internal position (medial or final) increases the probability of the approximant and fricative, while the affricate and stop are more likely to appear in word-initial position.

While not directly relevant to the study at hand, external factors such as speech rate, register and formality also affect the alternation. Faster and/or informal speech favors the approximant and fricative while slow, monitored, formal and/or emphatic speech increases the likelihood of the affricate and stop variants. As has been implied throughout this discussion, this is not to say, for example, that the approximant is categorically barred from contexts that tend to favor the affricate, or vice versa. To the contrary, every variant from the Spanish palatal continuum is POSSIBLE in any given context, yet particular contexts seem to FAVOR one variant (or one end of the continuum) relative to the other. The empirical goal of the present study is therefore to determine if native and L2 listeners use these TENDENCIES in speech segmentation and word recognition tasks. For example, we will argue in more detail in Sections 4 and 5 below that such tendencies aid word segmentation in Spanish; the likelihood of encountering a word boundary when presented with a palatal affricate or stop is higher than when presented with either the fricative or approximant given that word-initial position FAVORS the stop or affricate relative to the fricative or approximant. Even though the fricative and approximant are both POSSIBLE in this position, their appearance is LESS LIKELY compared to the likelihood of, for example, the affricate, even in running/connected speech (Piñeros, 2009, p. 207).

The fact that the three linguistic factors interact and cannot be applied independently further confirms that the alternation is one of general tendencies and not categorical absolutes. The issue is compounded when the external factors are also considered in natural speech outside of the empirical setting. To take as examples, consider *una llave* “a key” and *cónyuge* “spouse”. In the first case, the palatal segment orthographically corresponding to *ll* is stressed and word-initial, two factors that suggest the likelihood of a stop or affricate. On the other hand, it is simultaneously post-vocalic, which otherwise tends to favor the approximant or fricative. In the second case, the palatal segment, here represented by *y*, is internal and unstressed, both of which are normally associated with the approximant or fricative. Phonetically, however, it is post-nasal, which favors strengthening to the stop or affricate. It is this confluence of factors that demonstrates why any given alternant might be possible in any given position. Even so, tendencies such as those discussed above can

still be observed in the data and the present study seeks to determine if the two groups of listeners can apply these tendencies across four different speech tasks.

In the case of the L2 group, knowledge of the palatal alternation is a question of acquisition that we pursue experimentally considering that it is not “transferable” from English. English contains at least three palatal phonetic categories: approximant [j], voiced [ʒ] and voiceless [ç] palatoalveolar fricatives and voiced palatoalveolar affricate [dʒ]. At the moment we are only concerned with the palatal approximant [j] and palatoalveolar affricate [dʒ] in comparing the two systems as our L2 learners were not exposed to the Argentinean/Uruguayan Spanish dialects that employ [ʒ] and [ç], nor was our test stimuli representative of such dialects. Unlike Spanish, however, English palatals do not exist on a continuum and are not representative of predictable variants, as substituting one palatal for another in a given word runs the risk of changing the meaning. Both the approximant [j] and affricate [dʒ] can appear in initial stressed (*yet, jet*), initial unstressed (*yoursself, judgmental*), internal stressed (*beyond, pajamas*) and internal unstressed (*kayak, major*) positions. Furthermore, the preceding segments play no role, as both freely appear in post-vocalic (*a year, a jeer*), post-nasal (*unyielding, enjoin*) and post-lateral (*all year, all jeer*) positions. Additionally, the palatoalveolar affricate [dʒ] is a possible coda segment in English (*bridge*), which is categorically impossible in Spanish.

To summarize, given that English palatals [j] and [dʒ] are not predictable based on position, do not exist on a continuum and are not interchangeable without semantic consequence, the L2 Spanish learner coming from English is faced with the task of acquiring novel phonetic categories (palatal fricative, affricate and stop) and recognizing that, although each palatal segment is POSSIBLE in all positions, not unlike English, each position tends to FAVOR one category over the other, a fact for which English offers no guidance.

For the purposes of this paper, the positing of an underlying form is not necessary. We assume that the occurrence of either the less-consonant-like [j] or the more consonant-like [ç] is due to where in the word or phrase the sound occurs. Moreover, as stated above, the alternation is non-categorical in nature. Indeed, in probabilistic terms, the affricate alternant will be more likely to occur in phrase-initial position and the fricative alternant in word-medial position but this is never 100% and will depend, as stated above, upon phonotactics, orthography and dialectal variation. Nonetheless, our goal was to investigate whether learners are aware of the role played by the affricate variant in indicating word-onsets, rather than recognizing the difference between the two variants and their positional restrictions. Thus, even though both variants can probabilistically occur in word

onset, the affricate is highly unlikely to occur in word-medial, intervocalic position, suggesting that it plays a strong role in recognizing potential onsets to lexical candidates in Spanish.

L2 Speech perception and allophonic variants

Research has shown that the pattern of allophonic alternations in the listener’s native language influences speech perception (Boomershine, Hall, Hume & Johnson, 2008; Dupoux, Pallier, Sebastián-Gallés & Mehler, 1997). In a study directly related to English and Spanish, Boomershine et al. (2008) investigated the perception of [d], [ð] and [r] by speakers of Spanish and English. In English, [r] and [d] may alternate with each other while [ð] does not alternate with either; in Spanish, on the other hand, [ð] and [d] alternate while [r] is contrastive. Boomershine et al. found that English listeners rated [r] and [d] as more similar to each other than Spanish listeners did, while [ð] and [d] were more similar for Spanish listeners than for English listeners. In a subsequent speeded discrimination task, Boomershine et al. (2008) found that these cross-linguistic patterns of perceptual similarity also held. The authors concluded that the phonological relationships that hold in the listener’s native phonological inventory play a determining role in speech perception.

The results from Boomershine et al. show that L1 segmental relationships affect L2 perception. However, such alternations are by definition conditioned by the environment in which they occur and experienced listeners can take advantage of this probabilistic knowledge when carrying out speech perception tasks. One area where this knowledge is particularly useful is in the segmentation of the continuous speech stream. An extensive body of research has shown that infants and adult learners use all types of information in their search for word boundaries. While different segmentation heuristics are available, one particular mechanism has been shown to operate across all groups studied so far: the tracking of statistical information, such as transitional probabilities (TPs: Saffran, Aslin & Newport, 1996). The essential logic behind tracking TPs is that within-word probabilities across syllables are higher than across-word probabilities, and where a probabilistic trough occurs, the listener assumes a word boundary has been encountered.

Researchers have recently begun to explore how such a mechanism might operate in adults acquiring a second language. One key issue revolves around how previous learning potentially interferes with tracking a new set of statistical relations and how separate statistical relations between languages are maintained. Current work shows that in general, listeners exposed to two different “languages” (distinguished by different TPs) require either indexical information (Weiss, Gerfen & Mitchel, 2009) or pauses and explicit instruction (Gebhart, Aslin

& Newport, 2009) to inform them that there were two separate sets of statistics to be tracked.

Previous learning can also affect the way in which statistics are calculated across the input. For example, language-specific phonotactic knowledge can affect the way in which TPs are tracked and such knowledge, accumulated over years of experience with a native language, may impede successful storage and tracking of new regularities in the input. Research by Finn and Hudson Kam (2008) examined how native language phonotactic knowledge drives the segmentation of an artificial language stream by native English speakers. In their study, listeners preferred words that violated transitional probabilities but respected phonotactic regularities when tested on both after exposure to an artificial language stream. The authors interpret this finding to mean that linguistic knowledge takes priority over whatever statistical mechanism may be at work in speech segmentation. In other words, the tracking of transitional probabilities can be boosted by the presence of additional cues that may assist with segmenting the speech stream. More importantly, when such phonotactic cues are violated, listeners are prevented from successful segmentation altogether. Tyler and Cutler (2009) examined the role of prosodic cues in artificial speech segmentation by listeners of differing L1s. Their findings suggests that both language-universal (final vowel lengthening) and language-specific (pitch movement) information affected segmentation by native Dutch, English and French listeners.

The evidence suggests that both TP and phonotactic knowledge likely emerge from a distribution-based learning mechanism. However, there are important differences in terms of where these statistics come from. Using phonotactic knowledge for speech segmentation involves the application of generalized knowledge taken from specific instances (typically assumed to be type frequencies across the lexicon; see Pierrehumbert, 2006) and applying it to potential word forms. Thus, as Finn and Hudson Kam (2008) argue, phonotactic knowledge reflects input to which learners have been exposed over the course of their language experience, tracking TPs does not necessarily reflect such long term knowledge and can be easily manipulated over the short term.

In our third experiment, participants were exposed to an artificial language based upon a “new variety of Spanish” that either contained the palatal alternation or did not. Subsequently, they were tested on their ability to judge which member of a pair of words occurred in the language. We hypothesized that the native Spanish speaking group would benefit from the presence of the palatal alternation to a greater extent than the L2 Spanish group given their familiarity with the allophonic alternation in their native language.

In Experiment 4 we tap into another type of probabilistic knowledge, that of phonotactics. Native

listeners use phonotactic knowledge when segmenting the speech signal and thus benefit from the presence of phonotactically legal sequences when carrying out speech processing tasks. To test whether non-native speakers also benefit from such knowledge, we designed an elision task in which listeners had to strip away a context syllable to recognize the lexical item that remained. We predict that when the final segment of the context syllable favors a particular palatal variant and the following lexical item begins with the expected variant, listeners with greater Spanish experience will recognize the lexical item faster than when phonological expectations are violated. For example, context syllables ending in vowels will favor fricative onsets for the adjacent word while those ending in nasals or laterals will favor affricate-initial words. Again, we varied the onset palatal to follow phonotactic expectations (alternating condition) or not (non-alternating condition).

Some evidence for this prediction comes from Weber and Cutler (2006). They examined how such boundary effects play out in a word spotting experiment with highly proficient L1 German/L2 English and native American English speakers. The stimuli consisted of embedded English or (different condition) German words. They found that accuracy and response latencies were facilitated by boundaries that coincided with word-onset phonotactics for both languages (e.g., “wish” in *yarWISH* versus *plookWISH*) and boundaries that only occurred in English facilitated recognition by both groups, albeit to a lesser extent for the native German speakers.

The experiments reported on here help elucidate how task demands can modulate the performance of L2 listeners. Crucially, by examining the same variant under different task requirements, it is possible to see how information available for one task may or may not be available for another.

Experiment 1: Similarity rating task

In Experiment 1, native Spanish and L2 Spanish listeners heard pairs of nonwords that varied on the position and type of palatal variant and had to rate their similarity. The objective is to determine how language-specific perception affects the similarity judgments of two sounds that exhibit different relations of contrast across listeners’ first and second languages. Based upon previous work showing that the pattern of phonological relations between sounds in a listener’s native language plays a strong role in how listeners rate their similarity (Boomershine et al., 2008; Johnson & Babel, 2010), we predict that the native Spanish listeners will rate minimal pairs of nonwords with the palatal alternants as more similar than the L1 English/L2 Spanish listeners, given the L2 listeners’ limited experience with Spanish and, importantly, the phonemic status of the target sounds in English.

Table 1. *Biographical information.*

Group	Age	Years studying foreign language	Age started studying
Native Spanish (<i>n</i> = 29)	M = 21.3 (0.8)	M = 4.1 (1.1)	M = 10.1
L1 English/L2 Spanish (<i>n</i> = 30)	M = 20.2 (1.2)	M = 4.8 (1.2)	M = 14.5

Participants

Twenty-nine native (Mexican) Spanish speakers (NSS, 17 males, 12 females) participated in the experiment. They were recruited from the Center for Foreign Language Teaching at the National Autonomous University of Mexico (CELE-UNAM). They received \$10.00 for their participation in the experiment. Thirty L1 English/L2 Spanish listeners (L2 Spanish) were recruited from the University of Iowa. They received course credit for their participation. Biographical information on both groups of participants is presented in Table 1.

The two groups were similar in terms of age and years of studying their second language. None of the participants had lived in a country where their second language was spoken for a period longer than two weeks and none interacted in their second language outside the classroom. In Mexico (particularly Mexico City) it is difficult to find university-level students who do not listen to music or watch television in English. The L2 Spanish group did not listen to Spanish music or watch Spanish television outside of their classes. In terms of Spanish varieties, the L2 listeners were typical of college-level learners, in that they had been exposed to a wide variety of different accents over their course of study, with no particular variety being dominant for any of the learners. None expressed previous exposure to Buenos Aires/Uruguayan dialect.

Stimuli

A native female Mexican Spanish speaker recorded the stimuli. She read each word three times and the clearest tokens were spliced out and used for the experimental stimuli. The target stimuli consisted of six CV.CV or six V.CV bisyllabic words. The CV.CV items took the form of CV-*ma* (e.g., [ɟɟama] ~ [ɟama]), where onsets were either [ɟ] or [ɟɟ], combined with the vowels [a o u]. Tokens were produced with stress on the initial syllable. For the V.CV target items, the palatal sounds were combined with three vowels [a o u] (e.g., [uɟa] ~ [uɟja]). For the fillers, the onsets consisted of [s l r], which were combined with the vowels [a o u]. Stress was uniformly produced on the first syllable.

Stimuli were checked to guarantee that target segments in different positions shared the same acoustic properties. To this end, each of the affricate and fricative targets – in initial and medial positions – were measured and compared based upon their phonetic characteristics. For the affricate targets, we noted the amount of frication following the stop release, the overall segment duration and the presence (if any) of a release burst from the stop portion of the affricate. We further measured the intensity of the stop release burst relative to the intensity of the following vowel to factor out the effect of the differences in overall intensity across tokens. Burst intensity (dB) was subtracted from the vowel intensity (Sundara, 2005). A greater intensity difference will result with softer bursts, indicating a weaker release on the stop. For the fricative targets, we measured the segment duration and the intensity compared to the following vowel. In Table 2 we present the differences between the average values for each of the target token segments in onset and medial positions.

As can be seen from these acoustic measurements, the target segments were produced almost identically in onset and medial position. Moreover, as indicated by the presence of a release burst in the affricate segments, lower intensity ratios and shorter durations, we can also safely assume that the affricates were indeed distinct from the fricative targets. In Figure 1 we present examples of the stimuli used in Experiment 1 (and Experiment 2, see below).

Each token was combined into either a “same” or “different” pair taken from the same category, whether CV.CV or V.CV. For the same pairs, different tokens of the same item were presented. For the different pairs, different versions of the alternation were used, whether the affricate or fricative. In Table 3 we present the composition of the experimental trials.

Given the form of the stimuli, some were real words in Spanish.¹ There were a total of 36 trials (24 different, 12 same target word). The interval between each member of the pair was 500 ms.

¹ Some of the trials included real words (e.g., [oɟa] *olla* “pot”, see Table 3). To verify that the lexical status of certain stimuli was not affecting the results, we ran the same statistical analyses without the tokens that were possible words in Spanish and the results were consistent.

Table 2. *Acoustic description of stimuli.*

Segment position	Duration difference (mean)	Intensity ratio (segment: segment)	Token duration (mean)	Frication duration difference (affricate)	Burst intensity difference (affricate)
Within segments/across position					
fricative	5 ms	0.97	463 ms		
[joma]/[oja]		(SD = .025)	(SD = 11)		
affricate	4.7 ms	.98	460.2 ms	3.33 ms	2 dB
[ɟjuma]/[uɟja]			(SD = 9.8)		
Between segments/same position					
onset	22 ms	1.18	8 ms		
[joma]/[ɟjoma]					
medial	4 ms	1.23	4 ms		
[oja]/[oɟja]					

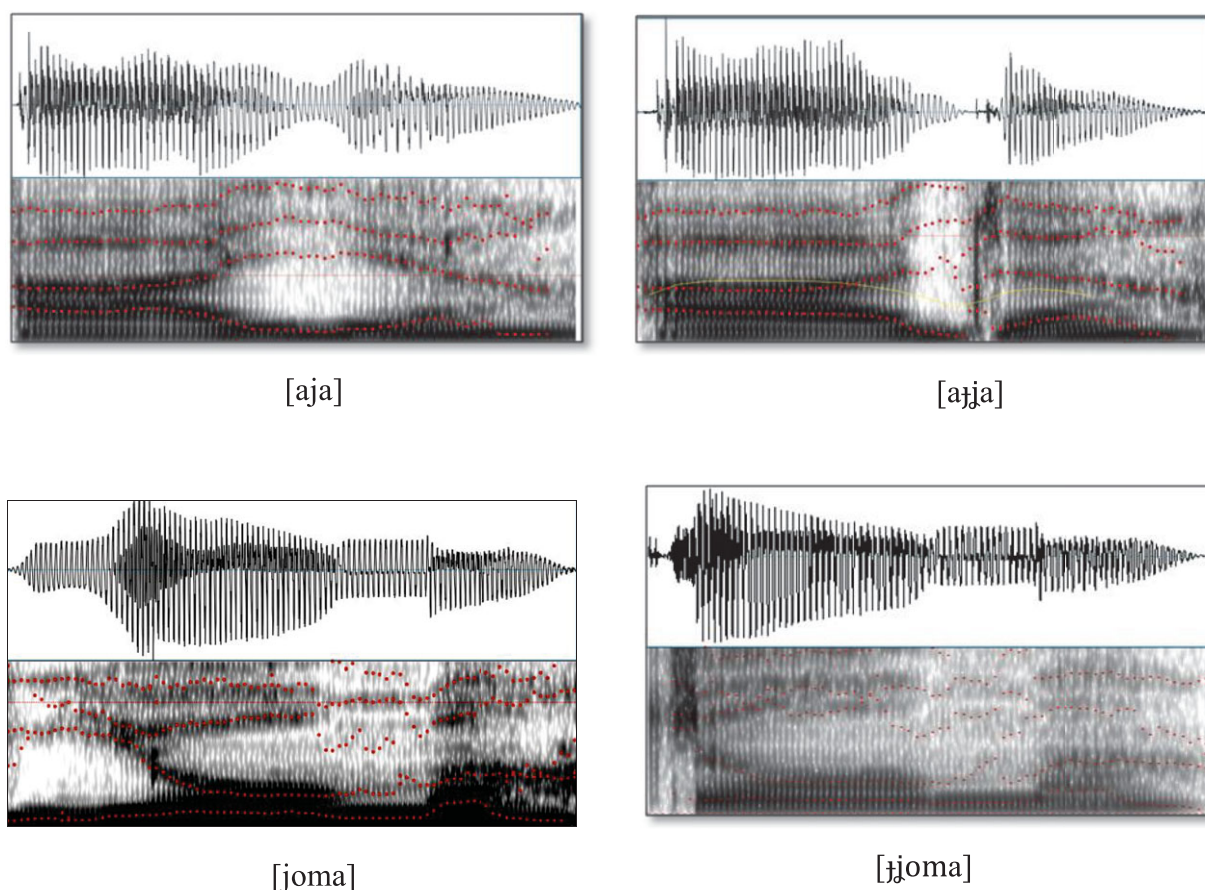


Figure 1. (Colour online) Examples of stimuli used for Experiments 1 and 2.

Procedure

The stimuli were presented in pairs using a Macbook Pro computer running Superlab experimental software. After the trial pair offset, participants had 3500 ms to circle the number on a sheet of paper that corresponded to their judgment of the two words: 1 signified that the words

in the pair were “the same” and 5 signified that the two stimuli in the pair were “very different”. The trials were randomized for each participant. Participants were given three practice trials before beginning. Practice items were not taken from the test trial items.

In an effort to have both groups listening in Spanish, they were told that the items were based upon possible

Table 3. AX rating task trials.

	CV.CV				V.CV			
	Same		Different		Same		Different	
Target pairs	[j]ama*	[j]ama	[j]ama	[j]ama	a[j]a	a[j]a	a[j]a	a[j]a
	[j]oma	[j]oma	[j]oma	[j]oma	o[j]a*	o[j]a	o[j]a	o[j]a
	[j]ama	[j]ama	[j]uma	[j]uma	o[j]a	o[j]a	u[j]a	u[j]a
	[j]uma	[j]uma	[j]ama	[j]ama	u[j]a	u[j]a	a[j]a	a[j]a
Filler pairs			[j]oma	[j]oma			o[j]a	o[j]a
			[j]uma	[j]uma			u[j]a	u[j]a
	[s]ama	[s]ama	[l]ama	[s]ama	o[s]a*	o[s]a	a[l]a*	a[s]a
	[s]oma	[s]oma	[l]oma	[s]oma	o[l]a*	o[l]a	o[l]a	o[s]a
	[l]oma*	[l]oma	[l]oma	[r]oma	u[l]a	u[l]a	o[l]a	o[r]a*
	[r]ama	[r]ama	[l]uma	[r]uma	a[r]a	a[r]a	u[l]a	u[r]a
			[r]ama	[s]ama			a[r]a	a[s]a
			[r]uma	[s]uma*			u[r]a	u[s]a*

* = possible Spanish word

words and combinations of sounds from that language. Their task was to decide how similar the two words were based upon the sounds in each. All communication with participants occurred in Spanish in an effort to guarantee that the L2 Spanish participants were carrying out all tasks in Spanish.

Results

All rhotic items were subsequently dropped because of inconsistent results.²

For the “same” trials, a similar rating pattern emerged for both groups, with 96% of responses corresponding to 1 (or “same”) and the remaining 4% were 2, or “similar”. Given these results, we will only address the “different” trials in the analysis that follows. In Table 4 we present the distribution of the rating scores from Experiment 1.

The data in Table 4 show that for the “different” stimuli, the native Spanish speakers’ scores clustered around two modes: “2” and “4” while the L2 listeners were almost uniform in their rating of “5”. Because the data have non-normal distribution and involve rating scores, we used the Mann-Whitney U-test to determine if there are differences among the rating scores associated with the two groups. The Mann-Whitney U test showed a significant difference

between the means on the rating scores for the two groups: $U = 41016$; exact $p < .001$, two-tailed.

We conducted a second analysis to determine if there were any differences in rating scores for the different types of target trials across the two groups. Again, because the data are not normally distributed, we used the non-parametric Friedman Test to investigate whether there was an effect for position of the allophone (onset vs. medial, CV.CV vs. V.CV). Extensive cross-linguistic research has shown that onset hardening is common in many disparate languages (Kenstowicz, 1994: Basque; Gordon, 1997: Estonian) and there is a general tendency across languages to strengthen consonant articulations in initial position of elements found in the prosodic hierarchy (Cho, 2001; Fougeron & Keating, 1997). This may mean listeners are more attuned to the differences between the two palatal variants in initial position than in medial position. Given this, we might expect a higher proportion of “very different” (5) responses for the CV.CV items than for the medial target items and further predict that these differences will be greater for the native Spanish listeners than for the L2 Spanish listeners. The results show that indeed, the native Spanish listeners are more sensitive to the variant in onset position than the L2 listeners, $\chi^2(4) = 22.6$, $p < .001$. For medial position, there was also a significant difference between groups, although it was not as pronounced as for onset position, $\chi^2(4) = 8.2$, $p < .05$. In order to determine if these differences were significant across positions for each group, we conducted another chi-square test. For the native Spanish listeners, there was a significant difference for position, $\chi^2(2) = 9.3$, $p < .01$, while for the L2 listeners, position was not significant, $\chi^2(2) = 6.2$, $p = .089$.

² The issue that arose with the rhotics was with the native Spanish speakers. Specifically, the stimuli were not uniformly perceivable as either the trill or tap phoneme. Where the rhotic was judged as more tap-like in word initial position, a violation of Spanish phonotactics occurred (only the trill can occur in initial position). The native Spanish listeners commented that it sounded more like a “d” onset than the desired trill.

Table 4. Percentage for each rating score on “different” responses for target item trials (raw total in parentheses, emboldened and underlined segment represents the target sound).

Group	Score (% of total “different” trials for group)										Mean (SD)
	1 (same)	2		3		4		5 (very different)			
NSS											
<i>n</i> = 696	2% (15)	29% (201)		1.6% (13)		62% (431)		6% (41)		3.42 (1.48)	
	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	
(% response for trial type)	1% (7)	1% (8)	8% (55)	21% (146)	0.7% (7)	0.9% (6)	36% (250)	26% (181)	1% (7)	5% (34)	
L2 Span											
<i>n</i> = 720	1.7% (12)	1.25% (9)		1.8% (13)		8.1% (58)		87% (628)		4.78 (0.705)	
	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	<u>C</u> V.ma	<u>V</u> . <u>C</u> V	
(% response for trial type)	0.8% (6)	0.8% (6)	0.5% (4)	0.5% (5)	0.8% (6)	0.9% (7)	3.1% (23)	4.7% (35)	40% (288)	47% (340)	

NSS = native Spanish speakers

Discussion

The results from this experiment show that native Spanish listeners rate the “different” trials as more similar than the L2 Spanish listeners. This suggests that the same input is grouped into separate modes by the different listener groups, based upon perceived similarity. Specifically, the native Spanish listeners demonstrate two modes in their rating of the palatal trials, with 29% of listeners rating them as similar (with a value of 2) and 62% rating them as different (with a value of 4). This result stands in contrast to that obtained from the L2 Spanish listeners, who rated the different trials almost uniformly (89%) with 5, or “very different”.

We further observed significantly different rating scores from the native Spanish listeners when the palatal alternation occurred in word-initial position versus word-medial position. In ONSET position, native Spanish listeners rated the stimuli as “similar” only 8% of the time while “different” was selected 36% of the time, or more than four times as often. However, when the alternation occurred in WORD-MEDIAL position, native Spanish listeners’ ratings were much more evenly split, specifically, “similar” was selected 21% of the time while “different” was selected 26% of the time. Finally, for those who preferred the 2 rating, this preference aligned 89% of the time with stimuli that had high back vowels, which may be due to the high feature of the vowel interacting with the palatal feature of the consonants.

The position-related perceptual rating differences suggest that native Spanish listeners perceive the alternation as more distinct in onset position than in medial position. In other words, the two variants were rated differently in terms of similarity depending upon the syllabic position. This suggests that native Spanish

listeners demonstrate position-specific perceptual patterns when making similarity judgments of this type. Indeed, the judgments of “different” were overwhelmingly more common for the word-onset position, suggesting a higher degree of sensitivity in this position (or, as a reviewer pointed out, it is also possible that there is much less variability in the input in this position, leading to more binary judgments).

In Experiment 2 we use a different task to examine the perception of the same contrast. In an effort to determine if such language-specific effects carried down to lower-level perception, listeners carried out a speeded AX discrimination task. Research examining adult speech perception using rating and speeded AX discrimination has brought mixed results. While Boomershine et al. (2008) found that language-specific listening occurred in both the rating and speeded AX discrimination tasks, Johnson and Babel (2010) found that language-specific listening only occurred for the speeded AX discrimination task when listeners had longer response times (RTs).

Experiment 2: Speeded AX discrimination task

Participants

The same 29 native Spanish and 30 L2 Spanish participants from Experiment 1 participated in Experiment 2. All experimental tasks were counterbalanced across participants and participants never completed Experiments 1 and 2 consecutively (a pseudo-counterbalanced order was used to avoid this, whereby Experiments 1 and 2 bookended the session). The results from five native Spanish speaker participants were discarded because they informed the experimenter after

finishing that they had confused the response buttons for many of their answers. Three L2 Spanish listener results were discarded because more than 50% of their responses were too slow (longer than 500 ms) and four were discarded because they confused the response buttons (again, they informed the experimenter after the experiment was finished). This gave a total of 24 and 22 participants for each group.

Stimuli

The target stimuli took the same form as that used for Experiment 1. There were two important differences, however. First, we added an extra vowel, [e], to the trial sets, giving [a e o u] for the CV.CV trials. We used the consonants [p b r r] and the clusters [bl tr dr kr kl] for the filler trials.³ There were 144 trials in total: 32 targets (16 “same”/“different” for CV.CV and V.CV targets) and 112 fillers.

Procedure

The goal of this experiment was to test auditory-level responses with as little interference from higher-level categories as possible. Thus, the trial pairs were presented with a very short 100 ms interstimulus interval (ISI) and a 500 ms response deadline. Listeners were told they would hear a pair of invented words based on Spanish and they were to determine if the second member of the pair was the same or different from the first member. They were told to respond as quickly and accurately as possible by means of a button on a button box attached to the computer. The experiment was conducted using Superlab experimental software on a Mac computer. If no response was recorded, the program continued to the next trial. Every eight to ten trials a screen appeared reminding listeners that they were to respond as quickly as possible. If they exceeded the 500 ms limit, a screen appeared telling them their response was too slow and that no answer was recorded. Listeners were provided with five practice trials before beginning the experiment. They received feedback on accuracy and reaction time for the practice trials.

Results

Across the two groups, responses which exceeded the 500 ms limit constituted 8.9% of the trials for the native Spanish listeners and 9.3% for the L2 Spanish

³ [tr dr] were only used for “same” trials after submitting the tokens to evaluation by three native Spanish speakers after running the experiment (not part of the study group) who stated that these particular tokens were not clearly audible. Three of the native Spanish speaker participants commented on the fact that they were not sure if they were in fact hearing the voiced or voiceless occlusive.

group. These responses were discarded. We then divided the responses into same/different and subsequently correct/incorrect categories. For the “same” trials, the accuracy rate was 97.3% for the native Spanish listeners and 95.2% for the L2 Spanish listeners. We conducted a one-way ANOVA for group on the proportion of correct responses for the “same” trials and found no significant differences in accuracy rates, $F(1,732) = .81$, $p = .21$. Subsequently we conducted a one-way ANOVA on LogRT latencies for the “same” trials and again found no significant differences between the two groups, $F(1,732) = 1.02$, $p = .19$. Given that the objective of this experiment is to investigate how speeded discrimination affects perception of the palatal variants on the “different” trials, we did not include the response time and latencies for the “same” trials in further statistical analyses. For the “different” trials, the accuracy rates were 91% for the native Spanish listeners and 93% for the L2 Spanish listeners.

We analyzed the data using a three-way mixed ANOVA, with group (native or L2 Spanish) as the between-subjects factor and variant and position as the within-subjects factors. LogRT was the dependent variable. The first member of the trial pair determined the condition for the trial. The results revealed a significant main effect for group, $F(1,44) = 14.2$, $p < .001$. There was also a significant interaction among the factors, $F(2,45) = 4.9$, $p < .01$. In Figure 2 we show the LogRTs for each condition, separated by group.

To examine the interaction between condition and group more fully, we conducted a series of independent *t*-tests on the means for each condition for each language group. Tests of the four a priori hypotheses were conducted using Bonferroni adjusted alpha levels of .0125 per test (.05/4) of each position and variant combination. For the native Spanish group, significant results emerged across the medial affricate-onset affricate pairs, $t(155) = 2.52$, $p = .013$, the medial fricative-onset fricative pairs, $t(155) = 5.1$, $p < .000$ and the medial affricate-medial fricative ($t(155) = 5.14$, $p < .001$). Only the onset affricate-onset fricative ($t(155) = -2.24$, $p = .027$) did not reach significance. For the L2 Spanish listeners, significant differences emerged only for the medial affricate-medial fricative ($t(154) = 4.24$, $p < .001$).

Discussion

The results from the speeded AX discrimination task support previous research (Babel & Johnson, 2010) demonstrating a language effect even in tasks that draw upon auditory representations. Babel and Johnson (2010) found language-specific discrimination of fricatives, where longer response latencies correlated with language-specific discrimination. The results obtained here appear to support this conclusion as well. The native Spanish

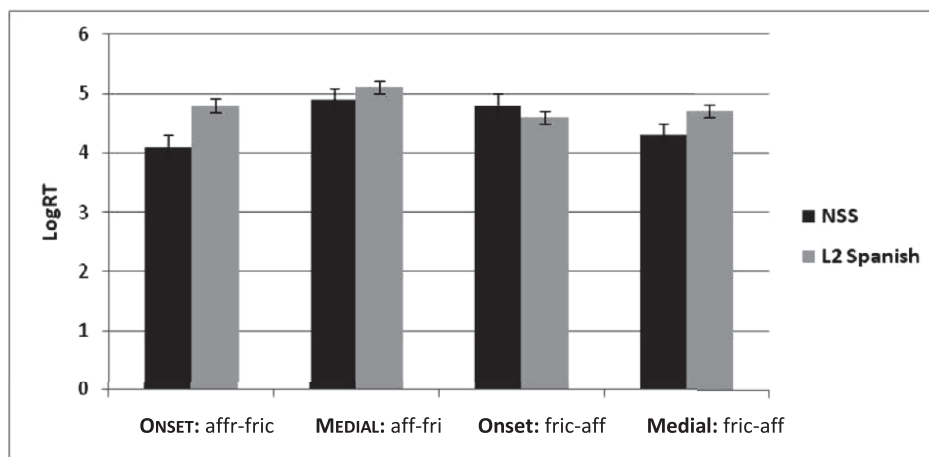


Figure 2. LogRT for speeded AX discrimination task across groups and conditions.

speakers demonstrated significant effects for both position and variant across three of the four pairs (the exception being in onset position), suggesting that this group has stored the variants in a position-sensitive manner that may be close to complementary distribution, in spite of the gradiency present in the linguistic input. (The experimental stimuli were binary, which may have influenced the outcome.) For the L2 listeners, significant effects were reached only for the variants in medial position.

These results are somewhat surprising, given those obtained in Experiment 1 where the native Spanish listeners perceived the two variants as similar-sounding while the learner group did not. One way of accounting for the apparently conflicting results is the nature of the task. When carrying out similarity judgments, L2 listeners heard the two palatal variants as very different. Yet when asked to discriminate between them, as in Experiment 2, the same listeners failed to do so in three of four possible conditions. For the native Spanish speakers, the opposite situation held. When given time to contemplate their answers, the palatal variants were classified as more “similar” than different. When given limited time to distinguish between the two, the native Spanish listeners did so successfully, with responses divided along position.

In Experiment 3 we examine how the two listener groups use the palatal alternation in an artificial speech segmentation task. Given our results thus far, we predict that the native Spanish speakers exposed to an artificial language stream WITH the palatal alternation will be more successful at segmenting word forms than those exposed to an input stream WITHOUT the alternation. The results from Experiment 1 suggest that when given tasks that appeal to higher-level representations, native Spanish speakers rate the sounds as different but not as different as unrelated phonemes. L2 listeners, on the other hand,

rated the palatal variants as “very different”, which may mean that they are not aware of these sounds are tightly linked to certain probabilistic distributions in the Spanish input. If this is true, L2 Spanish listeners will rely upon transitional probabilities as their only cue to segmentation and the availability of the palatal variant as a cue to “word” boundaries will not boost the segmentation accuracy for the L2 learners. They are predicted to be at chance for both input streams.

Experiment 3: Artificial language segmentation task

In an artificial language segmentation task, listeners are exposed to input streams of concatenated syllables with different co-occurrence probabilities. Listeners begin to recognize as potential word forms syllables that co-occur with greater frequency in the input; conversely, syllables that do not co-occur are not grouped together as possible word forms. For example, if a listener has no knowledge of Spanish, but yet hears the word *casa* “house” repeatedly in the speech stream surrounded by other words, she will eventually realize that the syllables *ca+sa* form a coherent word unit in Spanish. Typically, researchers manipulate the transitional probabilities between syllables to indicate possible ‘word’ boundaries. In Experiment 3, we added an extra element and instead of having listeners rely solely upon the co-occurrence probabilities across syllables when segmenting out possible word forms, half the participants were exposed to a speech stream with the palatal alternation present and the other half were exposed to a speech stream with no palatal alternation. Thus, half our listeners had the additional cue of the palatal alternation to boost their segmentation accuracy, provided they were sensitive to the role it plays in indicating a probabilistic word boundary in Spanish. In AltSpan (with the alternation), the transitional probabilities (TPs)

Table 5. Test items for artificial language segmentation task.

		Target words				
Alternating	[ɟja.pi.nu]	[ɟje.fru.so]	[mu.pi.ja]	[mi.fra.je]	[li.su.bro]	[lu.bre.ni]
Non-Alternating	[ja.pi.nu]	[je.fru.so]	[mu.pi.ja]	[mi.fra.je]	[li.su.bro]	[lu.bre.ni]
		Part words				
Alternating	[so.bro.ɟja]	[bre.ni.ɟje]	[ɟja.fra.lu]	[ɟje.mu.po]	[pi.nu.mi]	[fru.so.li]
Non-Alternating	[so.bro.ja]	[bre.ni.je]	[ja.fra.lu]	[je.mu.po]	[pi.nu.mi]	[fru.so.li]

were boosted by the presence of the palatal variants in their expected positions while in NonAltSpan (without the alternation), only the fricative variant occurred and listeners had to rely on TPs alone. Following 12 minutes of exposure to the artificial language speech stream, participants carried out a forced choice lexical decision task in which they had to decide which member of a pair of words was a potential word from the language they heard.

Participants

The same participants took part in Experiment 3 as in Experiments 1 and 2 above.

Stimuli

Table 5 provides the six trisyllabic words that were concatenated to form the artificial speech stream. They took the following forms: ((C)CV.(C)CV.(C)CV). Of the six word forms, two had palatal sounds in initial position, two had palatal sounds in the third syllable and two had no palatal sounds at all. No palatal-initial syllables occurred in the medial syllable. In an artificial speech segmentation task, the “words” are contrasted with “nonwords” based upon the probability of their syllable co-occurrence. For example, if the listener hears the syllables [ɟja] + [pi] + [nu] consistently together in the input stream, they have a transitional probability (i.e., linear co-occurrence probability) of 1.0. On the other hand, if the listener hears the syllables [nu] + [fru] + [li], these syllables never co-occur and the probability drops to zero. Moreover, the probability between of co-occurrence amongst the different word forms themselves was .167, that is, the word forms were combined in such a way that they only had a 1/6 chance of occurring one after the other. All syllables respected the phonotactic patterns of Spanish.

A native Mexican Spanish female speaker recorded each syllable in isolation and then in bisyllabic and in trisyllabic combinations. The dominant $\sigma(\sigma\sigma)$ trochaic stress pattern found in Spanish was maintained. To verify that there were indeed differences between the stressed and unstressed syllables, we conducted independent *t*-tests

on the pitch, intensity and duration of the vowels in the stressed vs. unstressed syllables.⁴ Significant differences emerged in all cases ($p < .05$), confirming that the stressed syllables were in fact phonetically distinct from the unstressed syllables. Since our goal was to have the artificial speech stream resemble real speech as closely as possible and also preserve the transitional information, we had to shift the stress to different syllables for the words not found in the speech stream as compared to the real words. Crucially, no target or non-target word had stressed palatal-initial syllables.

Some of the part-words did not occur in the speech stream, which meant that listeners did not need to be sensitive to TPs or phonotactics to correctly reject these stimuli. However, their status as “new words” should lead to an advantage for the occurring words as potential lexical items heard in the input stream. To address possible effects for this within and between groups, we conducted a mixed ANOVA (part-word status \times group) on the proportion of correct rejections for the non-occurring part-words (vs. the occurring part-words). The results revealed no main effect for non-occurring part-words over the occurring part-words ($M = 0.55$, $F(1,5) = 1.02$, $p = .12$) but did reveal an almost significant difference between groups, $F(1,59) = 32.1$, $p = .061$). There was no part-word status group interaction observed. These results suggest that the presence of some of the part-words in the input stream did not give these items an advantage over the others. Part of the reason for this may lie in the similarity between the vowels in the occurring and non-occurring part-words. The consonant onsets remained consistent and the only the vowels were switched, which may have been perceivable to the native Spanish speakers, who were listening in their “native language” but passed unperceived (or unnoticed) by the L2 listeners.

For the target word forms (high internal syllable co-occurrence probabilities) the average duration was 582 ms. For the non-target words, the average duration was 577 ms. The syllable with the affricate onset ([ɟj]) was 158 ms long and the affricate itself was 39 ms. The fricative-initial syllable was 146 ms long and the fricative

⁴ We averaged the values for the two unstressed syllables.

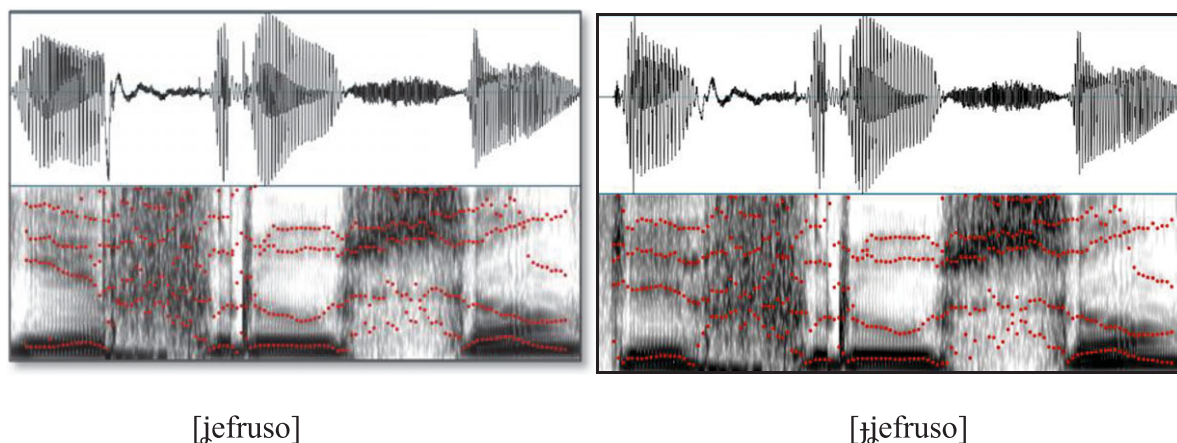


Figure 3. (Colour online) Examples of stimuli used for Experiment 3.

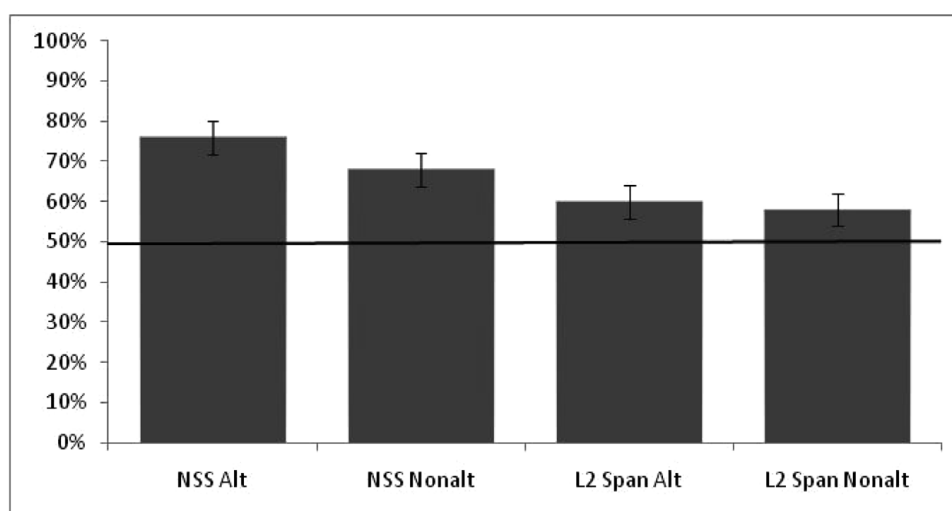


Figure 4. Mean accuracy rates across groups and conditions for artificial speech segmentation task.

itself was 49 ms. In Figure 3 we provide a spectrogram of the syllables with the target palatal segments.

Procedure

Participants were told they were going to listen to a new variety of Spanish and they should simply listen to the speech stream as closely as possible. To avoid possible overanalysis of the input, subjects were given a sheet of paper on which they were encouraged to draw or doodle. Participants wore headphones during presentation and testing. After the 12-minute exposure time, participants carried out the forced choice test. They were told they would hear pairs of words and had to choose which one constituted an example of a possible word from the language they just heard. Responses were indicated by means of pressing a button on a button box and participants were encouraged to guess where they were

not sure. The stimuli were presented in pairs using a Macbook Pro computer running Superlab experimental software. All trials were randomized; half the trials began with the word taken from the artificial language stream.

Results

Figure 4 shows participants' performance as accuracy percentages, broken down by language group and condition.

Performance accuracy for each group exceeded chance (50%). Native Spanish speakers in the alternating condition averaged 76% correct ($SD=1.2$) and the native Spanish speakers in the non-alternating condition averaged 68% correct ($SD=0.98$). For the L2 Spanish group, accuracy reached 60% ($SD=0.6$) for the alternating condition and 58% ($SD=0.44$) for the non-alternating condition.

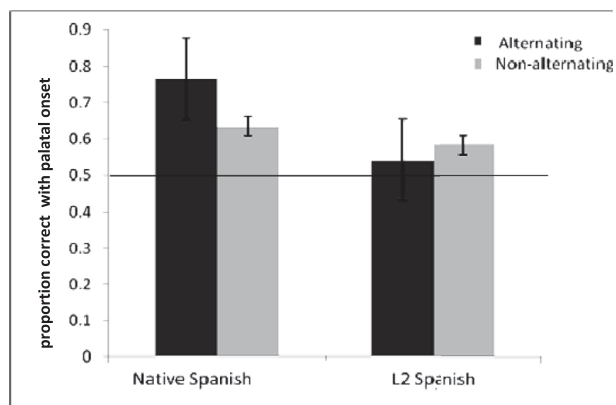


Figure 5. Percentage correct for palatal words.

We carried out a two-way mixed ANOVA with group (native Spanish vs. L2 Spanish) as the between-subjects variable and condition (alternating vs. non-alternating) as the within-subjects variable. Accuracy on the forced choice task was the dependent variable. Results showed a main effect for group, $F(1,58) = 28.1, p < .001$. There was also a main effect for condition, $F(1,58) = 12.3, p < .001$ and an interaction between group and condition, $F(1,55) = 9.1, p < .001$. Given the interaction, we carried out a Tukey's HSD and found that for the native Spanish group, there was a significant difference between the two conditions ($p < .01$) while for the L2 group, there was no such difference.

These results suggest that the presence of the allophonic palatal cue to word onset leads to more accurate identification of words on the forced choice task for the native Spanish listeners. For the L2 listeners, this result did not reach significance, suggesting that the allophonic benefit was weaker, if present at all.

In order to isolate potential effects for the allophone onset, we examined accuracy rates for test items that began with a palatal allophone. We hypothesized that native Spanish speakers exposed to the alternation might have higher accuracy rates for the trials that included only palatal onset items (e.g., [j]efruso (word) ~ [jemupo] (nonword)). For the L2 listeners, we predicted no significant difference. Figure 5 shows participants' accuracy on test trials with palatal words only, broken down by language group and condition.

We subsequently carried out a one-tailed independent samples *t*-tests for each group on the palatal-initial accuracy rates, which revealed a significant difference in mean accuracy across the native Spanish speaker groups in the two different conditions ($t(27) = 6.64, p < .001$). For the L2 listeners, there was no significant difference between the two conditions.

Greater variance was observed for the learners on the affricate-initial words in the alternating condition, as compared to the fricative-initial words in the non-

alternating condition. This suggests that the learners may be moving towards a more native-like pattern of sensitivity to the variants.⁵

Discussion

In Experiment 3, we showed that the extraction of word forms by native Spanish listeners benefits from the presence of the palatal allophonic cue to word onsets. This result is consistent with previous research showing that the TP-tracking mechanism in the input stream is rendered more powerful when combined with language-specific phonotactic information. The L2 Spanish listeners, on the other hand, did not benefit as fully from the presence of the affricate/stop onset.

In the artificial language segmentation task listeners had to track transitional probabilities present in the speech signal AND recognize language-specific allophonic alternations that coincided with the TPs in indicating word boundaries. Thus, language-specific allophonic information reinforced the transitional probabilities, which can be likened to the real-world task of segmenting the speech stream where phonotactic cues and TPs reliably coincide. Nonetheless, only the native Spanish speaker group was helped by the presence of the language-specific allophonic cue to word onsets. This suggests that while listeners do not necessarily have to use both TPs and the allophonic cue to segment the speech stream, the native Spanish speakers alone benefitted from the mutually reinforcing nature of these two cues to segmentation. The L2 listeners do not appear to do so and instead rely primarily upon the TPs. In order to benefit from the presence of the palatal alternation in the segmentation task, L2 listeners must be able to represent this knowledge abstractly and use it to segment the artificial speech stream in an on-line fashion. This suggests that the L2 listeners may not be aware of the distributional information linked to the affricate variant in (probabilistically) indicating word onset in Spanish. To further explore this, we carried out Experiment 4.

Experiment 4: Elision task using phonotactics

Experiment 4 uses an elision task to determine whether words are recognized faster when they occur in a phonotactically favored environment than when they do not. Elision tasks are typically used to evaluate phonological awareness in children learning to read. Participants are given a word and asked to eliminate a particular letter from it and say what word remains. In the present case, listeners heard a nonword and had to strip away a context syllable to identify the remaining word.

⁵ We thank an anonymous reviewer for pointing this out.

This task can be likened to a simplified version of word-spotting, which has been used in other research with more advanced second language learners (see Weber and Cutler, 2006, for German and English).

In Experiment 4, participants were exposed to context syllable+word combinations where the boundary between the context syllable and the word was either phonotactically highly probable or, conversely, highly improbable. For example, the affricate version of the palatal variant (or even the palatal occlusive [j]) occurs after nasal consonants and the lateral [l]. Thus, when listeners hear nonwords such as GLEN+llave “key” in the form [glenjjaβe], they are predicted to strip away, or elide, the context syllable and recognize llave more quickly than if it took the form [glenjaβe]. The nasal and lateral-final context-syllables probabilistically condition the more consonant-like version [jj] of the palatal variant in the following word. Vowel-final context syllables, on the other hand, favor the fricative in the onset.

Such phonotactic restrictions are not inviolable in Spanish, as compared to, say, the well-known inviolable constraint against [s]+obstruent clusters in onset position *skuela [skue.la] vs. escuela “school” [es.kue.la]. Instead, the substitution of an unexpected allophone variant is predicted to disrupt expectations regarding the initial palatal allophone and lead to longer latencies. Nonetheless, the recognition of the word itself should not be impeded.

Participants

The same participants who took part in Experiments 1, 2 and 3 also took part in this experiment. The results from one native Spanish and two L2 Spanish listeners were eliminated because they neglected to register the moment they recognized the embedded word and only wrote it on their sheet of paper.

Stimuli

All items were recorded via a Sennheiser microphone directly onto a PC computer. The speaker was a female native Mexican Spanish speaker, instructed to avoid any clear syllable boundaries in her productions.

Participants were exposed to 60 context syllable + target word combinations and 60 filler items. The 12 target words had a palatal target sound in onset position and were taken from a corpus of the 5000 most common words in Spanish (Davies, 2005) or from the first-year Spanish textbook used by the L1 English/L2 Spanish participants. The 60 filler items were among the 1000 most frequent Spanish words, according to Davies (2005) and, moreover, were taught as vocabulary items in the first eight chapters of the Spanish language textbook used in the university language program followed by the L2

participants. As Weber and Cutler (2006, p. 598) note, word frequency counts may not reflect the experience of L1 and L2 listeners in the same way; however, when the manipulation of interest is carried out within items, this problem is alleviated.

While all five vowels can occur in open syllables in Spanish, coronal consonants are much more frequent in coda position compared to consonants at other places of articulation (with exceptions mostly found in borrowings and learned Latinisms). Yet this tendency is not categorical, as a reviewer notes that labial and velar coda segments are possible; the reviewer lists: cá[p.s]ula “capsule”, ca[p.t]ar “to capture”, a[k.s]ión “action”, é[k.s]ito “success”, a[k.t]o “act”, corre[k.t]o “correct”, pa[k.t]ar “to agree on”, a[ɣ.n]óstico “agnostic”, i[ɣ.n]ición “ignition”, i[ɣ.n]ominia “disgrace”, i[ɣ.n]orar “to ignore” and ma[ɣ.n]animidad “magnanimity”. (To complete the paradigm, voiced labial codas such as a[β.s]urdo “absurd” might be included.)⁶ Within words, in non-velarizing dialects, target nasals in coda position assimilate place of articulation to the following consonant.

In terms of nasal consonants, most varieties of Spanish exhibit neutralization whereby nasals in word-final position are realized as either the alveolar [n] or the velar [ŋ]. Within words, in non-velarizing dialects target nasals in coda position assimilate place of articulation to the following consonant. We did not use any syllables with final [s] because of the high tendency for that segment to undergo either deletion or aspiration across Spanish dialects. In contrast to the nasal segments, [s] can be elided completely in certain varieties.

In Spanish, the palatal segment in word onset or medial position can be represented orthographically by the letter *y* (*yema* “egg yolk”) or, alternatively, by *ll* (*llama* “call”). Another way this sound is represented orthographically is by means of the combination *hiV*, as in *hielo* “ice” or *hierro* “iron”. Three of our test items exhibited this particular orthographic combination. According to Hualde (2005), Spanish speakers typically pronounce the first two orthographic patterns with the affricate when it occurs in the correct phonotactic context (after a pause or after a lateral/nasal segment) but

⁶ However, we would like to point out that the list does not constitute a counterexample to the claims that (a) codas TEND to be coronal, and (b) most exceptions are found in borrowings, as each example above is itself a learned word borrowed from Classical Latin. This is evinced by the maintenance of consonant clusters that were reduced or otherwise modified (via sound change) in the patrimonial lexicon of Spanish: CAPSA [p.s] > *caja* [x] “box”, CAPTĀRE [p.t] > *catar* “to look at”, LACTEM [k.t] > *leche* [tʃ] “milk”, SIGNU [g.n] > *seña* [ɲ] “sign”. Furthermore, such words are subject to synchronic simplification, often undergoing vocalization (*a[ɥ.s]urdo* < *a[β.s]urdo*, *corre[i.t]o* < *corre[k.t]o*; see Piñeros, 2001) or elision (*a[s]ión* < *a[k.s]ión*; Pharies, 2007) in casual speech, demonstrating their vulnerability.

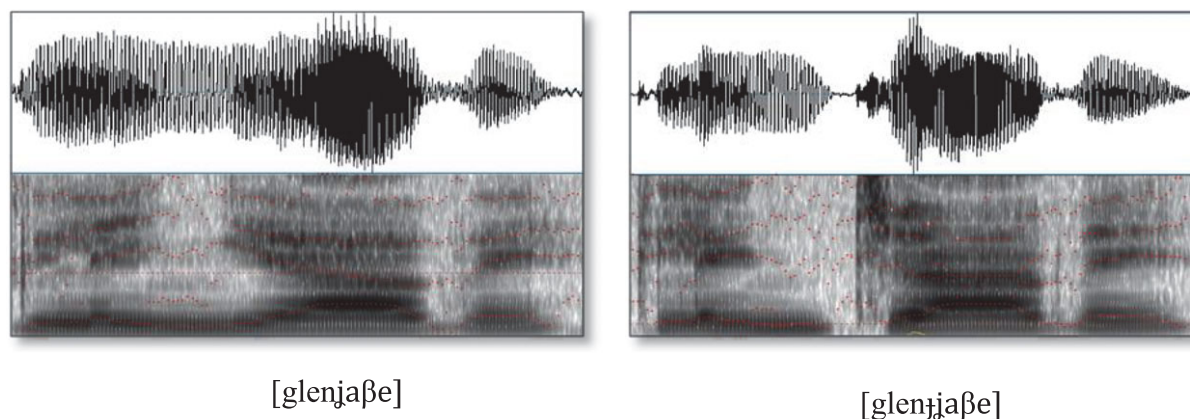


Figure 6. (Colour online) Examples of stimuli used for Experiment 4.

when the orthographic combination *hiV* occurs at the beginning of a word, speakers typically produce it more like the palatal fricative or even a glide, attributable to orthographic effects that manifest phonetically. For the present experiment, this meant that the three test items with *hiV* in initial position could potentially be recognized faster in the non-alternating condition because the palatal segment tends to be pronounced as a fricative rather than the affricate. In Figure 6 we present examples of the stimuli used for Experiment 4.

We selected six context syllable templates (initial syllables) that were subsequently combined with the target words and fillers. Table 6 presents the initial context syllables used with the target words.

We selected bisyllabic trochees, which, when combined with the context syllables, gave the trisyllabic form $\sigma(\sigma\sigma)$. Thus, in addition to whatever phonotactic and allophonic cues are available in the input, listeners could also take advantage of primary stress cues. In Spanish, between 75% and 80% of words follow the trochaic stress pattern (Harris, 1983; Quilis, 1984); specifically, penultimate (medial) stress accounts for 73.52% of trisyllabic words (LEXESP database: Sebastián-Gallés, Martí, Carreiras & Cuetos, 2000). In sum, the trochaic/penultimate stress pattern comprises nearly three-fourths of the Spanish lexicon. For English, Clopper (2002) analyzed tokens from the Hoosier Mental Lexicon (Luce & Pisoni, 1998) and found that three-syllable words in English exhibit primary stress most frequently on either the first or second syllable. However, when the type frequency for both accentual patterns is divided by the token frequency, second syllable stress is more common. Thus, based upon the distribution of stress patterns across the English lexicon, native speakers expect to encounter primary stress most often on the second syllable of three-syllable words, followed by the first syllable. Of the 12 context syllables that occur with the target items, four are vowel-final, four are lateral-

final and four are nasal-final. To avoid possible priming effects, we presented each word only once to the listeners, which was necessary given the relatively small number of possible and appropriate target items.

In this experiment we are testing linear phonotactic knowledge as well as lexical recognition in favorable vs. unfavorable contexts. If listeners recognize the phonotactic information regarding the expected context for each palatal variant and use this to activate their expectation regarding the palatal segment in the onset of the following word, there should be a significant difference between the alternating vs. non-alternating conditions. Using this type of elision task allows us to test the effect for linear phonotactic context and the knowledge listeners have of these expected contextual variants. The predictability of the context syllable in terms of length and boundary location does not interfere with the conclusion that phonotactic knowledge is involved in spotting the words.

We also conducted a small-scale control experiment to guarantee that our stimuli could be identified correctly without the context syllables. Following Weber and Cutler (2006), we presented three native Spanish and ten L1 English/L2 Spanish listeners with twenty-two items, ten of which were taken from the experimental fillers plus the 12 experimental items themselves. For the experimental items, we excised the context syllables and presented half with the fricative onset and half with the stop/affricate onset. The fillers were presented with their context syllables. Listeners were asked to press a button whenever they heard a Spanish word and subsequently write it down on a sheet of paper. The native Spanish speakers reached 100% accuracy (22/22, 12/12 on the target items) and a paired sample *t*-test revealed no significant differences among participants between target items and fillers with the context syllable ($t(11) = 0.82$, $p > .05$) in terms of response latencies; nor were there significant differences between the fricative onset and stop/affricate onset RTs,

Table 6. Context syllables and lexical items for elision task.

Context syllable	<i>ll-</i>	Context syllable	<i>hie-</i>	Context syllable	<i>y-</i>
<i>glen-</i>	<i>llave</i> “key”				
<i>kro-</i>	<i>llego</i> “arrive”	<i>ol-</i>	<i>hierro</i> “steel”		
<i>lon-</i>	<i>llora</i> “cry”				
<i>on-</i>	<i>llanta</i> “tire”				
<i>san-</i>	<i>llenar</i> “fill”	<i>grel-</i>	<i>hierba</i> “grass”	<i>prel-</i>	<i>yema</i> “yolk”
<i>so-</i>	<i>llama</i> “call”				
<i>bra-</i>	<i>lleva</i> “carry”	<i>pral-</i>	<i>hielo</i> “ice”		
<i>se-</i>	<i>lluvia</i> “rain”				

either ($t(11) = 0.71, p > .05$). The L1 English/L2 Spanish listeners reached 90% accuracy (average 20.2/22, 10.8/12 on target items) on the excised target items and 100% correct on the filler items. A paired sample *t*-test revealed significant differences between the excised target items and the embedded filler items, with the latter latencies being longer ($t(11) = 2.31, p < .05$). Given these results on the control task, we can safely assume that both the native Spanish and the L1 English/L2 Spanish groups are able to identify the target items with their context syllables excised and can identify the filler items embedded in their context syllables.

Procedure

The stimuli were presented in pairs using a Macbook Pro computer via Superlab experimental software. All trials were randomized. Participants were told they would hear nonwords in Spanish in which real words were embedded. They were to listen and as soon as they detected the real word, they were to press the corresponding button on the button box. They then had to write the real word down on a sheet of paper provided by the experimenter. Participants had four seconds to respond, after which the following item was presented. They were given four practice items that did not occur among the filler or experimental trials.

Results

We predicted that the different boundaries between the context syllable and target words would have an

effect on how quickly these words are recognized by listeners. Furthermore, we predicted that there would be an interaction between boundary and group. Specifically, following the results in the artificial speech segmentation task, native Spanish speakers exposed to target items with the stop/affricate variant in onset position should have shorter latencies than those exposed to target items with the fricative variant. For the L1 English/L2 Spanish group, on the other hand, we do not expect to find any significant differences in latencies or accuracy rates across the two types of input. In Table 7 we present the means and standard deviations for each group across each condition. RT latencies were calculated from the target word onset.⁷ No significant differences emerged across or within groups for accuracy.

We carried out a three-way mixed ANOVA on the latencies, with group and condition as the between-subjects factors and context syllable as the within-subjects factor. There was a main effect for group (the native Spanish speakers were faster overall than the L2 group, $F(1,55) = 19.2, p < .001$). There was also a main effect for condition, $F(1,55) = 10.2, p < .01$, due to the faster reaction times for both groups on the alternating condition. Finally, there was also a main effect for context syllable, $F(1,55) = 18.3, p < .001$. There was also a significant group, condition and

⁷ A reviewer points out that the context syllables varied in length, which may have affected the RT latencies. However, since the variability occurred across both alternating and non-alternating contexts, possible effects would be the same across both conditions.

Table 7. LogRT means (ms) and SD for elision task conditions.

	Context syllable		Overall mean
	(C)CV- (<i>se-lluvia</i>)	VC[l]/[n]- (C)CV[l]/[n]- (<i>ol-hierro</i>) (<i>on-llanta</i>)	
Native Spanish			
Alternating	2.83	2.77	2.79
Non-Alternating	2.88	2.94	2.93
L2 Spanish			
Alternating	2.86	2.94	2.92
Non-Alternating	2.94	3.1	3.02

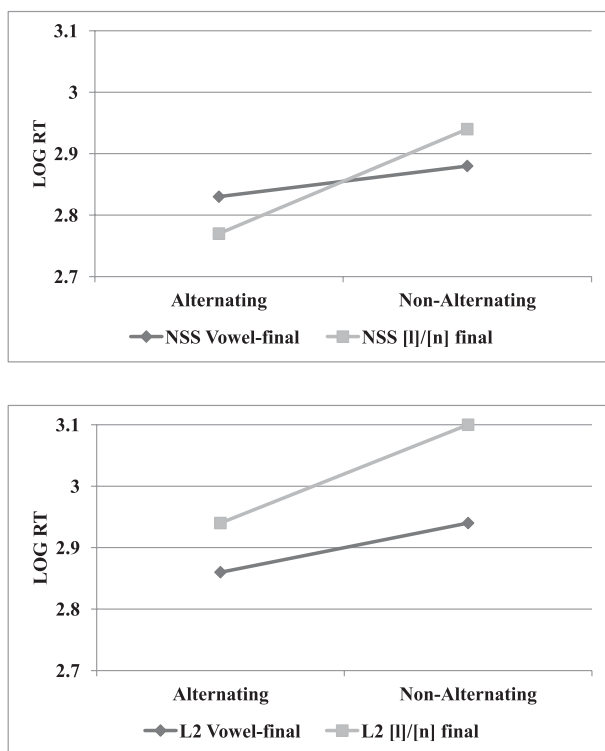


Figure 7. Average LogRT for each group across conditions for Experiment 4.

context-syllable interaction, $F(1,55) = 12.8, p < .01$. In Figure 7 we present graphs of the results.

Given that our hypothesis was related to group differences across the alternating/non-alternating condition, we followed up the significant three-way interaction with a series of *t*-tests for each group, comparing the condition \times context syllable (lateral/nasal or vowel-final). Tests of the three a priori hypotheses were conducted using Bonferroni adjusted alpha levels of .0167 per test (.05/3). For the native Spanish listeners, there was no significant

difference for the vowel boundary context syllable across the alternating/non-alternating condition ($t(27) = 0.780, p = .445$). However, there was a significant difference for the nasal boundary context ($t(27) = 4.96, p < .001$) and an almost significant difference for the lateral context ($t(27) = 3.137, p = .002$). For the L2 Spanish group, a significant difference emerged for the lateral ($t(27) = 2.56, p = .012$) and nasal boundary contexts ($t(27) = 2.16, p = .009$) as well, with no significant difference across the alternating/non-alternating conditions for the vowel boundary items ($t(27) = 2.187, p = .04$).

Discussion

The question addressed by Experiment 4 was whether the type of context syllable boundary would lead listeners to expect a certain palatal variant in the onset of the lexical item and if so, whether this effect would hold across the two groups of listeners. The significant interaction between context syllable and group suggests that our overall prediction held. There is a significant advantage for native Spanish speakers on the elision task when the embedded word occurs in a phonotactically expected context. After comparing within-group differences across the two conditions, we found that the L2 listeners followed the native Spanish speakers with longer latencies in the phonotactically disfavored context.

In this experiment, we examined the effect of linear phonotactic context on an elision task, using the same variant examined in the three previous experiments. Our results generally support those of Weber and Cutler (2006), albeit with a slightly different methodology. Specifically, Weber and Cutler found that phonotactically-expected word onsets were recognized faster than those which violated the phonotactics of the L2 (and L1 in Weber & Cutler). In our case, none of the stimuli violated English phonotactics (given that the closest English categories are not predictably distributed) but did require knowledge of probabilistic nature of the palatal variants in Spanish. Another important difference is that the target sounds in our study occur in probabilistic – not absolute – distributions. In the Weber and Cutler study, the phonotactically expected word onsets were categorically correct or incorrect. Thus, our results add to those of Weber and Cutler by showing that learners are also sensitive to allophonic information in the L2 when carrying out an elision task, suggesting that learners are tracking this information at some level.

The finding that L2 listeners are able to use this information to identify words more quickly in favorable contexts suggests that adult second language learners are sensitive to linear phonotactic restrictions in their second language. However, when these same listeners were exposed to the artificial speech stream in Experiment 3, they did not benefit from this knowledge. Added to this

are the results from Experiments 1 and 2 where the L2 Spanish listeners rated the sounds differently and then were unsuccessful in discriminating between them on a subsequent speeded AX task. In sum, it was only on Experiment 4, the elision task, where the L2 listeners' results aligned with those of the native Spanish listeners. Thus, we are left with the issue of how best to account for the differences across the distinct tasks. We turn to this in the discussion section.

General discussion and conclusions

In this study we presented four experiments that examined how native Spanish and L2 Spanish listeners perceive and use the Spanish palatal obstruent variants across a series of speech tasks. In Experiment 1, we showed that when asked to rate stimuli that contrasted only in the palatal variant, the native Spanish speakers rated the stimuli as more similar than the L2 learners. Moreover, we observed a significant difference between positions for these effects. Intriguingly, the native Spanish listeners rated the palatal variants as more "different" in onset position than in medial position. These results suggest that native language sound categories (both contrastive and non-contrastive) and the distributional information linked to each sound modulate the way listeners judge the similarity of sounds. In Experiment 2, listeners heard pairs of non-words and had to judge whether they were the same or different, under strict time pressure. Again, the results point towards a native-language effect even at low-level phonetic perception (see Babel & Johnson, 2010, for a similar result).

The results from the artificial speech segmentation task suggest that L2 Spanish learners do not benefit from the presence of the palatal alternation when segmenting an input stream. It is possible that L2 listeners are either not aware of the distributional information linked to the palatal alternation in the input or are potentially aware of it but unable to draw upon it when completing the task. When considered in the light of the elision task results, the latter explanation seems to be the most likely: L2 listeners ARE indeed sensitive to the presence of the stop/affricate allophone in onset position when it occurs at the start of a real word in a phonotactically-likely context. Together, these experiments provide a mixed picture in terms of how Spanish L2 listeners use the palatal allophone alternation to carry out tasks in their second language. The results further suggest that the functional role of native-language speech categories influences the perception of second language sounds and task effects play a role in the degree to which these effects play out. The question that naturally falls out from this is what prevents learners from accessing information consistently?

One way of accounting for this may be that learner representations are veridical in the sense of storing

detailed information present in the signal, but this information may not be consistently available across all task conditions, unlike in the case of native speakers (see Shea & Curtin, 2010, 2011, for similar conclusions). This may be the result of a type of attentional filter operating on L2 speech perception that prevents learners from accessing all the information they may have stored. Another possible explanation is that learners have not yet generalized across the input to form robust abstract representations. The L2 learners did not benefit from the palatal alternation when segmenting the artificial speech stream, which suggests that they may not have abstract representations of the alternations. In this task, learners did not have the advantage of phonotactic context to assist them with segmenting the input nor did they have the advantage of real words to trigger lexical recognition, as was the case in Experiment 4.

In speech perception research it is now accepted that distributional learning drives speech category learning. Nonetheless, it is not yet well understood how attention may drive (or not) distributional learning and how previous learning may interfere with it. We showed that adult L2 learners are sensitive to distributional information in the speech stream (Experiment 4) but cannot necessarily access it under all tasks (Experiments 1 and 3). In conclusion, this study demonstrated the complex nature of the interplay between language-specific relations of contrast and task effects.

References

- Aguilar Cuevas, L. (1997). De la vocal a la consonante. Ph.D. dissertation, Universidad de Santiago de Compostela.
- Alvar, M. (ed.) (1996). *Manual de dialectología hispánica: el español de España*. Barcelona: Editorial Ariel.
- Campos-Astorkiza, R. (2012). The phonemes of Spanish. In J.-I. Hualde, A. Olarrea & E. O'Rourke (eds.), *The handbook of Hispanic Linguistics* (vol. 69), pp. 89–100. Malden, MA: Blackwell
- Babel, M., & Johnson, K. (2010). Accessing psycho-acoustic perception and language-specific perception with speech sounds. *Journal of Laboratory Phonology*, 1, 179–205.
- Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony vs. contrast on speech perception. In P. Avery, E. Dresher & K. Rice (eds.), *Contrast in phonology*, pp. 143–172. New York: Mouton de Gruyter.
- Cho, T. (2001). Effects of morpheme boundaries on gestural timing: Evidence from Korean. *Phonetica*, 58, 129–162.
- Clopper, C. (2002). Frequency of stress patterns in English: A computational analysis. *IULC Working Papers Online*, 2. www.indiana.edu/~iulewp/contents.cgi?which=2.
- Davies, M. (2005). *A frequency dictionary of Spanish: Core vocabulary for learners*. New York: Routledge.
- Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A distressing "deafness" in French? *Journal of Memory and Language*, 36, 406–421.

- Finn, A., & Hudson Kam, C. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, 108, 477–499.
- Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domain. *Journal of the Acoustical Society of America*, 106, 3728–3740.
- Gebhart, A., Aslin, R., & Newport, E. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science*, 33, 1087–1116.
- Gordon, M. (1997). Syllable weight and the phonetics/phonology interface. *Berkeley Linguistics Society*, 23, 111–122.
- Harris, J. (1983). *Syllable structure and stress in Spanish: A nonlinear analysis*. Cambridge, MA: MIT Press.
- Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, 119, 3059–3071.
- Hualde, J.-I. (2004). Quasi-phonemic contrasts in Spanish. In V. Chand, A. Kelleher, A. J. Rodríguez & B. Schmeiser (eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, pp. 374–398. Somerville, MA: Cascadilla Press.
- Hualde, J.-I. (2005). *Sounds of Spanish*. New York: Cambridge University Press.
- Jimenez Sabater, M. (1975). *Más datos sobre el español en la República Dominicana*. Santo Domingo: Ediciones Intec.
- Johnson, K., & Babel, M. (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics*, 38, 127–136.
- Kenstowicz, M. (1994). *Base-Identity and Uniform Exponence: Alternatives to cyclicity*. Rutgers Optimality Archive [ROA 103].
- Lipski (1990). Aspects of Ecuadoran vowel reduction. *Hispanic Linguistics*, 4, 1–19.
- Lope Blanch, J. (1989). *La complejidad dialectal de México* (Estudios de lingüística hispanoamericana 1410158). México: Universidad Nacional Autónoma de México.
- Lope Blanch, J. (1996). México. In Alvar (ed.), pp. 81–89.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1–36. Martin
- Butragueño, P. (forthcoming). Estructura del yeísmo en la geografía fónica de México. In R. Gómez & I. Molina (eds.), *El yeísmo en el mundo hispánico*. Madrid: Iberoamericana Verveurt.
- Martínez Celdrán, E. (2008). Some chimeras of traditional Spanish phonetics. In L. Colantoni & J. Steele (eds.), *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, pp. 32–46. Somerville, MA: Cascadilla Press.
- Martínez Celdrán, E., & Fernández Planas, A. (2007). *Manual de fonética Española*. Barcelona: Editorial Ariana.
- Pharies, D. A. (2007). *A brief history of the Spanish language*. Chicago, IL: The University of Chicago Press.
- Pierrehumbert, J. (2006). The statistical basis of an unnatural alternation. In L. Goldstein, D. H. Whalen & C. Best (eds.), *Laboratory phonology 8*, pp. 81–107. Berlin: Mouton de Gruyter.
- Piñeros, C.-E. (2009). *Estructura de los sonidos de español*. Saddle River, NY: Prentice Hall.
- Piñeros, C.-E. (2001). Segment-to-syllable alignment and vocalization in Chilean Spanish. *Lingua*, 111, 163–188.
- Quilis, A. (1984). *Métrica española*. Barcelona: Ariel.
- Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Sebastián-Gallés, N., Martí, M. A., Carreiras, M., & Cuetos, F. (2000). *LEXESP: Léxico informatizado del español*. Barcelona: Ediciones Universitat de Barcelona.
- Shea, C., & Curtin, S. (2010). Discovering the relationship between context and allophones in a second language: Evidence for distribution-based learning. *Studies in Second Language Acquisition*, 32, 581–606.
- Shea, C., & Curtin, S. (2011). Experience, representations and the production of second language allophones. *Second Language Research*, 27, 229–250.
- Sundara, M. (2005). Acoustic-phonetics of coronal stops: A cross-language study of Canadian English & Canadian French. *Journal of the Acoustical Society of America*, 118, 1026–1037.
- Tyler, M., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *Journal of the Acoustical Society of America*, 126, 367–376.
- Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *Journal of the Acoustical Society of America*, 119, 597–607.
- Weiss, D., Gerfen, C., & Mitchel, A. (2009). Speech segmentation in a simulated bilingual environment: A challenge of statistical learning? *Language Learning and Development*, 5, 30–49.