second
language
research

Second Language Research
27(2) 229–250

DOI: 10.1177/0267658310375753
slr.sagepub.com
SAGE

Experience, representations and the production of second language allophones

Christine E Shea
University of Iowa, USA

Suzanne Curtin
University of Calgary, Canada

Abstract
In this study we examined the effect of language experience on the production of second language (L2) allophones. We analysed production data of the Spanish stop–approximant alternation (b d g ~ β ð ɣ) from Low Intermediate and High Intermediate level native English/Spanish L2 speakers and five native Mexican Spanish speakers. This allophonic alternation is conditioned primarily by position in the word and lexical stress. We examined the use of two cues to the alternation – consonant intensity and the presence of a release burst – and analysed how these cues varied in participants' productions in distinct contexts. Results show that the use of these cues differs with experience; that is, learners with greater language experience exhibit cue use that is closer to the native speakers' cue use. Results further suggest that Low Intermediate learners may be using a basic rule for producing the alternation, but that over time shift to a more nuanced production pattern. These results indicate that more experienced learners' ability to use these phonetic cues in a native-like fashion emerges over the course of allophone acquisition.

Keywords
allophones, adult second language phonetics, adult second language phonology, Spanish

1 Introduction

Learning a language involves not only acquiring the contrastive sound categories that form its phonemic inventory but also acquiring allophones, i.e. the non-contrastive sounds that surface in predictable contexts (Crystal, 1997). An important part of allophone acquisition involves determining the correct context for each variant. One way learners might do this

Corresponding author:
Christine E Shea, Department of Spanish and Portuguese, University of Iowa, Iowa City, IA 52242, USA
Email: christine-shea@uiowa.edu

is by attending to the specific, context-dependent cues that characterize each category. For example, an English learner is exposed to input that contains an alveolar sound with aspiration in word initial position (e.g. [$\mathbf{t^h}$ap]) and is also exposed to input that contains a voiceless alveolar sound with no aspiration (e.g. [s**t**ap]). These two sounds share similar acoustic and articulatory characteristics but do not contrast lexical entries in English. In other words, they fulfil many of the typical characteristics of allophones found across the world's languages. Moreover, [t^h] and [t] represent a very particular type of allophonic relationship: complementary distribution. The likelihood of encountering one sound in the phonological environment where the other occurs is close to zero. The phonological environment can include sounds directly adjacent to the sound itself, sounds that occur at a predetermined distance from it, as well as the prosodic structure that directly contains the sound, such as the syllable, the foot or the prosodic word (Hall, 2009). Part of learning an allophonic distribution of this type involves connecting the correct allophone to its phonological environment, or context.[1]

In this article we examine the role of contextual factors in the production of speech sounds, focusing on second language (L2) learners' ability to produce appropriate allophones in the expected context. Further, we explore whether learning an allophonic alternation happens in an across-the-board fashion. That is, if this alternation targets a natural class of speech sounds, will L2 learners treat all members of the class in the same way? Moreover, will the contextual factors that condition the alternation affect learner productions in a uniform fashion?

We consider these questions by examining first language (L1) English speaker's productions of the Spanish stop–approximant allophonic alternation ([b d g ~ β ð ɣ]. While this alternation is traditionally described as stop-voiced spirant allophones (Zampini, 1994; Lléo and Rakow, 2005), recent work has demonstrated that the relationship may be better characterized as involving stops and approximants (Hualde, 2005). Martínez-Celdrán (2004) argues that these sounds are approximants because they do not exhibit turbulent airflow and moreover, [β ð ɣ] have a lower degree of articulatory precision than the spirants or fricatives and they can be close to vowels in their openness or close to stops, depending upon the tension involved (2004: 4). In traditional IPA, the symbols used for these allophones are [β̞ ð̞ ɣ̞], accompanied by the subscript for lowering, which reflects the more open nature of the approximants versus the fricatives or spirants. For the remainder of this article, we will refer to these allophones as approximants but for ease of exposition we will not use the subscript.

In Spanish, the voiced stops /b d̪ g/ alternate with the approximants /β ð ɣ/, conditioned primarily by position in the word and position relative to stress: Stops occur post-pause ([**g**]*ato* 'cat') and tend to co-occur with stress. Approximants, on the other hand, occur intervocalically and do not tend to co-occur with stress (*a*[ɣ]*uila* 'eagle') (see Cole et al., 1999 for [g]; Lavoie, 2001; Ortega-Llebaría, 2003; Martínez-Celdrán, 2004). Examples are provided in (1):

(1) Examples of allophones across contexts:

Word-initial stressed	Word-Initial unstressed	Word-medial stressed	Word-medial unstressed	Phrase-medial
bicho	**gu**sano	a**de**ntro	ca**ba**lgar	la **b**ata
[ˈbit͡ʃo]	[guˈsano]	[aˈð/dentɾo]	[kaβalˈgaɾ]	[laˈβata]
'bug'	'worm'	'inside'	'to trot'	'the housecoat'

The conditioning factors of word position and stress have a variable effect not only on the alternation as a whole but also on the different consonants themselves: More stop-like productions are observed with the bilabial segments than with the velars (Cole et al., 1997; Ortega-Llebaría, 2003). Given this variability, Spanish L2 learners do not receive an equal amount of input for each of the stops. Moreover, this input is tempered by two other factors: where in the word the stop occurs and whether or not the syllable is stressed.

Beyond the input, L2 learners have to contend with their L1 sound categories. We assume that L1 English speakers will have voiced stops as their underlying form when acquiring the approximant allophones in Spanish, which means they must learn to lenite or weaken the voiced stops in the correct contexts. Initially, L1 English / L2 Spanish learners have one sound category with two variants for the alveolar stop.

Spanish stops and approximants share the same orthographic symbols, which are also shared by English. Our learners were exposed to Spanish in the classroom, where the written form of the language accompanies the spoken forms even in the initial stages. Thus, we assume that L1 transfer and orthography will play a determining role in leading the L1 English speakers to assume that the stop is the base, or underlying, form and the approximant is the alternant (Zampini, 1994; Bassetti, 2008).[2]

Typologically, approximants are more marked than stops across the world's languages. In terms of contextual markedness, however, approximants are less marked than stops in intervocalic position (Ohala, 1994). This is due to an effort on the part of the speaker to keep the closure duration short, but still avoid de-voicing the stop. Excessive shortening may lead to an incomplete closure and a spirant or approximant may result. Nonetheless, while intervocalic approximants are contextually less marked, they do not form the ideal onset because of the relative lack of sonority increase (Prince and Smolensky, 1993; Beckman, 1997). Thus, in phonetic terms, approximants are contextually less marked in intervocalic position, but more marked as syllable onsets. Arguably, L2 learners could be more sensitive to the latter restriction than the first. There is evidence from child L1 Spanish acquisition indicating that stops are substituted for approximants much more often than the other way around (Shea and Curtin, in preparation).

The task facing L1 English / L2 Spanish learners involves creating a new, functionally non-contrastive category for the approximants. To do so, learners must track the cues that indicate there are two allophones in their target language, Spanish, and connect these cues to the context in which each alternant occurs. We examined two phonetic cues to the alternation: release bursts (presence vs. absence) and consonant intensity. More stop-like segments will have a release burst and lower intensity than the approximant-like segments, which will have no release burst and higher overall intensity.

We propose that there are (at least) two ways in which learners might carry out this task. The first possibility involves a rule-based phonological system that leads to categorical acquisition patterns (see Chomsky and Halle, 1968). The second involves a more gradual input-based system that leads to gradient, non-categorical acquisition patterns (Pierrehumbert, 2003a; Ellis, 2008). If L2 learners are acquiring a rule, then they should treat all three members involved in this allophonic alternation in the same way. If not, we might see asymmetries depending on the place of articulation of the contrast involved in the alternation. That is, more frequent speech sounds and their alternates might be acquired first.

In the generative phonology tradition, phonological knowledge is posited as a series of rules that operate across minimal, abstract representations of lexical items (see, for example, Chomsky and Halle, 1968) or constraints (e.g. Optimality Theory; Prince and Smolensky, 1993) that operate over possible outputs. Allophones are the product of rule application or constraint interaction. Because they are entirely predictable, allophones are not stored. Only contrastive sounds (i.e. phonemes) are stored in representations, and the lexicon is considered fully separated from the rules and constraints that form the grammatical output (i.e. allophones). Learning is assumed to be categorical and systematic: rules are applied across natural classes in a non-gradient fashion. More recently, research has shown that language users are sensitive to non-categorical aspects of the signal. For example, frequency and fine phonetic details have been shown to affect lexical recognition and production patterns cross-linguistically (Frisch et al., 2004; Dahan et al., 2008). In the present context, evidence for a categorical phonological system would be across-the-board productions of the alternation, with no differences for place of articulation. On the other hand, if differences in place of articulation arise, a more gradient conceptualization of phonology may be required.

II Present study

The goal of this study is to investigate whether L1 English / L2 Spanish learners of different proficiency levels make use of a categorical or gradient phonological system, as shown by their production of allophones. Our theoretical position follows that expressed by Munson et al. (in press), whereby we assume that representations are latent variables and therefore cannot be directly observed. Instead, they can only be inferred from behavioural patterns. Part of the goal of this research is precisely to infer the types of representations created over the course of L2 phonological acquisition and to consider what type of phonological system is required to create them. Our research questions are as follows:

- Nature of representations: Do learner productions provide evidence for rule-based or gradient representations?
- Proficiency and contextual sensitivity: Does proficiency play a role in adult L2 sensitivity to phonological environment?

In terms of the first research question, allophone acquisition provides an excellent testing ground for comparing models of categorical and gradient phonological systems because, arguably, learners could be using either a categorical or gradient phonological system to carry out the learning task. Seeing as the distribution of the stop–approximant alternation is predictable, it is possible to analyse how learner productions are directly affected by the cues in the input and infer how they are using such information over the course of acquisition. In the present context, evidence for a categorical phonological system would be the finding that no differences across place of articulation emerge. This would support a model that allows for phonological encoding of alternations (Chomsky and Halle, 1968). If, on the other hand, learners are using a more gradient system, such differences are predicted to emerge. This would support a model that allows for richer representations that store phonetic details such as place of articulation.

In terms of the second research question, proficiency effects could be realized in (at least) two ways in terms of phonological environment and cue-use. First, it is possible that learners of different proficiency levels show distinct integration patterns and do not produce the allophones in the correct manner for the phonological environment: the presence of release bursts should co-occur with stops in word onset and stressed syllable contexts.

Another possible effect for proficiency could occur at the level of the phonological environment cues themselves: the contextual factors of stress and position could have differential effects on learner production. For example, learners could be more sensitive to position than to stress in their realization of the phonetic cues.

III Experimental design

1 Participants

Three groups of participants took part in this study. One group is classed as Low Intermediate, one as High Intermediate, and the final one is a Native Speaker group. For the Low Intermediate group, five L1 English / L2 Spanish learners were recruited from third semester Spanish university level classes. The High Intermediate participants were recruited from fifth semester Spanish classes. They were paid $10.00 for their participation. Participants filled out an autobiographical questionnaire regarding their experience with Spanish. None had spent more than six weeks in a Spanish-speaking country and none spoke Spanish outside of the classroom context. Two members of the High Intermediate group also spoke French. In order to confirm their placement in either the Low or High Intermediate groups, participants were asked to self-rate their Spanish ability on a scale of 1–9, where one represented 'my ability on my first day of Spanish class' and nine represented 'my Spanish teacher'. Subsequently, participants were recorded taking part in a 5-minute conversation in Spanish with a speaker who has a near-native level of fluency in Spanish.

The Low-Intermediate group had taken two university-level Spanish courses, with a total class time of approximately 67 hours, over eight months of the same academic year. Two had taken one year of Spanish in high school, three years previous to the data collection. The High Intermediate group had taken four university-level Spanish courses, with a total class time of approximately 135 hours, over two academic years. All had taken Spanish for two years or had taken Spanish in high school. In terms of the input they received in their Spanish class, their instructors were Mexico (Mexico City and Guadalajara) and Spain (both from Madrid). These two varieties of Spanish are relatively conservative in their realizations of the stop–approximant alternations and follow the phonological characterization detailed above. Specifically, stops follow nasals and for [d̪], the lateral. Otherwise, approximants are expected intervocalically.

Two native Mexican Spanish-speakers who were unaware of the study's goal listened to the conversation and classified the speakers into two groups, based upon their accent, grammar and speech rate, on a scale of 1(low)–5(high). The ratings coincided with the initial recruitment levels in all but two cases, where one participant was moved to the High Intermediate group and another was moved to the Low Intermediate group.

Table 1 L2 participant biographical data

Group	Age at testing		L1 English participants' self-rating, out of 9		Native Spanish speaker judges, out of 5	
	Average	Range	Average	Range	Average	Range
Low Intermediate	26.6	19–53	2.2	1–3	2.2	1.0–3.0
High Intermediate	22.4	21–24	5.2	4–7	4.1	3.3–4.5

Table 1 presents the result from these two classification tasks. The native Spanish speakers group was composed of five female Mexican Spanish speakers, from the central region of Mexico (Mexico City [2], Jalisco [1], Puebla [2]). They were all living in an English-speaking environment at the time of data collection and all spoke Spanish and English. Four of the five participants reported speaking Spanish at home and at least 50% of the time outside of the home.[3]

2 Materials and equipment

In selecting words for inclusion, we crossed the following factors: consonant (b, d or g), following vowel context (i, a, u), position (initial or medial), and stress (stressed or unstressed), yielding 36 words (see Appendix 1). The word list included real and nonce lexical items. Where the segments of interest were in initial position (50%, 18/36) 14 of the lexical items were bisyllabic. Where the segments of interest were in medial position, lexical items were either three or four syllables in length. The segment of interest never occurred in syllable-final position. Recordings were carried out in a soundproof booth and made directly onto a PowerMac computer (GI A417 soundcard) and a Sennheiser microphone. The microphone was placed into a stand and maintained at a 45 degree angle at all times, approximately 3.5 cm from the speaker's lips. The speech tokens were sampled at a rate of 44.1 Hz with a quantization of 16 bits and saved directly onto the computer's hard drive.

3 Procedure

All communication with the researcher was conducted in Spanish to avoid possible effects for language mode on the learner groups. However, the self-rating questionnaire for the lower-level learners was in English.

Participants were asked to read three lists of the same words, with semi-counterbalanced order, at a moderate pace, using the carrier phrase *Diga ___, por favor* or 'Say ____, please'.[4] Each participant read the same three lists and the third reading was used for analysis in order to counteract possible novelty effects for the lexical items. Novelty effects occur when words are new to the speaker and may result in a slower, more deliberate reading of the lexical item. Even words that exist in Spanish may exhibit novelty effects for the low-level learners.

4 Phonetic analyses

Once recordings were made, all target words were labelled using Praat 5.0 (Boersma and Weenink 2009). Labels were inserted at the following points for each token: consonant onset–offset, CV onset–offset, burst onset–offset (where present) and vowel onset–offset. Both the waveform and the spectrogram were consulted during labelling. The offset of the previous vowel's F2 served as the onset of the following consonant and the onset of the following vowel's F2 served as the offset of the previous consonant (Lavoie, 2001). Where there was doubt, intensity and other formants were also taken into account. Bursts were identified after a visual inspection of the waveform and spectrogram and also labelled for their onsets and offsets. A 15 ms Hamming window was used for analyses. Window size for burst measurements was based upon the duration of the burst itself and thus varied from token to token. Figure 1 provides an example.

The labelling procedure served the purpose of allowing scripts[5] to be run on the sound files, guaranteeing accurate recording of the data and also allowing verification of labelling decisions where required. There were a total number of 36 tokens per speaker.

Recordings were analysed for consonant intensity and the presence of release bursts. One of the main acoustic features associated with stop production is a noise burst at the moment of release (Kent and Read, 2002). The burst is a very brief acoustic event (10–30 ms in duration) and is the manifestation of the initial release of the air pressure behind the constriction for the stop.[6] According to Stevens and Keyser (1989: 90), bursts can be interpreted as an enhancing feature of a stop. Phonologically, bursts are said to be licensed in onsets: they are often missing in syllable codas or in word final position. Thus, the presence of a release burst indicates a stronger manifestation of the stop, and its absence indicates a weaker segment. Given that there is no closure for approximants, there is no release burst. The implementation of the release burst cue was determined by examining the spectrogram and calculating a ratio based upon the number of bursts present/number of possible contexts. There were nine possible contexts for burst production for each context (stress/unstressed × initial/medial).

The other phonetic cue is consonant intensity. In traditional phonological approaches (e.g. Mascaró, 1984) the stop–approximant alternation in Spanish has been characterized as a process of increasing the sonority of the stop, through lenition, when it occurs between two vowels. According to Parker (2002), intensity is the most reliable correlate of phonological sonority, a fact which is also noted by Ladefoged (1975: 219): 'The sonority of a sound is its loudness relative to that of other sounds with the same length, stress, and pitch,' which is based on intensity or the perceived loudness of a sound. Thus, intensity is connected directly, albeit non-linearly, to the loudness of a sound (Raphael et al., 2007). Because intensity can vary across speakers and also across words with different phonemic composition,[7] we used a ratio measurement of consonant intensity/CV intensity.[8] The ratio was calculated as follows:

$$\frac{\text{target consonant intensity (C)}}{\text{target consonant + following vowel intensity (CV)}} = \text{RATIO}$$

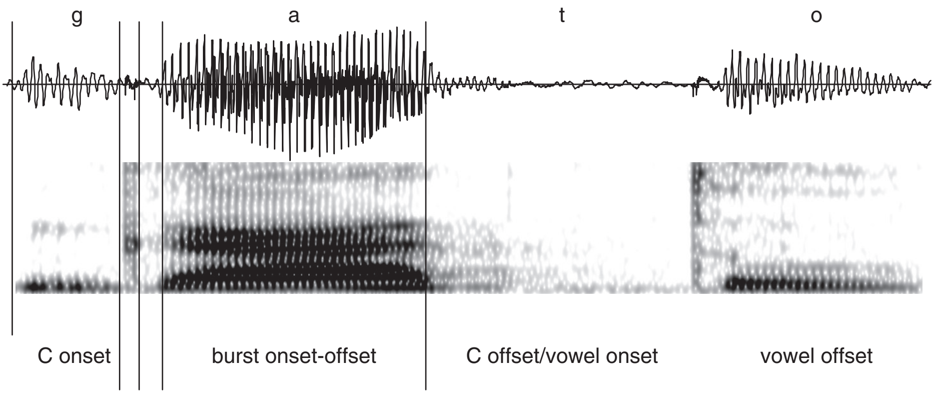

Figure 1 Spectrogram of *gato* 'cat'

Where the target segment has lower intensity, the ratios will be close to zero, indicating the presence of a more stop-like segment. Where the target segment has higher intensity, the ratios will be closer to 1, indicating the presence of a more approximant-like segment.

IV Results

To determine if there were any significant differences between the real and nonce words, a mixed ANOVA was conducted with group as the between-subject factor and word-type (real vs. nonce) as the within-subjects factor. The dependent variable was the average consonant intensity.[9] The main effect for word-type was not significant overall ($F(1,12) = .031$, $p > 0.05$, partial $\eta^2 = 0.003$). This permitted collapsing across word types for subsequent analyses.

Recall that our first objective is to determine whether learning is systematic and categorical or if gradient effects are observed. The second objective is to determine whether proficiency plays a role in sensitivity to phonological environment factors. Thus, we examined which cues (if any) best separate the three groups and how to interpret these dimensions of difference in terms of the phonetic and phonological environment cues. Because we know the native Spanish speakers produce the alternation, their data can serve as a baseline against which to compare the learner groups.

Discriminant Analysis (DA) is the data analysis method that best serves this purpose. DA allows researchers to determine along which dimensions groups differ reliably and how those dimensions can be interpreted (Tabachnik and Fidell, 2007). The DA was run using the two cues in each of the four phonological environments (stressed/unstressed, initial/medial). This gave a total of eight potentially significant predictors. The grouping variables were formed by the three proficiency levels: Low Intermediate, High Intermediate and Native Spanish speakers.[10] In view of the fact that the groups have an *n* of five, each run of the DA could only use four predictors (Tabachnik and Fidell, 2007). As a consequence, multiple DAs were run to determine which predictor variables were most important in separating our three groups.

The relative importance of each predictor variable was determined by their structure correlations, or discriminant loadings, which represent the correlation between the predictors and the discriminant functions (Huberty and Olejnik, 2006). The four predictors with the highest structure r's (all greater than 0.5, $p < 0.05$) were kept. Using these criteria, the following four predictor variables were included in the DA:

- unstressed syllable C–CV intensity ratios;
- medial syllables C–CV intensity ratios;
- unstressed syllable burst ratios;
- medial syllable burst ratios.

These four predictor variables that emerged from the DAs are all related to medial position and unstressed syllables. Table 2 presents the descriptive statistics for the data. There were no missing data nor were there any outliers. The correlations are in the small to moderate range and the equality of variance assumption is not violated (Box test, (F(20, 516.9) = 48.3, p = .277)).

To better determine how the four predictor variables separated the three groups, we examined the two linear discriminant functions (LDFs) which emerged as significant. Table 3 presents these results. Function 1 (Wilks' = 0.001, $p < 0.001$) accounts for 93.3% of the variance found in the data while function 2 (Wilks' = 0.15, $p < 0.001$) accounts for 6.7% of the variance. Function 1 is best defined by C–CV intensity, related to both stress and word position: the intensity of the allophone segments in unstressed (.592) and medial syllable (.408) onsets serve to maximally separate the three groups, with intensity values rising relative to amount of Spanish experience. All three groups are separated maximally by this function. This is consistent with the hypothesis that experience with Spanish will lead to a differentiation in phonetic cue-use between word-initial, stressed syllable context and word-medial, unstressed syllable context. Function 2, on the other hand, loads primarily on the positional predictors. That is, burst ratios in medial position (.804) and C–CV intensity in medial position (.358). Function 2 separates the Native Spanish speakers and the Low Intermediate speakers from the High Intermediate speakers. This can be seen in the two-dimensional plot of group centroids provided in Figure 2.

The discriminant analysis presented in this section provides a general picture of the two constructs separating the three groups. The first function in the DA revealed that consonant intensity ratios in unstressed and medial syllables contributed greatest to group separation. For the second function, position contributed greatest to group separation. Thus, the three groups are best separated by consonant intensity ratios in the first instance and position in the second. These results suggest that the two learner groups implement the phonetic cues to the stop–approximant alternation in a way that differs from the Native Spanish speaker group and also differs from each other. What the DA did not reveal, however, were more precise details regarding inter-group differences for each predictor. To investigate this, a one-way multivariate analysis of variance (one-way MANOVA) was conducted. The four predictors used in the discriminant analysis (C–CV medial, C–CV unstressed, burst medial position, burst unstressed) served as the dependent variables. Group was the independent variable and all tests were conducted at $p < .05$.

Table 2 Means and standard deviations on the dependent variables for the three groups; standard deviations are in brackets

Predictor	Low Intermediate	High Intermediate	Native Spanish speakers
Unstressed: C–CV (1)	.853 (.458)	.925 (.071)	.967 (.040)
Medial: C–CV (2)	.849 (.421)	.919 (.092)	.951 (.044)
Medial: Burst (3)	.906 (.671)	.763 (.049)	.235 (.083)
Unstressed: Burst (4)	.889 (.781)	.711 (.149)	.367 (.165)

Table 3 Results of discriminant analysis for phonetic and phonological environment cues

Predictor	*r*'s for predictor variables: Discriminant functions		Within-groups correlations among predictors matrix		
	Function 1	Function 2	Medial C–CV	Medial burst	Unstressed burst
Unstressed C–CV	.592	n.s.	.47	–.48	.14
Medial C–CV	.408	.358	–	.54	–.51
Medial burst	n.s.	.804	–	–	–.16
Unstressed burst (4)	n.s.	n.s.	–	–	–

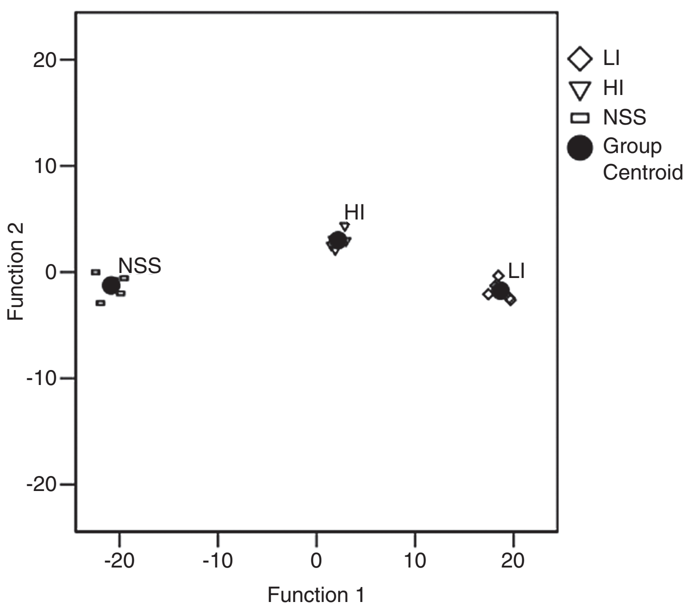

Figure 2 Plot of group centroids

The results for the multivariate test show that we can safely reject the hypothesis that, overall, the means for the dependent variables are the same for the three groups (Wilkes's Lambda = 0.002, significant at $F(4,8) = 103.25$, $p < 0.001$). The multivariate $\eta^2 = .88$, indicating that 88% of the variance of the dependent variables is associated with the group factor. Means and standard deviations are presented in Tables 2 and 3 above. Figures 3 and 4 present the means for the two dependent variables related to C–CV intensity ratios and the two variables related to burst ratios.

The univariate results on the four dependent variables were all significant across all four groups: C–CV medial position ($F(2,12) = 332.65$, $p < 0.005$, $\eta^2 = .78$); C–CV unstressed ($F(2,12) = 690.92$, $\eta^2 = .69$, $p < 0.01$); burst unstressed ($F(2, 12) = 20.153$), $\eta^2 = .77$, $p < 0.05$); burst medial ($F(2,12) = 135.15$, $p < 0.05$, $\eta^2 = .61$). To determine if there were any significant differences between the groups, we conducted post hoc analyses to the univariate ANOVA for the four dependent variables. Tukey's pairwise comparison revealed that the Native Spanish speaker group had significantly different mean scores on all four dependent variables in comparison with the other two groups (all *p*s < 0.05). The Low Intermediate and High Intermediate pairwise comparisons were significant for all dependent variables except for burst unstressed ($p = .126$).

Conjointly, the results from the DA and MANOVA demonstrate that proficiency affects sensitivity to the contextual factors of stress and position. The three groups produce significantly different cue values overall and across the four variables that serve to best distinguish between them on the DA analysis. They further suggest that learners demonstrate non-systematic learning effects, given that the two conditioning factors affected the learners of different levels in different ways.

The question remains, however, whether the non-systematic effects occur across different places of articulation. If speakers are applying a systematic rule to the production of the stop–approximant alternation, such a rule would target a natural class, in phonological terms. Therefore, if learners are applying an abstract rule, there should be little, if any, significant differences across places of articulation. On the other hand, if speakers are drawing upon stored phonetic details when executing the articulatory plan for a specific sound, we expect differences across the three places of articulation.

In order to examine this, we conducted a two-factor mixed ANOVA on the C–CV intensity ratios, with context (stressed, unstressed, initial, medial) as the within-subjects variable and place of articulation as the between subjects variable. We analysed each group separately, since we were not interested in whether between-group differences exist but rather whether differences exist across places of articulation within the groups. Thus, we had a total of 60 tokens for each run of the ANOVA (three places of articulation, four contexts, five cases). Figures 5–7 show the differences in means among the consonants in each of the four contexts, for the three groups.

The results for the Low Intermediate group demonstrate a significant main effect for context ($F(3,36) = 17.645$, $p < 0.001$) but not for consonant ($F(2,12) = .107$, $p > 0.05$). Pairwise comparisons (Tukey's, $p < 0.05$) revealed that this was due to significant differences between initial ($M = .83$) and medial contexts ($M = .87$). These results suggest that the Low Intermediate group productions are affected by context but not by place of articulation, indicating a systematic acquisition pattern has emerged for this group.

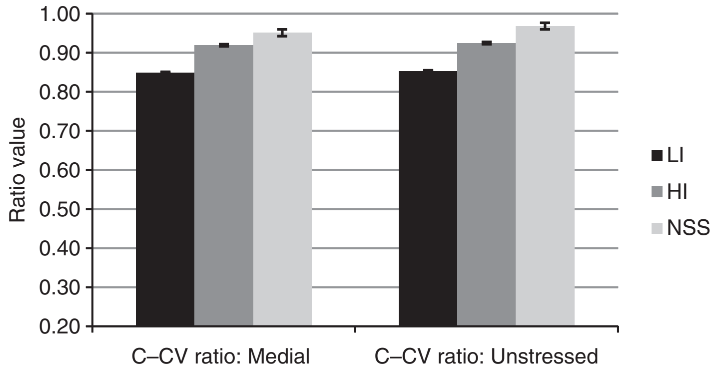

Figure 3 C–CV ratio values for the MANOVA dependent variables

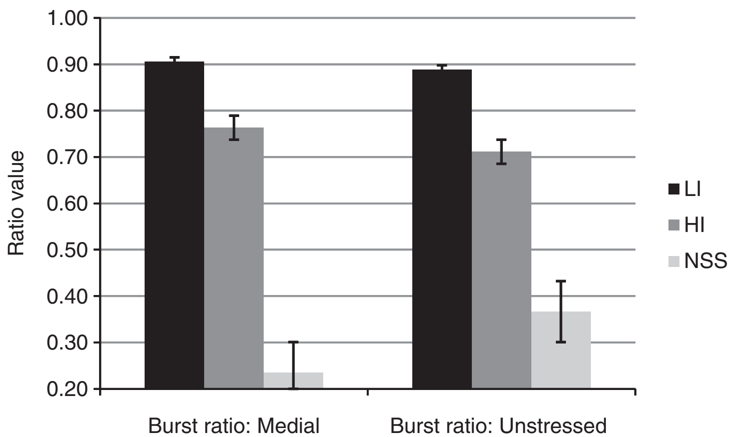

Figure 4 Burst intensity ratio values for the MANOVA dependent variables

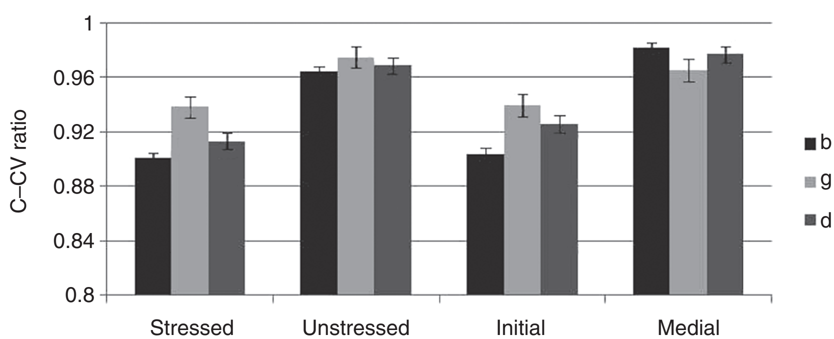

Figure 5 Consonant by context values for Native Spanish Speakers

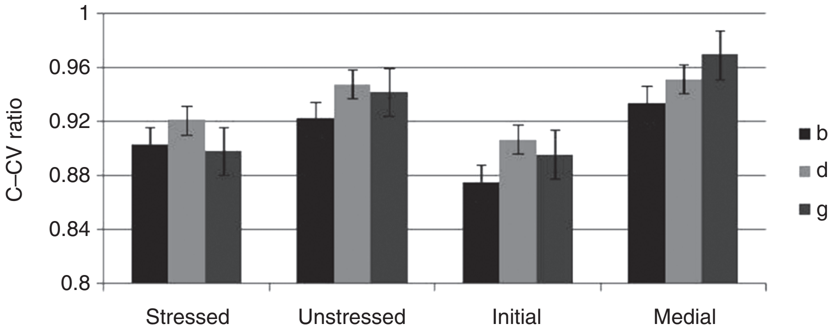

Figure 6 Consonant by context values for High Intermediate speakers

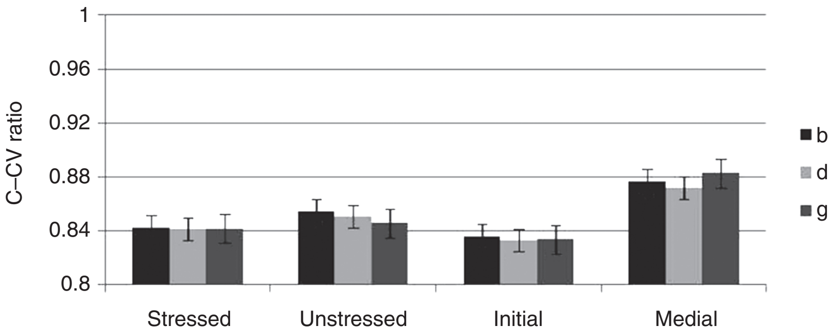

Figure 7 Consonant by context values for Low Intermediate speakers

For the High Intermediate group, there were main effects for context ($F(3,36) = 58.96$, $p < 0.001$) and consonant ($F(2,12) = 13.8$, $p < 0.001$). A significant interaction between context and consonant also occurred ($F(6,36) = 2.83$, $p < 0.05$). Subsequent post hoc tests revealed significant differences amongst b and d/g ($p < 0.001$). Thus, it appears that the High Intermediate group productions demonstrate a more gradient pattern than those of the Low Intermediate group.

Finally, the Native Spanish speaker group productions showed a main effect for context ($F(3,36) = 68.5$, $p < 0.001$) and consonant ($F(2, 12) = 5.03$, $p < 0.05$). There was a significant interaction between the two factors as well ($F(6,36) = 3.71$, $p < 0.01$). Post hoc tests revealed significant differences between b and g ($p < 0.05$).

The results from this section indicate that gradiency in productions across context and consonant emerges with more Spanish experience. The Low Intermediate group may be applying a rule along the lines of 'b, d and g become softer' (i.e. more lenited/more vowel-like) when in the middle of the word. Because there were no significant effects across the places of articulation, we can assume that this is due to the systematic effect of categorical and/or explicit learning at this early stage. As speakers gain experience, their productions become less categorical and more gradient. At the beginning stages, learners may be applying a rule to the natural class of voiced stops and only with more experience do they begin to differentiate across the places of articulation. One way to explain this is that learner representations actually shift over the course of acquisition. Another possibility is that representations remain consistent but the way in which learners access the information they contain is subject to developmentally-dependent modulation. These possibilities are explored in more detail below.

V General discussion and conclusions

The results from this study show that the answer to our first research question – whether proficiency affects sensitivity to contextual factors in adult L2 allophone acquisition – is affirmative. The results from the Discriminant Analysis revealed two significant functions separating the three groups. Function 1 loaded primarily on the consonant intensity phonetic cue, in medial and unstressed position. Function 2 loaded primarily on the medial position phonetic environment cue, for both release burst and consonant intensity. Both significant functions that maximally separate the groups are associated with cues that differentiate the phonological environment of approximants from that of stops. They show that significant differences exist across the three groups for the implementation of the contextual factors of stress and position.

Our second research question was related to the nature of the phonological system and learner knowledge. Specifically, we hypothesized that learning an allophonic alternation could involve either categorical or gradient knowledge. The evidence provided here supports more gradient knowledge, albeit with certain caveats.

The results from the two-way ANOVA for place of articulation and context indicate that detailed phonetic knowledge is also stored, as shown by the differences across the places of articulation in the more experienced groups' productions. Crucially, this detailed phonetic knowledge does not emerge in learners of lower proficiency. The High Intermediate group's productions revealed an interaction between place of articulation and context. These results support the hypothesis that experience with a language is required in order for such subtle effects to emerge in learner productions. Learners with less experience did not produce the fine-grained differences across place of articulation and context that were observed in the productions of the High Intermediate learners.

We propose that the nature of L2 classroom learning may play a role in accounting for these results. It is quite common in the Spanish second language classroom for instructors to mention that *b*, *d* and *g* become 'softer' when they occur between vowels. In fact, the textbook used by the Low Intermediate learners who participated in this study mentions this rule in an explicit manner, which may explain why their productions were most influenced by position. As for the more proficient learners, they may still have the categorical pattern but it has been enhanced and rendered more gradient by increased amounts of experience.

There has been a great deal of research in adult L2 acquisition on the role of explicit vs. implicit learning, most of which has concentrated on the acquisition of morphosyntax. In general, this research suggests that teaching explicit rules to adult learners can lead to faster integration of these rules in production and comprehension. However, the rule must fulfil certain characteristics – for example, it must be relatively simple and transparent in its application – in order for learners to benefit (for a general discussion of this, see Ellis, 2008, amongst others). Given the results observed here, this may hold for phonological acquisition as well. Explicit instruction may lead to categorical effects but implicit learning may be required in order for finer-grained phonetic differences to emerge, such as place of articulation effects. These distinctions may only emerge once the speaker has had experience with the language and can draw upon sufficiently robust representations (Pierrehumbert, 2003a; Ellis, 2008).

According to models of L2 speech acquisition, a key step in any sort of perceptual learning is the realization that differences exist between the L1 and L2 categories, required in order to initiate the acquisition process (Flege, 1995; Best and Tyler, 2007). Again, this can occur implicitly or explicitly.

When speaking of sound category acquisition, it is often difficult to draw a clear line between implicit and explicit learning, given that some sort of awareness is required for the creation of a new category. A difference can be made, however, at the level of explicit instruction vs. more naturalistic acquisition. For example, learners could be explicitly taught that a particular phonological contrast exists in their target language or a particular phonological process occurs across a natural class of sounds. Orthography is another explicit cue to sound category distinctions.[11] Semantic and orthographic contrasts have been shown to assist L2 learners with lexically-based categorization (Cutler et al., 2006; Escudero et al., 2008). In Spanish and English, the allophones are represented by the same orthographic symbols,[12] which may impede the formation of separate phonetic categories for each allophone. L1 English speakers are exposed to the orthographic symbols *b*, *d* and *g*, and associate them with their phonetic/phonological equivalents in English, which are voiced stops. In order to overcome this automatic response, adult learners may initially employ a rule. Indeed, the results seem to suggest that learners shift from an abstract, categorical 'rule' at the early stages of acquisition – which may be the result of explicit classroom instruction – to an implicit mechanism that can track detailed phonetic information across places of articulation. The difficulty with this explanation, however, is the incompatibility of the assumptions regarding the phonological system. We would require two different mechanisms to account for the differences between the two groups and an explanation for how and why they would shift between them.

As an alternative, we propose that all learners – regardless of proficiency level – use the same mechanism and create the same types of representations. However, not all the information that is stored in these representations will be consistently available to all learners, nor will the representations themselves be equally robust. For example, the Low Intermediate group could be abstracting across representations that do not support place of articulation details. In other words, these learners could be accessing information related to position only.

The High Intermediate learners, on the other hand, may be using different levels of information in their productions of the stop–approximant alternation, information that allows them to carry out abstractions that could include place of articulation details. This explanation can also account for why we did not observe differences between the real and nonce words on this production task. The Low Intermediate learners use levels of information in their productions that include positional details, allowing them to abstract from known sublexical patterns (i.e. 'soften the stops in word medial position') to new lexical items. The High Intermediate learners can use this positional information level and also place of articulation information. These learners have stored information regarding sublexical patterns in Spanish that allows them to support generalizations to new words. The High Intermediate group's additional experience means more detailed, robust representations can be drawn upon when carrying out the articulatory plan. Thus, learners of different proficiency levels access different information over the course of perception and production.

Further support for the fact that learners of different proficiency levels access different information was shown in the DA results. The three groups are separated along both position and stress environment cues, suggesting that learners are storing this information and subsequently drawing upon it. However, proficiency will play a key role in precisely how this information is implemented in production. Learners at the early stages do not connect the phonological environment factors of stress and position to the phonetic cues for the approximant and instead produce similar phonetic cue values across the four contexts.

The question that remains is what type of model might best account for the data presented here? We require a framework that can explain how learners track and implement phonetic and contextual cues over the course of allophone acquisition and how experience plays a role in this process. The PRIMIR framework (Processing Rich Information from Multidimensional Interaction Representations; Werker and Curtin, 2005; Curtin et al., forthcoming) provides a developmentally oriented account of how this might be possible. PRIMIR (2005) is designed to account for how infants begin to acquire the sound system of their target language(s) and organize the input into representations. Representations in PRIMIR are exemplar-based and include information from three interactive planes. The first plane organizes sound structure on a pre-lexical dimension called the General Perceptual Plane, the outcome of which is language-specific phonetic categories. These representations interact with the Word Form Plane, which organizes and represents extracted word forms. Words cluster together in multidimensional lexical neighbourhoods based upon similarities in phonetic features. Generalizations across these clusters eventually lead to the formation of phoneme categories, which are represented in the Phonemic Plane.[13] The exemplar-based representations that form are sensitive to context – segments in word-initial position will cluster with similar positionally-occurring segments. Context sensitivity to word-onset position has been demonstrated in infants as young as 9 months (Jusczyk et al., 1999), and Zamuner (2006) found that 10-month-olds could not discriminate word-final contrasts that they could discriminate when in word-initial position.

Exemplar-based models (see, e.g., Goldinger 1996; Johnson, 1997, 2007; Bybee 2000, 2001b, 2003; Pierrehumbert 2001a, 2001b, 2003a, 2003b) of phonological and phonetic knowledge assume that all information found in the input is stored in the multidimensional phonetic space. Grammar emerges as a consequence of generalizations across these stored exemplars once there is a large cluster of similar exemplars that can be identified as a category (Pierrehumbert, 2002, 2003; Hall, 2009). Categories emerge when the connections amongst certain exemplars are stronger than the connections to other exemplars, which may be the consequence of higher-level, top-down factors, such as spelling or lexical knowledge (Werker and Curtin, 2005). Allophones will share more connections in the multidimensional space than phonemes (Johnson, 2007; Hall, 2009), and frequency will necessarily play a strong role in this process, as more frequently encountered items will coalesce into more robust representations. The more often listeners hear a word the more entrenched that word becomes as does its sub-lexical parts (Pierrehumbert, 2001a; Munson et al., 2005). Exemplar approaches can account for the types of positional effects that are the topic of this article and can also account for positional asymmetries that have been observed in adult L2 allophone acquisition.

While all information is stored in detailed, rich, exemplar representations, not all information is available to the learner under all conditions, due to the interaction of three dynamic filters that operate to limit learners' access to the stored information. These filters include task effects, developmental level and natural biases (for example, relative articulatory difficulty). Collectively, they serve to highlight certain information in the input and to modify or even prevent other information from being accessed by the learner. For example, it has been shown that task effects are active in infant speech acquisition: Infants will use certain information at particular developmental stages and under certain task demands and ignore other information under different circumstances (see Fennell and Werker, 2003; Werker et al., 2002).

While PRIMIR was developed to understand how infants understand and process speech, it also provides a ready account for much of the data presented here. For example, our learners showed evidence of tracking the phonetic cues in the input and recognizing their co-occurrence with the phonological environment factors of stress and position. This is predicted by PRIMIR's assumption that contextual information is stored in learner representations. Indeed, there were clear effects for developmental level in our data: the Low Intermediate group produced the phonetic cues in a significantly different manner than the other three groups in terms of the phonological environment in which they occurred. However, in order to fully account for our data, we propose that in addition to the three dynamic filters that are posited in PRIMIR for infant speech development, there is an L1 filter that functions for adult L2 learners. This filter also directs learner attention to certain elements of the input and away from others and, crucially, determines what information adult L2 learners can actually use when carrying out perception and production tasks in their target language. Moreover, there must be some sort of mechanism that allows learners who are acquiring a second language to keep their linguistic input separate and permit the tracking of distributional information across two sets of input statistics.

In a recent bilingual extension to PRIMIR, Curtin et al. (forthcoming) propose that learners use a comparison and contrast mechanism that allows them to maintain the input information for each language separate. For example, in classroom contexts, L2 learners are aware which language is being used and can use this knowledge to initiate a separate statistical tracking mechanism. In a recent study, Weiss et al. (2009) showed that adult learners of an artificial language were capable of forming multiple representations based upon cues that separated different input streams. When their learners were given explicit cues to the existence of different speech streams, their learning increased significantly. This can be likened to the situation for adult L2 learners, who must encapsulate the distributional information from their target language from that which they have stored for their L1. PRIMIR's comparison and contrast mechanism facilitates this process.

To conclude, the complementary distribution of the stop–approximant alternation in Spanish allowed us to examine how learners are generalizing across the information they are exposed to over the time course of acquisition. Adult L2 learners are capable of producing allophones in their second language and doing so requires representations capable of storing rich phonetic detail. In our study, we showed that phonological environment effects interact with learners' developmental level and affect the production of phonetic cues in distinct ways. By appealing to exemplar-based representations and learning filters that allow for differential access to information stored within these exemplars, it is possible to account not only for allophonic learning but also for learning of native and non-native sound categories in general.

Notes

1 Complementary distribution is often cited as a necessary but not sufficient condition for an allophonic relationship to exist (Crystal, 1997). In addition, allophones also generally share certain acoustic and/or articulatory features. For example, in English the sounds /h/ and /ŋ/ occur in complementary distribution; /h/ only occurs in syllable-initial position and /ŋ/ in syllable-final position, but no native speaker of English would ever consider these sounds to be allophones in the same way as [t] and [tʰ].

2 In the case of learners who are not exposed to orthographic input – i.e. only auditory input – we predict that they would also recognize these segments as allophones because of their acoustic and articulatory similarities. However, since this study is not specifically examining phonetic categorization, we will not directly address these points.

3 While we acknowledge that five speakers per group is a relatively small *n*, there was an extensive amount of data collected from each participant, allowing for a full picture of the phenomena under investigation.

4 The final vowel in *diga* may have influenced the production of the following stop-initial word. However, if true, this influence is expected to be in the direction of more approximant-like segments, running counter to our hypothesis that speakers would produce stops in post-pause position. Therefore, if the preceding vowel had an effect on the following stop, its effect would run counter to our predictions.

5 The authors would like to thank Titia Benders for writing the Praat scripts.

6 Burst intensity values are an acoustic cue to place of articulation (Raphael, Harris and Borden, 2007: 150). For the labial stops /p/ and /b/ the bursts have low frequencies, while for the alveolar stops these frequencies are high. Velar stops exhibit more variability in their burst frequency, linked closely to the F2 frequency of the following vowel.

7 Intensity also varies across phonemes. However, given that the segments of interest form a natural class we assume that inherent intensity will not vary greatly across the three segments.

8 Intensity is measured in dB, which are on a logarithmic scale. In order to calculate ratios using logarithmic values, normally one value is subtracted from the other. Given that the objective of this study is to compare productions of cues across proficiency levels, we deemed a pure ratio value sufficient.

9 We selected the consonant intensity variable because the burst ratio values were generally either very low (for the word medial positions, where there were few release bursts produced) or very high (i.e. for word-initial position, where there were a high number of release bursts). Thus, an average score for these groups would not have been indicative of their variability.

10 In order to guarantee that each participant only contributed one score to each variable and thus ensure independent error effects (Max and Onghena, 1999), an average for each cue in each context was calculated. For example, to calculate the C–CV intensity and burst production values for the phonological environment of stressed syllables, all occurrences of the segments for each phonetic cue in stressed syllables were counted, regardless of their position in the word. To calculate the C–CV intensity and burst production for word medial position, all occurrences of the segments in word medial position were counted, regardless of whether the syllable was stressed or not. Again, the creation of these variables ensured that each subject contributed only one score per context.

11 For additional evidence that orthography plays a role in L2 acquisition, see recent work by Bassetti (2006; 2008).

12 In Spanish, the orthographic symbols *b* and *v* are realized in the same manner phonetically, and it is claimed that phonologically they also share a representation. None of the target words had the letter *v* in them, so this was not relevant to the present analysis.

13 This is not meant to imply that the planes respect a strict hierarchy; that is, a plane need not be fully developed before the emergence of another plane. There is interaction among planes and feedback from one to another. Additionally, the planes do not have to be fully adult-like in their representations in order to influence subsequent development (Werker and Curtin, 2005).

References

Bassetti B (2006) Orthographic conventions and phonological representations in learners of Chinese as a Foreign Language. *Written Language and Literacy* 9: 95–114.

Bassetti B (2008) Orthographic input and second language phonology. In: Piske T and Young-Scholten M (eds) *Input matters in SLA*. Clevedon: Multilingual Matters, 191–206.

Beckman J (1997) Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14: 1–46.

Best C and Tyler M (2007) Nonnative and second-language speech perception: Commonalities and complementarities. In: Bohn O-S and Munro M (eds) *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins, 13–34.

Boersma P and Weenink D (2009) *Praat: Doing phonetics by computer, Version 5.1* [Computer program]. Retrieved from http://www.praat.org (February 2009).

Brown CA (1998) The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14: 136–93.

Bybee JL (2000) The phonology of the lexicon: Evidence from lexical diffusion. In: Barlow M and Kemmer S (eds) *Usage-based models of language*. Stanford, CA: Center for the Study of Language and Information (CSLI), 65–85.

Bybee JL (2001) *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee JL (2003) Mechanisms of change in grammaticization: The role of frequency. In: Janda R and Joseph BD (eds) *Handbook of historical linguistics*. Oxford: Blackwell, 602–23.

Chomsky N and Halle M (1968) *The sound pattern of English*. New York: Harper and Row.

Cole J, Hualde JI, and Iskarous K (1999) Effects of prosodic and segmental context on /g/-lenition in Spanish. In Fujimura O, Joseph BD, and Palek B (eds) *Proceedings of the 4th International Linguistics and Phonetics Conference, volume 2*. Prague: Karolinum Press, 575–89.

Crystal D (1997) *The Cambridge encyclopedia of the English language*. Cambridge: Cambridge University Press.

Curtin S, Byers-Heinlein K, and Werker J (forthcoming) Bilingual beginnings as a lens for theory development. *Special Issue of Journal of Phonetics*.

Cutler A, Weber A, and Otake T (2006) Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics* 34: 269–84.

Dahan D, Drucker SJ, Scarborough RA (2008) Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition* 108: 710–18.

Ellis NC (2008) Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention, and the limited L2 endstate. In: Robinson P and Ellis NC (eds) *Handbook of cognitive linguistics and second language acquisition*. London: Routledge, 372–405.

Escudero P, Hayes-Harb R, and Mitterer H (2008) Novel second-language words and asymmetric lexical access. *Journal of Phonetics* 36: 345–60.

Fennell CT and Werker JF (2003) Early word learners' ability to access phonetic detail in well-known words. *Language and Speech* 46, 2003, 245–64.

Flege J (1995) Second language speech learning: Theory, findings and problems. In: Strange W (ed.) *Speech perception and linguistic experience: Theoretical and methodological issues*. Timonium, MD: York Press, 233–77.

Frisch S, Pierrehumbert JB, and Broe MB (2004) Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22: 179–228.

Goldinger SD (1996) Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1166–83.

Hall KC (2009) A probabilistic model of phonological relationships from contrast to allophony. Unpublished PhD Dissertation, Ohio State University, Columbus, Ohio.

Hualde J-I (2005) *The sounds of Spanish.* Cambridge: Cambridge University Press.

Huberty C and Olejnik S (2006) *Applied MANOVA and discriminant analysis.* Hoboken, NJ: Wiley-Interscience.

Jakobson R (1941/1968) *Kindersprache, aphasie und allgemeine lautgestze [Child language, aphasia, and phonological universals]*. The Hague: Mouton. Originally published in 1941; citation from 1968 translation by AR Keilor.

Johnson K (1997) Speech perception without speaker normalization. In: Johnson K and Mullennix JW (eds) *Talker variability in speech processing*. San Diego, CA: Academic Press, 145–65.

Johnson K (2007) Decisions and mechanisms in exemplar-based phonology. In: Sole MJ, Speeter Beddor P and Ohala M (eds) *Experimental approaches to phonology*. New York: Oxford University Press, 25–40.

Jusczyk PW, Goodman M, and Bauman A (1999) 9-month-olds attention to sound similarities in syllables. *Journal of Memory and Language* 40: 62–82.

Kent RD and Read WC (2002) *Acoustic analysis of speech*. 2nd edition. San Diego, CA: Singular.

Ladefoged P (1975) *A course in phonetics*. Orlando, FL: Harcourt Brace.

Lavoie L (2001) *Consonant strength: Phonological patterns and phonetic manifestations.* New York: Garland.

Lléo C and Rakow M (2005) Markedness effects in the acquisition of voiced stop spirantization by Spanish–German bilinguals. In: Cohen J, McAlister KT, Rolstad K, and MacSwan J (eds) *ISB4: Proceedings of the 4th International Symposium on Bilingualism.* Somerville, MA: Cascadilla Press, 1353–71.

Martínez-Celdrán E (2004) Problems in the classification of approximants. *Journal of the International Phonetic Association* 34: 201–10.

Mascaró J (1984) Continuant spreading in Basque, Catalan and Spanish. In: Aronoff M and Oehrle RT (eds) *Language sound and structure*. Cambridge, MA: MIT Press, 287–98.

Max L and Onghena P (1999) Some issues in the statistical analysis of completely randomized and repeated measures designs. *Journal of Speech, Language, and Hearing Research* 42: 261–70.
Munson B, Edwards J, and Beckman ME (in press) Phonological representations in language acquisition: Climbing the ladder of abstraction. In: Cohn AC, Fougeron C, and Huffman MK (eds) *Handbook of laboratory phonology.* Oxford: Oxford University Press.
Munson B, Edwards J, and Beckman ME (2005) Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research* 48: 61–78.
Ohala JJ (1994) Speech aerodynamics. In: Asher RE and Simpson JMY (eds) *The encyclopedia of language and linguistics*. Oxford: Pergamon, 4144–48.
Ortega-Llebaría M (2003) Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/ in Spanish and English. In: Face T (ed.) *Laboratory approaches to Spanish phonology*. New York: Mouton de Gruyter, 237–55.
Parker S (2002) Quantifying the sonority hierarchy. Unpublished PhD dissertation, University of Massachusetts, Amherst, MA. Reproduced and distributed by GLSA (Graduate Linguistics Students' Association).
Pierrehumbert JB (2001a) Exemplar dynamics: Word frequency, lenition, and contrast. In: Bybee JL and Hopper P (eds) *Frequency and the emergence of linguistic structure*. Philadelphia, PA: John Benjamins, 137–57.
Pierrehumbert JB (2001b) Stochastic phonology. *Glot International* 5: 195–207.
Pierrehumbert JB (2002) Word-specific phonetics. In: Gussenhoven C and Warner N (eds) *Papers in laboratory phonology 7.* Berlin: Mouton de Gruyter, 101–40.
Pierrehumbert JB (2003a) Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46: 115–54.
Pierrehumbert JB (2003b) Probabilistic phonology: Discrimination and robustness. In: Bod R, Hay J, and Jannedy S (eds) *Probabilistic linguistics*. Cambridge, MA: MIT Press, 177–228.
Prince A and Smolensky P (1993) Optimality theory: Constraint interaction in generative grammar. *Rutgers University Center for Cognitive Science Technical Report* 2.
Raphael LJ, Borden GJ, and Harris KS (2007) *Speech science primer, 5th Edition.* Baltimore, MD: Lippincott, Williams and Wilkins.
Shea C and Curtin S (in preparation) Phonetic biases and frequency in the acquisition of L1 alternations.
Stampe D (1973) A dissertation on natural phonology. Unpublished doctoral dissertation, University of Chicago, IL.
Stevens KN and Keyser SJ (1989) Primary features and their enhancement in consonants. *Language* 65: 81–106.
Tabachnik BG and Fidell LS (2007) *Using multivariate statistics.* Boston, MA: Pearson Education.
Weiss DJ, Gerfen C, and Mitchel AD (2009) Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development* 5: 30–49.
Werker JF and Curtin S (2005) PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development* 1: 197–234.
Werker JF, Fennell CT, Corcoran KM, and Stager CL (2002) Developmental changes in infants' ability to learn similar sounding words. *Infancy* 3: 1–33.
Zampini ML (1994) The role of native language transfer and task formality on the acquisition of Spanish spirantization. *Hispania* 77: 470–81.
Zamuner TS (2006) Sensitivity to word-final phonotactics in 9- to 16-month-old infants. *Infancy* 10: 77–95.

Appendix 1 Word list

a. Word onset stressed syllable

Ci	Ca	Cu
bicho	*bato*	burro
gita	gato	*gudo*
díba	dado	duda

b. Word onset unstressed syllable

bidán	banana	buró
guitarra	galán	gusano
dilató	dató	dudó

c. Word medial stressed syllable

cabina	*abague*	aburro
meguilla	abogado	laguna
bedita	adapto	maduro

d. Unstressed word medial

tabila	sábado	aburró
mudinó	*idagó*	*laduró*
guila	*regadó*	agua

Note: Words in italics are nonce words.