Studies in Second Language Acquisition, 2010, **32**, 581–606.
doi:10.1017/S0272263110000276

DISCOVERING THE RELATIONSHIP BETWEEN CONTEXT AND ALLOPHONES IN A SECOND LANGUAGE

Evidence for Distribution-Based Learning

Christine E. Shea and Suzanne Curtin
University of Calgary

The identification of stressed syllables by adult second-language (L2) Spanish learners was examined for evidence of influence of an allophonic alternation driven by word position and stress. The Spanish voiced stop-approximant alternation, whereby stops occur in stressed-syllable and word onsets, was utilized. If L2 learners track the distribution of this alternation, they should tend to link stops to stressed syllables in word-onset position and approximants to unstressed, word-medial position. Low- and high-intermediate-level first-language English learners of Spanish as well as native Spanish and monolingual English speakers listened to a series of nonce words and determined which of the two consonant-vowel (CV) syllables they perceived as stressed. In Experiment 1, onset allophone and vowel stress were crossed. In Experiment 2, the onset allophone alternated and a vowel

The authors received helpful comments and feedback from colleagues in the Department of Linguistics at the University of Calgary as well as from audiences at New Sounds 2010 in Poznan, Poland and the Second Workshop on Second Language Speech Perception, ASA, in Portland, OR. Additionally, the first author would like to thank colleagues at the Centro de Enseñanza de Lenguas Extranjeras de la Universidad Nacional Autónoma de México (CELE-UNAM) for assistance in obtaining participants.

Address correspondence to: Christine E. Shea, Department of Linguistics, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2K 1N4, Canada; e-mail: christine-shea@uiowa.edu.

© Cambridge University Press 2010 0272-2631/10 $15.00

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

unmarked for prominence was used. The results show that the monolingual English and low-intermediate groups were more likely to perceive syllables with stressed vowels as stressed, regardless of the allophone onset. In contrast, listeners with greater Spanish proficiency performed similarly to native Spanish speakers and were more likely to perceive stress on syllables with stop onsets, a pattern that follows the distributional information of Spanish. This finding suggests that learning the interplay between allophonic distributions and their conditioning factors is possible with experience and that knowledge of this relationship plays a role in the acquisition of L2 allophones.

Studies of second-language (L2) speech perception have primarily explored how target language sounds fit into the sound system of the speaker's native language (L1). In particular, these studies examine whether nonnative sounds represent new categories, are classified into existing L1 phoneme categories, or are similar to existing allophones. For example, Flege's (1995) speech learning model (SLM) and Best's perceptual assimilation model (PAM; Best, 1994, 1995; Best & Tyler, 2007) emphasized L1-L2 perceptual similarity as predictive of difficulties in the discrimination of nonnative contrasts. Flege characterized nonnative phones along a continuum of similarity to L1 phones that ranged from identical to similar to new. A new phone may be difficult to perceive for inexperienced listeners, but the sound will eventually become differentiated from L1 phones (and from other L2 phones) as learners gain experience in the L2. Best's (1994, 1995; Best & Tyler, 2007) PAM maintains that listeners assimilate nonnative phones into L1 categories based on a perceived continuum of category goodness. If two nonnative phones are considered to be of the same native category, they will be very difficult to discriminate. If the two phones are perceptually assimilated to the same native category but differ in their perceived category goodness, they will be discriminated more easily. Finally, if nonnative phones are assimilated to different native categories, discrimination will be very easy. Although these models have provided testable hypotheses regarding how L2 learners acquire new sounds, the effect of context on the acquisition of target language sounds is not explicitly addressed, and it is thus difficult to draw clear predictions for the SLA of sounds that differ in the nature of their contrastiveness. For instance, allophones are perceived as members of the same category by native listeners on phonological identification tasks but as different sounds on phonetic discrimination tasks (Jaeger, 1980; Kazanina, Phillips, & Idsardi, 2006).

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

In the phonological literature, it is traditionally assumed that segments can be related either through contrast or allophony (see, e.g., Steriade, 2007). Two segments contrast if their distribution in the lexicon of a language is not predictable. Sounds that are related through allophony, however, occur in conditioned distributions: Each allophone occurs in a regular, predictable context (Crystal, 1995). Furthermore, for certain cases of allophony, the contexts are in complementary distribution; that is, their distributions are nonoverlapping and distinctive.

Most L2 phonological and phonetic models regard allophone acquisition as similar to that of phoneme acquisition and thus assume that the way in which the new sound is assimilated into existing sound categories will determine the ease or difficulty of its acquisition. In the present study, the allophone acquisition task is approached from a slightly different perspective. Instead of focusing on the acquisition of new sound categories and the way in which L2 allophones assimilate into the L1 sound inventory, the current study examines whether learners are aware of the factors that condition their distribution. In other words, the focus is whether learners recognize and store information about the specific context in which each variant occurs. It is predicted that learners with greater language experience will be aware of the factors that condition the allophonic alternation and that this awareness will shift their perception of the target language.

The present study examines how L1 English learners of Spanish perceive the conditioning factors that drive the stop-approximant alternation in the target language. In Spanish, voiced stops [b, d̪, g] alternate with the approximants [β̞, ð̞, ɣ̞].[1] According to phonological descriptions of this alternation (see Harris, 1969), the stop is the underlying form and the approximant is the allophone; nonetheless, in distributional terms, the approximant is in fact more common than the stop. Stops appear after a pause or a nasal, and the dental stop appears after the lateral /l/. Traditional analyses have characterized the alternation in terms of feature spreading. The feature [+ continuant] spreads from the adjacent vowels to the [– continuant] stops, which then become approximants. However, the difficulty quickly encountered with this analysis was how to account for the [– continuant] nature of /d/ next to the lateral. This anomaly was addressed by assuming that /l/ is underlyingly unspecified for [continuant] in Spanish. In an alternative account, Face (2002) argued that the dental stop surfaces next to /l/ because of the proximity in place of articulation. According to Face, this overlapping place of articulation causes the stop to have a strong release.

More recent research has examined the noncategorical phonetic realizations of the alternation and considered conditioning factors such as

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

word position and stress. According to Hualde (2005), more open, approximant-like articulations occur posttonically than in the onset of a stressed syllable (see also Lavoie, 2001; Ortega-Llebaría, 2003; Shea & Curtin, in press). Sounds that occur in word-medial, unstressed position will be more approximantlike, with higher intensity and no release burst. In contrast, sounds that occur in word-initial, stressed position will tend to be more stoplike in terms of their acoustic realization, with lower intensity and a strong release burst. If L2 learners can track this distribution, then upon hearing the phonetic cues that indicate that one variant consistently occurs in the same context, learners will begin to build the distribution for this allophone. Another set of cues will be associated with the other variant, stored as part of a separate context. Because this information is probabilistic (stop cues are more likely to be heard in word-initial, stressed position), listeners will optimize their perception based on prior probabilities, which emerge directly from experience with the language.

The current study analyzes whether L2 learners of Spanish associate the acoustic cues that characterize the stop-approximant alternation to the contextual factors of stress and word position. The focus is on whether learners can identify a stressed syllable based on the onset consonant that accompanies the syllable, or on the position in the word, or both. This focus will determine whether hearing a stop or an approximant in certain word positions or positions relative to a stressed vowel indirectly influences the likelihood of perceiving stress. The appearance of these relationships will provide evidence to support the claim that learners associate contextual factors to the allophonic variants and have fairly sophisticated knowledge about sound alternations.

One way learners might develop this contextual knowledge is by means of a distribution-based learning mechanism. Such models define phonetic categories as distributions of speech sounds. A distribution-based learning mechanism can track information in the input, such as the co-occurrence of sounds, and allows learners to draw upon this information over the course of acquisition. Exemplars are classified together due to similarity along one or multiple dimensions, and speech sounds are assumed to have Gaussian, or normal, distributions, given that variability in production (whether across or within speakers) causes alterations to the prototypical speech category (Pierrehumbert, 2003). Sound category acquisition involves building robust distributions in the phonetic space to support the categorization of incoming sounds. Tokens that share characteristics with the most frequent exemplars in the input will be stored in the highest part of the curve, whereas those of extremely low frequency will be in the tails of the distribution. Consequently, as the phonetic category space becomes more densely populated, it becomes more difficult to bias perception one way or another

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

because more recent exemplars will not carry as much perceptual weight in terms of categorization. As the listener gains experience with a language, it becomes less likely that the most recently experienced token or outlying tokens will alter the distribution. By their very nature, distribution-based models assume that learners create and store highly detailed and rich exemplar-type representations and that grammar emerges as a result of generalizations across all of the stored items in the mental lexicon (see, e.g., Goldinger, 1997; Johnson & Mullennix, 1997; Pierrehumbert, 2001, 2003).

Researchers have shown that both adult and infant listeners are able to form categories based on the distributions of speech sounds and to shift their perception of allophones by using the same mechanism (e.g., Goudbeek, Cutler, & Smits, 2006; Holt & Lotto, 2006; Maye & Gerken, 2001; Maye, Werker, & Gerken, 2002). In the case of the stop-approximant alternation, the learning task for L1 English learners of Spanish involves the creation of a new distribution for the approximant category separate from the stop category that occurs in their L1. However, accumulating the knowledge required to determine how context affects the production of the voiced stops in Spanish requires time and exposure to the language. The separate categories will emerge only when the learners have accumulated sufficient examples of each category and the phonetic details that distinguish these tokens.[2]

It is assumed, first, that if adult L2 learners of Spanish track distributional information in the input, they will expect to hear more stops in word-initial, stressed position than in word-medial, unstressed position. Learners will thus be more likely to perceive stress when the stressed syllable is accompanied by a stop consonant and less likely to perceive a syllable as stressed when it begins with an approximant. Second, given that it takes extensive linguistic experience to build sound category distributions, more advanced learners will associate the alternant with its most probable context of occurrence, whereas less advanced learners will not make this association. The fact that the distribution of the stop-approximant alternation in Spanish is not categorical—other than the postnasal contexts for stops (and postlateral in the case of [d̪])—suggests that acquisition must also be probabilistic in nature. A stressed syllable is much more likely to occur in word-initial position with a stop consonant in the onset than in word-medial position accompanied by an approximant in the onset, although this likelihood is still not 100%.

The first experiment used nonce words that consisted of two consonant-vowel (CV) syllables crossed for allophone onset and stressed versus unstressed vowels to determine if the perception of stress shifts in relation to the allophone onset. Because stress is one of the conditioning factors that drives the allophonic alternation and because stress is more likely to co-occur with stop onsets than with approximant

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

onsets in the input, listeners with (more) knowledge of Spanish were predicted to perceive stress with greater likelihood on stop-initial syllables than on approximant-initial syllables. In the second experiment, allophone onsets alternated between the two allophones and the vowels were equated for prominence. The groups with more Spanish experience are predicted to select stop-onset syllables as stressed with greater likelihood than groups with less Spanish experience, given increased knowledge of Spanish distributional information.

EXPERIMENT 1: CONSONANT AND VOWEL STRESS SHIFT

In Experiment 1, participants listened to bisyllabic stimuli and were asked to select the syllable they perceived as stressed. Stress detection served as an indirect method of determining the perceptual association of stress with stop-initial syllables. This indirect behavioral method circumvents problems with phonetic versus phonological representations and also gets at the main question that motivates this study: Are learners aware of the contextual factors that drive allophonic alternations in their target language?[3]

Method

Participants. Nineteen low-intermediate and 20 high-intermediate L1 English-L2 Spanish learners were invited to participate in the current study. These learners were recruited from different sections of the same second- and third-year university-level Spanish courses (with eight different instructors). Participants filled out an autobiographical questionnaire about their experience with Spanish, which revealed that no participant from either group had spent more than 6 weeks in a Spanish-speaking country and that these participants did not use Spanish outside of the classroom context. Participants had received explicit instruction on the stop-approximant alternation during their class sessions. This alternation is described as a softening of the /b, d, g/ sounds between vowels. This was covered in the low- and high-level classes. None of the participants had previously taken courses in Spanish phonology or phonetics.

Fifteen native Spanish speakers (NSSs) were recruited from the Center for the Teaching of Foreign Languages, at the National Autonomous University of Mexico in Mexico City. They were age-matched with the L1 English learner groups. The NSSs had never lived abroad or attended bilingual schools and did not have more than 3 hr per week of contact with English. Finally, 15 monolingual native English speakers (NESs)

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

were recruited from a university psychology subject pool and were also age-matched with the two learner groups. None of these participants spoke any language other than English. All participants were paid the equivalent of $15 for their time or, in the case of the NESs, received course credit. None had any reported hearing difficulties.

Stimuli. The stimuli were created from naturalistic speech samples, recorded by a native female speaker of Spanish from Mexico City. Recordings were made in a soundproof booth directly onto a PowerMac computer (GIA417″ Soundcard) and a Sennheiser microphone. The microphone was placed into a stand and maintained at a 45º angle at all times, approximately 5.5 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.

The speaker read a list of bisyllabic CVCV nonce words in which the first syllable was stressed, following the expected stress pattern of Spanish. She was asked to read these nonwords in Spanish. The consonants were [b], [d̪], or [g] and the vowel was [a]. Using Praat 5.1 (Boersma & Weenink, 2008), the consonants were spliced from the vowels to create four separate sounds: stop (word onset), approximant (second syllable onset), stressed vowel, and unstressed vowel. For example, the nonce word *baba* [báβa] provided four separate segments: [b], stressed [a], [β], and unstressed [a]. These four sounds were combined to create four different tokens: [báβa], [βába], [baβá], and [βabá]. This procedure was repeated for both [d] and [g], which created a total of 12 tokens. Stimuli ranged in length from 67 ms to 78 ms.

All stimuli were presented to two NESs and two NSSs and judged for how natural or speechlike the sounds seemed on a scale of 1 (*extremely artificial sounding*) to 5 (*extremely natural sounding*). Stimuli that did not receive an initial rating of 4.5 or higher were respliced and presented to the judges again. Given that approximants do not—or at least very rarely—occur in word-initial position following a pause, any stimuli that exhibits this phonotactic pattern will necessarily be deemed unnatural by NSSs.[4] Asking native speakers to rate stimuli that violate the phonotactic constraints of their L1 may be difficult. However, the raters were asked for a global impression of how Spanish-like or English-like—that is, how speechlike—the sounds seemed to them. Care was taken not to direct their attention to any specific aspect of the stimuli, which minimized the likelihood that they would pay too close attention to the allophone onsets.

Procedure. Testing in Mexico was carried out in a small, quiet room with the door closed. Outside of Mexico, testing was conducted in a soundproof booth. Participants were seated at a table in front of a

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

PowerMac computer and stimuli were presented through headphones at a comfortable volume using PsyScope experimental software program (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants were told that they were going to listen to nonwords in Spanish. For the NSS group and the two learner groups, all communication occurred in Spanish (for the low-intermediate group, clarifications were given in English when requested). The NES group was told that the nonwords were from a foreign language. Participants were instructed to select which syllable they perceived as stressed by pressing a key on the computer keyboard. The keys were marked 1 and 2 with stickers. Subsequent tokens were played after the participant made their selection, with an interstimulus interval of 1,000 ms. If no decision was made within 1,500 ms, the trial timed out and the following trial was played. This occurred in 2.5% of all trials.

Participants were given five pretest trials before beginning the experimental trials. The pretest trials were randomly selected by the experimenter from among the test stimuli and were different for each participant. Because there were technically no right answers—listeners were predicted to perceive stress either according to the vowel or the consonant onset, based on language experience—no feedback was provided during the pretest trials.

All participants first completed a stress-detection task prior to the pretest trials, to ensure that the results would not be compromised due to an inability to detect stressed syllables. The stress-detection task involved listening to a series of 20 nonce words that followed Spanish phonotactic and prosodic constraints, read by a Mexican female NSS. Words that could be analogically compared to conjugated verbs or infinitives were avoided where possible (see the Appendix). Participants indicated by means of pressing keys on the computer keyboard whether they thought stress fell on syllable 1, 2, or 3. Sixteen of the 20 lexical items were bisyllabic, 8 with stress on the first syllable and 8 with stress on the second. The four remaining items were trisyllabic and stress fell either on the first (n = 2) or the third (n = 2) syllable. Only the results of participants who obtained at least 75% accuracy on this task were included for analysis.

Results

Stress Perception: Allophone Plus Stressed Vowel. Four participants from the low-intermediate group and six from the high-intermediate group were excluded because they did not reach the 75% criterion on the stress-detection task. This left a total of 15 participants for the NES, low-intermediate, and NSS groups and 14 participants for the high-intermediate group.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

The prediction is that the stress perception of the NES and low-intermediate groups will not be affected by the allophone in the onset position of the syllable but instead will only be affected by the vowel. In other words, these listeners will perceive stress in relation to the prominence of the vowel. To test this prediction, a one-way ANOVA was carried out to see whether there were overall differences among the groups in terms of a connection between stressed vowels and one of the allophone variants. Group was the independent variable and the dependent variable was a calculated ratio value as in (1).

$$\frac{\text{Stops} + \text{Stressed vowel perceived as stressed}}{\text{Approximants} + \text{Stressed vowel perceived as stressed}} = \text{Ratio}. \quad (1)$$

Ratios greater than 1 are predicted for the NSS and high-intermediate groups, which would indicate that more syllables with initial stops and stressed vowels were perceived as stressed than syllables with initial approximants and stressed vowels. For the low-intermediate and NES groups, the ratios are predicted to be around 1, which would indicate a lack of preference for either allophone. Figure 1 presents the means for each of the groups.

The results from the one-way ANOVA revealed a significant effect for group, $F(3, 56) = 30.85$, $p < .001$. The ratios for the NES and low-intermediate groups were close to 1 (NES: $M = 0.95$, $SD = 0.15$; low-intermediate: $M = 1.02$, $SD = 0.25$), whereas the ratios for the high-intermediate and NSS groups were significantly greater (high-intermediate: $M = 2.1$, $SD = 0.7$; NSS: $M = 2.78$, $SD = 0.122$). Tukey's honestly significant difference (HSD) post hoc tests showed significant differences between the NSS group and the NES ($p < .01$) and low-intermediate ($p < .01$) groups, and between the high-intermediate group and the NES ($p < .05$) and low-intermediate ($p < .05$) groups but not between the high intermediate and NSS groups ($p = .18$). These results show that the high-intermediate

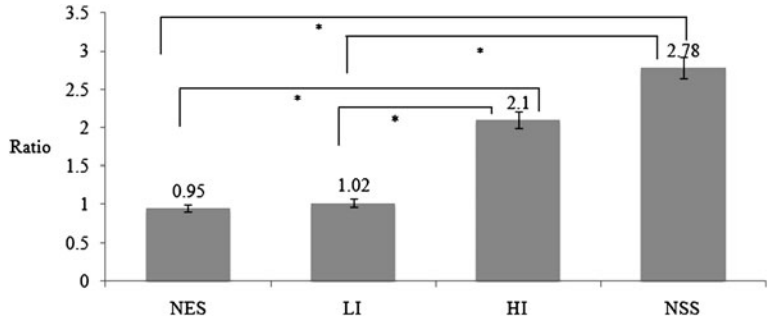


Figure 1. Ratio values for Experiment 1; *$p < .05$.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

and NSS groups perceived stress significantly more often on stop-initial than approximant-initial syllables, which suggests that listeners with more Spanish experience associate stress with the stop allophone.

It is possible, however, that the selection of stressed syllables is also influenced by knowledge of the predominant stress pattern in English and Spanish. To determine whether such a generalization occurred, the selection of stressed syllables was compared with the most frequent stress patterns in Spanish and English.

Testing for Trochaic Bias. Of the 4,829 most frequent polysyllabic words in Spanish, those that end with a vowel follow a trochaic stress pattern 87.5% of the time (Alameda & Cuetos, 1995). In English, the most common word type is bisyllabic with a trochaic stress pattern. Only about a quarter of English words are polysyllabic with a weak initial syllable (Cutler & Carter, 1987). It is possible that participants simply relied on predominant stress patterns of English and Spanish rather than using the allophone or the perception of vowel stress—the two independent variables used in the main analysis—to select the stressed syllable.

To investigate the proportion of stressed-syllable responses that corresponded to initial stress and to determine whether positional biases affected listeners' stress perception, the proportion of initial syllables perceived as stressed was examined, independently of the allophone. The total number of trials successfully completed by each group (15 participants × 12 trials = 180 per group, except for the high-intermediate group [n = 14], for which there were only 165 responses) was considered. The proportion of first syllables perceived as stressed, independent of vowel prominence, was calculated, followed by the proportion of trials perceived as stress-initial, with stop allophones.[5] Although it is possible that all groups could potentially demonstrate a trochaic bias for CVCV forms, it is predicted that only the high-intermediate and NSS groups will show a preference for syllables with stops over approximants in trochaic contexts. The other two groups are predicted to perform at around chance (.5). Table 1 presents these results.

To permit adequate comparisons among groups, ratio values (stressed $\text{vowel}_{\sigma 1}$/total number of syllables) were calculated and a one-way ANOVA was carried out, with groups as the independent variable and the ratio values as the dependent variable. There was no significant effect for stress detection on the first syllable, $F(3, 55) = 1.97$, $p = .22$. All groups perceived stress on the first syllable at proportions above chance, consistent with the lexical statistics of Spanish and English. A second ANOVA was conducted with stop-onset syllables in initial position as the dependent variable (also a ratio) and the results revealed a significant main effect for language experience, $F(1, 55) = 22.03$, $p < .001$. Tukey's HSD post hoc tests revealed significant differences between the NSS

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

Table 1. Stress perception on first syllable across groups and onsets

Syllable	NES	LI	HI	NSS
First				
Trials	105/174	98/179	95/164	100/176
Proportion	.60	.55	.59	.57
Stop onset				
Trials	57/87	55/90	59/82	73/88
Proportion	.65	.62	.71	.81

Note. LI corresponds to low-intermediate and HI corresponds to high-intermediate participant groups.

group and the other three groups ($p < .001$), but no significant differences emerged among these three groups: high- versus low-intermediate, $p = .134$; high-intermediate versus NES, $p = .109$; NES versus low-intermediate, $p = .08$.

It can be concluded from these results that all four groups of listeners show a bias toward hearing trochaic stress patterns, but the NSSs demonstrate an additional bias toward perceiving stress on syllables that have stop onsets, consistent with the Spanish distributional information. The high-intermediate group did not differ from the other three groups in their preference for stop onsets in initial position of trochaic nonwords. This result was somewhat surprising, given that the high-intermediate group was expected to pattern largely along the same general lines as the NSSs. One explanation for this result may be that the preference for trochaic bias in stress perception overrides any preference for stop onsets, even for groups that have had extensive experience with Spanish.

According to the predictions, the high-intermediate and NSS groups should be most likely to perceive stress on stop-initial syllables with stressed vowels, regardless of the position in the word. To determine whether this prediction held, a goodness-of-fit chi-square (χ^2) test on the proportion of syllables perceived as stressed for each of the four possible onset-vowel combinations was conducted for each group. If the high-intermediate and NSS groups prefer stops as onsets to stressed syllables, a difference among the four combinations is expected, with the different allophone types clustering together for these two groups. In contrast, stressed-unstressed vowel factor clustering is expected for the low-intermediate and NES groups. These results are presented in Figure 2. For the NSS group, preference for syllable stress was not equally distributed, $\chi^2(3, N = 176) = 4.21$, $p < .05$. A similar pattern was found for the high-intermediate group, $\chi^2(3, N = 174) = 7.22$, $p < .05$. There were no significant differences across the

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

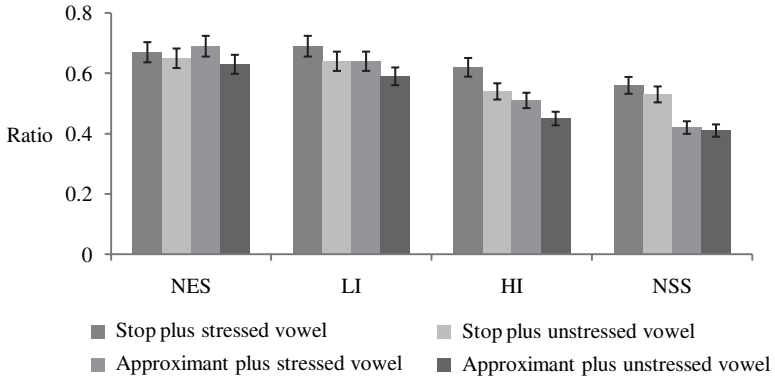


Figure 2. Proportion of syllables perceived as stressed.

four contexts for the low-intermediate (p = .091) or NES group (p = .128). The finding that language experience led to significant differences in the perception of stress across the four contexts shows the pivotal role of this variable in terms of how the allophone drives stress perception in Spanish.

These results show that L1 trochaic stress perception preference affects both L1 Spanish and L1 English listeners. However, NSSs perceived initial syllables with stops as stressed at a significantly higher rate than the L1 English groups. The hypothesis that experience determines this particular aspect of stress perception is thus only partially supported.

Having established that all groups are affected by a bias toward trochaic stress, which can be attributed to their L1, an analysis is required that can reflect the likelihood that a particular group will be affected by vowel stress or allophone onset. Thus, a logistic regression analysis was carried out to examine which factor (allophone onset or vowel stress) is most likely to affect the stress perception of each group.

Logistic Regression Analysis. Through logistic regression analysis, the predictor variables can be connected to the probability that they affect the dependent, or outcome, variable. Hierarchical logistic regression was used because it allows for testing the significance of each predictor in a cumulative manner and because it is better suited to analyses with a small n (Tabachnik & Fidell, 2007).

Three levels were used in this analysis: (a) stressed vowel (level 1) included all syllables with a stressed vowel collapsed across allophone type; (b) allophone type and stressed vowel (level 2) included allophone variant and vowel stress (approximant plus stressed vowel, e.g., βába, and stop plus stressed vowel, e.g., báβa); and (c) position (level 3)

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

included position and allophone (approximants in word-medial position and stops in word-initial position, collapsed across stress). See Table 2.

Overall, the results indicate that listeners with Spanish experience are more likely to rely on the allophones when detecting a stressed syllable. The results show that at level 1, the stressed vowel reached significance for the low-intermediate and NES groups, which suggests that this predictor was particularly good at identifying members of these groups. For the high-intermediate group, significant results were also obtained, but the odds ratio was less than 1, which indicates that members of the high-intermediate group are less likely to perceive stress on syllables with a stressed vowel than those of the NES and low-intermediate groups. This predictor did not reach significance for the NSS group. Thus, participants who relied primarily on whether the syllable contained a stressed vowel were more likely to be members of the low-intermediate group and the NES group and less likely to belong to the high-intermediate group.

Table 2. Odds ratios from hierarchical logistic regression

Predictor	NES	LI	HI	NSS
Level 1				
Stressed vowel	2.026** (.223)	1.707* (.193)	0.688** (.153)	0.542 (.196)
Level 2				
Stressed vowel	1.285 (.243)	1.430 (.250)	1.116 (.261)	0.733 (.314)
Stop plus stressed vowel	0.964 (.239)	0.935 (.196)	0.820 (.204)	1.362 (.214)
Approximant plus stressed vowel	1.930* (.899)	1.207 (.164)	0.613** (.234)	0.593* (.264)
Level 3				
Stressed vowel	0.966 (.331)	1.390 (.332)	1.026 (.271)	0.636 (.373)
Stop plus stressed vowel	0.724 (.355)	1.316 (.269)	0.888 (.215)	1.369 (.279)
Approximant plus stressed vowel	1.825* (.306)	1.345 (.253)	0.663 (.284)	0.849 (.337)
Approximant medial	2.323 (.590)	1.145 (.357)	0.866 (.326)	0.474 (.462)
Stop onset	2.093* (.734)	1.115** (.711)	0.649 (.452)	0.545 (.544)

Note. LI corresponds to low-intermediate and HI corresponds to high-intermediate participant groups. Odds ratios greater than 1 indicate likelihoods greater than chance. Standard error values appear in parentheses.
* Wald $\chi 2$, $df = 1$, $p < .01$.
** Wald $\chi 2$, $df = 1$, $p <. 05$.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

The level 2 predictors demonstrated that if a participant was likely to perceive stress on an approximant-initial syllable, he or she was twice as likely to be a NES as a member of one of the other speaker groups. This shows that experience with Spanish affects the likelihood of perceiving stress on an approximant-initial syllable.

The results for the level 3 predictor showed that the NES group was significantly more likely than the other three groups to perceive stress on approximant-medial and stop-initial syllables. Moreover, only this group was more likely to perceive stress on approximant, word-medial syllables. If the results for the stop-initial syllables are considered, the fact that the NES group was more likely to perceive stress on these syllables independent of vowel stress indicates a strong effect for position on the probabilities of stress detection by this group. This appears to contradict the initial hypothesis that stress perception for the groups with Spanish experience would be driven primarily by the allophone. However, a possible explanation may lie in perceptual biases that drive the NES listeners to prefer stress on initial syllables (i.e., trochaic stress pattern) to a greater extent than the other three listener groups. Because the NESs are completely unfamiliar with the sounds used in the stimuli, when forced to rely purely on the onset allophone this group reverts to the bias for initial stressed syllables, consistent with the pattern that predominates in English.

It still remains to be seen whether the perception of stress is being driven by the onset alone or whether the presence of vowel stress plays a key role. To investigate these two possibilities, a second experiment was carried out in which only the onset—whether stop or approximant—was shifted and the vowel stress held steady in both syllables of the bisyllabic nonword (i.e., the vowel prominence was neutral).

EXPERIMENT 2: ALLOPHONE ALTERNATION WITH VOWEL HELD STEADY

Experiment 2 examined whether shifting the consonant onset influences the perception of stress when the prominence of the vowel remains constant across syllables. Participants listened to bisyllabic CVCV sequences in which the vowel was held steady and the syllable onsets alternated between the two allophones (e.g., *baβa* vs. *βaba*). Thus, the perception of stress is an illusion in this experiment. That is to say, stress is not explicitly present in the signal but rather inferable from the presence of a stop onset, provided the listener is sensitive to the distributional information that connects stress and stops in the Spanish input. The presence of a stop onset is predicted to be one of the cues that NSSs will use to detect stress in the signal. Other cues may come from the vowel and contribute to the perception of stress for the listener.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

NSSs and listeners of more advanced Spanish proficiency are expected to use their knowledge of Spanish phonotactics to perceive stress on the syllable with the stop onset and not on the syllable with the approximant onset. The NES and low-intermediate groups will have limited or incomplete knowledge to draw upon and are therefore predicted to perform at or around chance.

Method

Participants. The same participants as in Experiment 1 took part in Experiment 2.

Stimuli. The stimuli for this experiment consisted of CVCV nonwords, created from the same naturalistic speech samples used in Experiment 1. However, instead of shifting the vowel [a] between stressed and unstressed values, the vowel was held steady and only the consonants switched. A stressed vowel token was chosen from among the CVCV stimuli used in Experiment 1 and the intensity adjusted to 75 dB. The F1 value was 806 Hz and F2 was 1628 Hz and the duration was 74 ms. Because stressed syllables are detectable only in comparison to unstressed syllables, by maintaining both vowels equal in terms of vowel duration, intensity, and pitch, a stress-neutral CVCV item was created. In addition to the onset shifts, the duration of the onset consonants was manipulated to test the hypothesis that shorter segments are more likely to be perceived as the onset to unstressed syllables than longer segments. Lavoie (2001) found that, in addition to manner contrasts, the length of the allophone served to distinguish stoplike productions from approximantlike productions in a group of native Mexican Spanish speakers. The length of the onset consonant varied in terms of duration, from values of 33%, 67%, or 100% of the original duration. Finally, the consonants were spliced onto the vowel and counterbalanced for allophone variant. The place of articulation was held constant within each CVCV sequence. This created 18 possible CVCV combinations. For example, a nonce word with the velar versions of the allophones and with the approximant in word-initial position would take the form [γaga]. Stimuli ranged in length from 171 ms to 201 ms. The 54 stimuli were randomly presented together with the 36 stimuli from Experiment 1 in one block.

As in Experiment 1, two NESs and two NSSs were asked to judge the stimuli for naturalness on a scale of 1 (*extremely artificial sounding*) to 5 (*extremely natural sounding*). Only stimuli rated 4.5 or higher were used for the experiment. The same caveats hold for the explanation of naturalness in terms of these stimuli as for the first experiment.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

Procedure. The same procedure as in Experiment 1 was used.

Results

A mixed ANOVA was conducted to determine if there was any effect for the three different onset lengths. Group was the between-subjects variable and segment length (33%, 67%, or 100%) was the within-subjects variable; the results revealed a nonsignificant main effect for group, $F(3, 42) = 1.8$, $p = .08$, and a nonsignificant main effect for the within-subjects variable of length, $F(6, 126) = 2.1$, $p = .162$. Therefore, consonant lengths were collapsed into a single variable for subsequent analyses.

Stress Perception: Allophone Plus Nonprominent Vowel. As in Experiment 1, a one-way ANOVA was conducted to determine whether participants perceived stress in higher proportions on stop or approximant syllables. Group was the independent variable and the dependent variable was a ratio measure calculated as in (2):

$$\frac{\text{Stop-initial syllables perceived as stressed}}{\text{Approximant-initial syllables perceived as stressed}} = \text{Ratio.} \quad (2)$$

There was an effect for group, $F(3, 55) = 20.18$, $p < .001$. Post hoc Tukey's HSD tests revealed significant differences between the NSS group and the other three groups (NES, $p < .01$; low-intermediate, $p < .01$; high-intermediate, $p < .05$) and between the NES group and the other three groups (low-intermediate, $p < .01$; high-intermediate, $p < .01$; NSS, $p < .01$). There were no significant differences between the high-intermediate group and the low-intermediate group ($p = .212$).

Testing for Trochaic Bias. As with Experiment 1, the responses were examined to see if there was a bias for perceiving stress on the first syllable of the word. Of the 1,566 possible test trial responses, 4% were discarded due to timing out. This left a total of 1,503 responses for analysis. Table 3 gives the proportion of first syllables perceived as stressed and the raw totals for each group. The totals are also broken down into the percentage of syllables perceived with stress on the first syllable and the percentage of stop-initial stressed syllables. These data are presented in Table 3.[6]

The results did not reveal a significant effect for group, $F(1, 55) = 1.8$, $p = .221$, possibly because the NSS and high-intermediate groups clustered together, as did the low-intermediate and NES groups. The mean ratio for the NSSs was 2.4 ($SD = 0.6$), and for the high-intermediate group, it was 1.9 ($SD = 0.7$), which demonstrates that these participants prefer to associate stress with stop syllable onsets. For the

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

Table 3. Stress perception on first syllable across groups and onsets

Syllable	NES	LI	HI	NSS
First				
Trials	440/746	444/792	397/747	380/785
Proportion	.60	.56	.53	.48
Stop onset				
Trials	228/440	229/444	215/397	229/380
Proportion	.53	.52	.54	.60

Note. LI corresponds to low-intermediate and HI corresponds to high-intermediate participant groups.

low-intermediate group, the mean ratio was 1.1 (*SD* = 0.37), and for the NESs, the mean was 0.93 (*SD* = 0.4), which suggests that participants in these groups did not distinguish between stop and approximant onsets as related to stress. These results confirm the predictions about the effect of the onset allophone and the perception of stress as a function of group membership. See Figure 3.

The results of this analysis indicate that, with Spanish experience, the onset allophone—whether stop or approximant—can lead to an illusory stress perception effect. The NSS group heard stress significantly more often on stop-initial syllables than on approximant-initial syllables compared to the other three groups, and the NES group perceived stress significantly less often than the other three groups when the syllable had a stop in onset position. However, the finding that the two learner groups were not significantly different from one another but

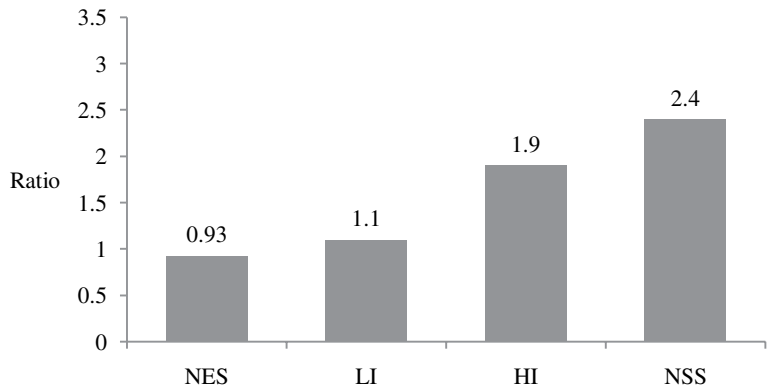


Figure 3. Ratio values for Experiment 2.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

were in fact significantly different from the NSS and NES groups suggests that experience has affected the perception of the illusory stress effect. The high-intermediate group does not perceive stress in Spanish purely by relying on their L1, which would have meant patterning with the NES group. Nonetheless, they still do not exhibit the illusory stress effect in the same way as the NSSs.

Logistic Regression Analysis. As in Experiment 1, a logistic regression analysis was run, with the predictors of approximant onset or stop onset.

The predictor variable that most likely affects the likelihood of belonging to the different groups is expected to be the perception of stress on a syllable with an approximant. The likelihood of perceiving stress on an approximant syllable should be close to chance for both the low-intermediate and NES groups. The likelihood of perceiving stress on an approximant syllable should be lower for the two groups with more Spanish experience. Table 4 presents the results from this analysis.

The approximant allophone predictor reached significance and was able to correctly separate participants into the four speaker groups. The low-intermediate participants were more likely to select a syllable with an approximant in the onset as the stressed syllable, and the likelihood for the NES group was around chance. The NSS and high-intermediate participants were less likely to select a syllable as stressed if it had an approximant in the onset.

These results suggest that with increased L2 experience, L1 English learners of Spanish perceive an illusory stress effect induced by the onset allophone in bisyllabic nonwords. Learners come to associate the stop allophone with stress and to associate the approximant allophone with an absence of stress after considerable experience with Spanish. These results add to those from Experiment 1 and further suggest that when vowel stress is not accessible as a direct cue to word stress, listeners with more knowledge of Spanish follow the distributional information found in Spanish and disprefer approximants as the onset to stressed syllables. The fact that stop onset did not reach

Table 4. Odds ratios from logistic regression

Predictor	NES	LI	HI	NSS
Approximant	1.048** (.207)	1.986* (.242)	0.687* (.138)	0.547** (.175)
Stop	0.898 (.264)	1.407 (.277)	1.096 (.192)	1.07 (.201)

Note. LI corresponds to low-intermediate and HI corresponds to high-intermediate participant groups. Standard error values appear in parentheses.
* Wald $\chi 2$, $df = 1$, $p < .01$.
** Wald $\chi 2$, $df = 1$, $p < .05$.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

significance for any of the groups indicates that having stops in onset position is expected by listeners of both L1 backgrounds. In other words, none of the groups was more likely than the others to perceive stop onsets as stressed, which follows from the expectation that a natural bias toward preferring stops is at work. It has been extensively documented in the literature on phonology that stops are ideal onsets, whether due to markedness or phonetically grounded motives (Archangeli & Pulleyblank, 1994; Prince & Smolensky, 2004).

GENERAL DISCUSSION

This study investigated whether L2 learners are aware of the factors that influence the stop-approximant alternation in Spanish and, if so, whether language experience plays a role. These results indicate that learners are able to track the distribution of the allophones and that, over time, the relationship between the allophones and the contexts in which they surface can gradually be learned.

One possible way to explain how L2 learners acquire allophones is through distributional learning, which was demonstrated in Experiment 1. As expected, given the predominant pattern for main stress in both Spanish and English, all four groups showed a preference for perceiving stress on the first syllable. However, upon closer examination, the bias in favor of stop-initial, stressed syllables only occurred with the high-intermediate and NSS groups. This suggests that participants in the NSS and high-intermediate groups have acquired knowledge about the distribution of these allophones. In particular, a stressed syllable is more likely to have a stop onset than an approximant onset, and an unstressed syllable is more likely to have an approximant onset than a stop onset. Additionally, the L1 of the participants (both Spanish and English) also appears to influence stress perception and leads to a default trochaic stress preference. However, listeners with sufficient knowledge of Spanish further demonstrate a bias in favor of perceiving stress on initial syllables with a stop onset, which reflects Spanish distributional information. This shows that experience with Spanish can actually shift the perception of stress in nonnative speakers in the direction of the distributional information found in the target language.

Distribution-based learning mechanisms play a fundamental role in exemplar-based models of phonological acquisition. In these types of models, phonological categories are represented as probability distributions over a mental phonetic acoustic or auditory map (Pierrehumbert, 2003). Categories emerge when multiple exemplars that are phonetically similar accumulate in the same location on the phonetic map. For example, in the case of the stop-approximant alternation, NSSs would have a large number of exemplars at the coordinates for the

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

voiced bilabial approximant [β̞] and the voiced bilabial stop [b]. These two categories share many articulatory and acoustic characteristics in addition to being represented by the same orthographic character. In most exemplar models, allophones are not necessarily tied to the notion of belonging to the same overarching category; that is, there is no need to specify a /b/ category that subsumes all exemplars of [b] and [β̞]. Instead, allophones and contrastive segments are at two ends of a continuum that can be understood as end points of a similarity relationship (Ladd, 2006).

Results in support of distribution-based learning in L1 allophone perception were shown by Maye and Gerken (2001), who demonstrated that the perception of allophonic contrasts can be modified after exposure to an artificial language that contains tokens of the allophones with a certain statistical distribution. They tested the perception of the allophonic contrast between voiced [d] and the voiceless unaspirated [t] in English. Adult native speakers of American English were exposed to either a monomodal or a bimodal distribution of tokens along a continuum between these two sounds. After exposure, participants in the bimodal group performed significantly better than those in the monomodal group in a discrimination task, which suggests that the bimodal group participants had constructed two separate categories, whereas the other group had not.

Peperkamp, Pettinato, and Dupoux (2002) explored the effect of contextual information on distribution-based modifications to L1 allophone categories. They exposed native speakers of French to a bimodal and monomodal distribution of voiced and voiceless uvular trills in their L1. For one group, the trills were presented in context, whereby voicing was the result of assimilation to the following sound. The other group was presented with the same voiced-voiceless trill stimuli, following the same distributions but without the contextual information. The group without context improved their perception of the contrast, whereas the group exposed to the allophones in context did not. Peperkamp et al. argued that the condition without context led to greater improvement because learners relied more on phonetic perception, whereas the other group relied on phonological knowledge. According to phonological theory, both French trills are members of the same phoneme class. In other words, the phonetic differences between the voiced and voiceless versions of the uvular trills were perceivable, but, in the correct phonological context for the realization of the allophones, this perceptibility was diminished. These results indicate that distributional information—in this case, immediate contextual information regarding voicing assimilation patterns—plays a role in perceived noncontrastiveness of allophones. Peperkamp et al.'s results are reflected in the current finding that learners rely on a distribution-based mechanism in their perception of allophones. Furthermore, the results of the current

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

study show that the likelihood of learners' use of contextual cues is a function of language experience.

The data also suggest tentative support for Escudero and Boersma (2004), who proposed that adult L2 learners can incorporate cues used to distinguish categories in their target language but not in their L1. Escudero and Boersma's L1 Spanish speakers learned to use the duration cue in their perception of the English /i/-/I/ contrast, although their L1 relies primarily on spectral information to distinguish among vowels. The learners noticed that English vowels were functionally differentiated by the duration cue and, with increased language experience, L1 Spanish speakers learned to differentiate between these two vowels. According to Escudero and Boersma, L1 Spanish learners of English use a general learning mechanism that interacts with language-specific experience. It is possible that the learners in the current study are also using a similar general learning mechanism. Stops are crosslinguistically less marked as segmental onsets for words and syllables because of their low sonority (Archangeli & Pulleyblank, 1994; Prince & Smolensky, 2004). The transition between low-sonority consonant onsets and high-sonority vowels reflects a general acoustic preference for maximizing sonority distance to render the syllable more salient. Listeners from all groups could be showing a general preference for this crosslinguistic pattern by preferring stops as onsets to stressed syllables. However, given that listeners with more Spanish experience clearly do not associate approximant-initial syllables with stress as often as less experienced listeners, language-specific experience also played a key role.

Another possibility is that listeners transfer cue use from their L1. In English, the alveolar stops /d/ and /t/ are flapped in word-medial unstressed syllables, but /b/ and /g/ do not undergo similar phonological processes. It is possible that the listeners use their knowledge of flapping in English when perceiving the approximant allophones in Spanish. The evidence presented in this study indicates that this does not appear to be the case, however, as the low-intermediate learners and, more particularly, the NESs do not show any preference for stop or approximant onsets as stressed or unstressed, which indicates that they do not associate either of these allophones as being more probabilistically related to stress than the other.

Under an exemplar-based model, such effects arise when listeners rely on information that has been stored and probabilistically drawn upon through exposure to the input. Listeners from the high-intermediate and NSS groups have representations that probabilistically associate stress with stop onsets. These listeners' perception is biased toward perceiving stops and stress, which yields phonotactic sequences that are highly probable in Spanish. For the same reason, perception is biased against hearing stress on approximant-initial syllables. The

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

groups with less Spanish experience have not built up sufficiently dense representations and are thus not biased in one direction or the other.

Language-specific phonotactics have been shown to bias the perception of individual segments. Massaro and Cohen (1983) found that synthetic stimuli ambiguous between /r/ and /l/ tended to be perceived as /r/ when preceded by /t/ and as /l/ when preceded by /s/. Massaro and Cohen argued that perception is biased toward segments that yield the legal clusters /tr/ and /sl/ rather than the illegal clusters /tl/ and /sr/. Similarly, Hallé, Segui, Frauenfelder, and Meunier (1998) found that illegal onset clusters in French were perceived as legal ones. In particular, illegal /dl/ and /tl/ were perceived as legal /gl/ and /kl/, respectively. This phenomenon could contribute to the explanation for the results from Experiment 2, in which the high-intermediate and NSS listeners demonstrated an allophone-induced illusion of stress.

As language experience increases, adult L2 learners are more affected by the contextual cues, or conditioning factors that drive allophonic alternation. Specifically, knowledge of probabilistic, distribution-based information allowed the high-intermediate group to recognize the factors that condition the allophonic alternation. In the present experiments, context effects (i.e., the onset allophone) actually shifted the perception of stress in the high-intermediate learners. Lower proficiency learners did not demonstrate any such effects. This finding suggests that adult L2 speech perception shifts over time and becomes sensitive to the phonotactics of the target language. Similar results were obtained by Dupoux, Kakehi, Hirose, Pallier, and Mehler (1999), who compared French and Japanese native-speaker perception of sequences that were respectively phonotactically legal and illegal in their L1. They found that the phonotactic properties of Japanese (i.e., the very reduced set of syllable types) drove L1 Japanese listeners to perceive illusory vowels inside consonant clusters in VCCV stimuli. French listeners, for whom these sequences were legal, did not perceive such vowels. The current results indicate that stop allophone onsets can induce the perception of stress in L1 Spanish listeners and L2 English listeners with high Spanish proficiency, an illusion consistent with the distributional information found in Spanish but not English.

According to PAM (Best, 1995; Best & Tyler, 2007), contextual factors will change how the target language sound is assimilated into L1 categories. For example, the Spanish /b/ category may be realized as an approximant or as a stop, depending on the context in which it occurs, and PAM predicts that this will affect the way in which the L2 /b/ allophone is assimilated into the L1 /b/ category. The approximant and stop realizations will each assimilate into a different L1 category—or possibly not into any L1 category at all. Thus, context can play a key

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

role in PAM in terms of phonological categories. However, it is less clear as to how these allophones might be assimilated into the L1 on a phonological level. L2 listeners may hear the two target language allophones as separate sounds and may even initially want to assimilate them into two completely separate L1 categories (see Boomershine, Currie Hall, Hume, & Johnson, 2007, for an example of this), but knowledge of target language orthography or classroom instruction will drive learners to classify these allophones as variants of the L1 category.

According to Flege's (1995) SLM, phonetic characteristics of speech sounds are stored and production targets are taken from these stored acoustic memories. L2 speech learning occurs across the life span, causing adaptations and changes to the L1 phonological system. For a new category to be learned, the SLM posits that the listener must notice the difference between the new category and the native language categories. All learning involves the acquisition of positional allophones in the target language—the acquisition of phonological categories is not addressed. The current results lend support to two important hypotheses of the SLM—namely, that learning is possible and will occur as experience with the target language accumulates and that learners store phonetic details in their representations. However, the claim that phonetic differences drive the formation of new sound categories is less clear in terms of allophone acquisition. In the case of the stop-approximant alternation, learners must realize that there are two allophonic variants of the voiced stop category but that these two variants are in contrastive distribution. Thus, some sort of mechanism is required by which differences can be noted (following the SLM), but category unity can still be maintained on an abstract level.

In conclusion, the results of the current study suggest that adult L2 learners use contextual information in their acquisition of target language allophones: The perception of stress was conditioned by the onset allophone and the position in the word, as a function of language experience. In a broader sense, these results point to the availability of a distribution-based mechanism for adult SLA and further suggest that language experience plays an important role in how this mechanism is used over the time course of SLA.

(*Received 22 January 2010*)

NOTES

1. We follow Hualde (2005) and Mártinez-Celdrán (2004) in calling the allophones stops and approximants.

2. It is assumed that L1 English learners of Spanish will not create contrastive categories for the approximant allophones because of the high number of connections between the allophones at the level of orthography, articulation, and perception.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

3. If listeners had been asked for either discrimination or categorization responses, different learners of distinct proficiency levels may have tapped into separate levels of representation. The L1 Spanish speakers may tap into the allophones at either the phonetic (resulting in two categories) or phonological (one category) level. This may also be the case with learners; however, their perception could potentially be based either on the L1 (in which case learners would not hear the contrast) or the L2 (they would hear the contrast but use phonological categories to classify the two variants together).

4. We thank an anonymous *SSLA* reviewer for pointing out this.

5. The totals do not add up to 180 (or to 168 for the high-intermediate group) because 2.5% of the data were discarded due to timing out. A total of 18 trials were discarded: 6 from the NES data, 1 for the low-intermediate group, 4 for the high-intermediate group, and 4 for the NSSs.

6. The totals do not add up to 405 (or to 378 for the high-intermediate group) because 4% of the data were discarded due to timing out. A total of 64 trials were discarded: 12 for the NESs, 18 for the low-intermediate group, 9 for the high-intermediate group, and 25 for the NSSs.

REFERENCES

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano* [Frequency dictionary of the lexical units of Spanish]. Oviedo, Spain: University of Oviedo Press.

Archangeli, D., & Pulleyblank, D. (1994). *Grounded phonology*. Cambridge, MA: MIT Press.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

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

Boersma, P., & Weenink, D. (2008). Praat: Doing phonetics by computer [Computer program]. Retrieved 8 Sept 2008 from http://www.praat.org/

Boomershine, A., Currie Hall, K., Hume, E., & Johnson, K., (2007). The influence of allophony vs. contrast on perception: The case of Spanish and English. In P. Avery, B. Elan Dresher, & K. Rice (Eds.), *Contrast in phonology: Perception and acquisition* (pp. 145–171). Berlin: Mouton de Gruyter.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*, 257–271.

Crystal, D. (1995). *The Cambridge encyclopedia of the English language*. New York: Cambridge University Press.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133–142.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1568–1578.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551–585.

Face, T. L. (2002). Disentangling the necessarily entangled: The phonology and phonetics of Spanish spirantization. *Southwest Journal of Linguistics*, *21*, 56–71.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

Goldinger, S. D. (1997). Words and voices: Perception and production in an episodic lexicon. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 33–66). San Diego, CA: Academic Press.

Goudbeek, M., Cutler, A., & Smits, R. (2006). Supervised and unsupervised learning of multidimensionally varying non-native speech categories. *Speech Communication*, *50*, 109–125.

Hallé, P., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 592–608.

Harris, J. (1969). *Spanish phonology*. 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. (2005). *The sounds of Spanish.* New York: Cambridge University Press.

Jaeger, J. J. (1980). Testing the psychological reality of phonemes. *Language and Speech*, *23*, 233–253.

Johnson, K., & Mullennix, J. W. (Eds.). (1997). *Talker variability in speech processing*. San Diego, CA: Academic Press.

Kazanina, N., Phillips, C., & Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences United States of America*, *103*, 11,381–11,386.

Ladd, D. R. (2006). 'Distinctive phones' in surface representation. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory phonology 8* (pp. 3–26). Berlin: Mouton de Gruyter.

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

Martínez-Celdran, E. (2004). Problems in the classification of approximants. *Journal of the International Phonetic Association*, *34*, 201–210.

Massaro, D., & Cohen, M. (1983). Phonological context in speech perception. *Perception and Psychophysics*, *34*, 338–348.

Maye, J., & Gerken, L. (2001). Learning phonemes: How far can the input take us? In A. H.-J. Do, L. Dominguez, & A. Johansen (Eds.), *Proceedings of the 25th annual Boston University conference on language development* (Vol. 2, pp. 480–490). Somerville, MA: Cascadilla Press.

Maye, J., Werker, J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.

Ortega-Llebaria, 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 T. L. Face (Ed.), *Laboratory approaches to Spanish phonetics and phonology* (pp. 237–255). Berlin: Mouton de Gruyter.

Peperkamp, S., Pettinato, M., & Dupoux, E. (2002). Allophonic variation and the acquisition of phoneme categories. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of 26th annual Boston University conference on language development* (Vol. 2, pp. 650–661). Somerville, MA: Cascadilla Press.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. L. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–157). Amsterdam: Benjamins.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, *46*, 115–154.

Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. (Tech. Rep. No. CU-CS-696-93). Boulder: University of Colorado.

Shea, C., & Curtin, S. (in press). Cue-use and context in L2 allophone production. *Second Language Research*.

Steriade, D. (2007). Contrast. In P. de Lacy (Ed.), *Handbook of phonology* (pp. 139–157). New York: Cambridge University Press.

Tabachnik, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson Education, Inc.

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press

APPENDIX

STRESS DETECTION WORD LIST

Type	Initial stress	Final stress
Bisyllabic	ive, bera, bito, dabo, dida, duno, gita, mave	barad, belen, budan, dagú, gonam, gudin, magú, dabun
Trisyllabic	lúguida, vásigo	dibanó, adigué

https://doi.org/10.1017/S0272263110000276 Published online by Cambridge University Press