Foreign-accented phonetic detail in L1 word processing: L1 Spanish L2 English listeners and English-accented Spanish

Hadley Forst[1], Miquel Simonet[2]

University of Arizona[1, 2]
hforst@arizona.edu, simonet@arizona.edu

ABSTRACT

This study investigates to what extent English-like VOT in Spanish words affects lexical access for L1 Spanish listeners with English as an L2. Will English-accented Spanish words be recognized as easily as words with Spanish-like VOT? L1 Spanish L2 English bilinguals born and raised in Mexico completed an auditory lexical decision task online. Participants heard /p/-initial Spanish words and nonwords featuring English-like (long-lag) VOT or Spanish-like (short-lag) VOT and indicated whether each item was a real word or not. Results suggest that, for these listeners, English-like VOT in Spanish words causes a processing delay at best, with significantly slower response times to words with English-like VOT than words with Spanish-like VOT, and at worst prevents lexical access entirely, as Spanish words with English-like VOT were accepted as real words in fewer than half of the trials. These findings indicate that, for these bilinguals, L2 phonetic detail is not included in L1 word entries.

Keywords: accented speech perception, phonetic detail, lexical representation, bilingualism, Spanish

1. INTRODUCTION

It is well known that accented speech can be challenging to understand. A growing body of evidence suggests that phonetic detail is stored in the lexicon [1-6], which helps explain why phonetic variation, like that in accented speech, affects lexical access [7-9]. When a listener is confronted with phonetic variants which deviate from the ones they are accustomed to hearing, mismatches between the signal and information in the listener's lexical representation or expectations may cause processing delays and interfere with lexical access [10-12]. Low accuracy on word judgment tasks and delayed lexical access have been found in studies with monolinguals listening to foreign-accented speech [10, 13], L1 regional accents [6, 12] and artificially controlled accents [14]. However, when confronted with variation that is familiar, such as after exposure to a particular accent in the lab, or when listening to an accent that one is accustomed to hearing, the processing delay and interference decreases or disappears [6, 10-14].

Studies on pronunciation and dialect variant recognition have reached similar conclusions: experience with certain types of phonetic variants reduces or eliminates any processing delay present for listeners initially unfamiliar with those variants. This experience with variant phonetic forms contributes to the formation of lexical representations that are sensitive to phonetic detail [1-6]. For example, hearing common pronunciation variants of words—such as the casual pronunciation "cenner" for the English word *center*—does not cause a delay in lexical access when compared with hearing the canonical form "center" [5]. Furthermore, stronger activation of the target word may be triggered by the more frequent variant, indicating that lexical entries store detailed phonetic information, and that a single lexical entry can be linked to several phonetic forms. Similarly, there is evidence that, due to their linguistic backgrounds, speakers of some dialect varieties may have multiple representations for dialect variants (such as "slender" for General American English and "slenda" in a regional New York dialect of English) and can activate the intended word in their lexicon via either phonetic variant, whereas speakers of other dialect varieties cannot [6]. For researchers interested in bilingual speech processing and representation, these findings suggest that some bilinguals may possess multiple representations of words in their lexicons, including accented and unaccented phonetic forms, just as some bidialectal speakers do.

Research into second-language (L2) perception of accented speech also reveals a processing benefit associated with accent familiarity [15-17]. Weber, Broersma and Aoyagi [17] found that while listening in L2 English, native (L1) Dutch listeners processed Dutch-accented English words faster than Japanese-accented English words. The authors conclude that linguistic experience with the phonetic features of a particular accent, such as being a speaker and frequent overhearer of the accent in question, allows for faster lexical access and word recognition of accented speech. In a different study of bilingual word recognition, Shea [18] found that L1 English L2 Spanish bilinguals showed similar levels of lexical activation when hearing both English-accented and Spanish-accented variants of Spanish words (words featuring either an intervocalic voiced stop or intervocalic spirant, respectively). L1 English L2 Spanish learners are known to produce the stop instead of the spirant intervocalically, at least in initial stages of acquisition, and may be accustomed to hearing it in the speech of other L2 Spanish speakers. In contrast, the native speakers in Shea's study only showed lexical activation from the Spanish-accented forms (with intervocalic spirant), and not the English-accented forms. These results also suggest that bilinguals store phonetic information in the lexicon: the L1 Spanish listeners were unable to match the signal to their representation because they had insufficient experience with the phonetic variants characterizing English-accented Spanish. In contrast, the L1 English L2 Spanish speakers have had ample experience with English-accented Spanish, being both producers and overhearers of it, and they were able to match the signal to an existing representation.

The present study examines English-accented Spanish word recognition, but with a focus on a different salient characteristic of English-accented Spanish. In word-initial position, L1 English L2 speakers of Spanish often produce an English-like [p^h], with aspiration or long-lag voice onset time (VOT), instead of Spanish-like [p], with short-lag VOT. Thus, L1 English speakers commonly pronounce Spanish words like *pelo* 'hair' as [p^h]elo instead of [p]elo.

Instead of examining the processing of L1-accented L2 speech, the present study focuses on L2-accented L1 speech and tests a new population: L1 Spanish L2 English listeners. Learning more about how bilinguals recognize their L1 as likely produced by native speakers of their L2 will help fill in a gap in the research. The participants in our study are neither producers nor frequent overhearers of English-accented Spanish, as they are L1 Spanish speakers who have lived in Mexico since birth. However, they have had exposure to English, as it is their L2, and may implicitly be familiar with some basic pronunciation differences between Spanish and English. Do L1 Spanish L2 English speakers recognize English-accented Spanish words as readily as Spanish-accented words? If so, this would indicate they had formed lexical representations including English-accented detail. Or, alternatively, will they show a processing delay upon hearing English-accented Spanish words?

L1 Spanish L2 English listeners completed an auditory lexical decision task featuring /p/-initial Spanish words differing in their phonetic detail. Recordings of critical words and pseudowords were manipulated so that each item had two phonetic realizations: one with Spanish-like, short-lag VOT, [p], and one with English-like, long-lag VOT, [p^h]. Due to their relatively low experience with English-accented Spanish, we hypothesized that our L1 Spanish listeners would show longer response times (RTs) for word tokens featuring long-lag VOT than for word tokens featuring short-lag VOT.

2. METHOD

2.1. Participants

Forty-five participants were recruited online through the participant recruitment platform Prolific [20]. Participants were L1 Spanish L2 English speakers who were born and raised in a monolingual Spanish environment in Mexico and had never resided outside of Mexico for more than six months. Scores from the History and Attitudes modules of the Bilingual Language Profile questionnaire [19] indicate that all participants were dominant in Spanish. Responses to the English proficiency module of the BLP [19] indicate a median self-assessment of 4.5 on a scale of 0–6 (*0 = not well*, *6 = very well*) when asked how well they speak, understand, read, and write in English ($M = 4.3$, $SD = 0.9$, $min = 2.25$, $max = 6$).

2.2. Stimuli

Disyllabic /p/-initial Spanish words with penultimate stress were used as auditory stimuli for a lexical decision task. The 20 most frequent words meeting these criteria were chosen from the stimuli search engine NIM [21], based on the relative frequency-per-million-words for each word. To create a balance between 'yes' and 'no' responses in the lexical decision task, 20 disyllabic /p/-initial pseudowords were generated, as well as 160 fillers. The fillers were evenly divided between real words and pseudowords, all disyllabic and beginning with vowels, fricatives, nasals, laterals and stops, including /b/ and /p/.

An English-dominant Spanish-English early bilingual talker (with phonetic training) was recorded in a sound-attenuated booth using professional equipment at a sampling rate of 44.1 kHz. The talker repeated each item in the carrier phrase "__ también es palabra" (*__ is also a word*) three times each in random order. After recording all items without any specific instructions regarding pronunciation, the talker was asked to produce the critical words and pseudowords an additional three times but was instructed to "exaggerate" the VOT of the initial /p/ to imitate an aspirated "English" word-initial [p^h].

Stimuli were resynthesized in Praat [22]. The VOT values of the initial /p/ of the 40 critical items (20 words and 20 pseudowords) were manipulated using a progressive cutback and replacement method [23]. This method begins with recorded minimal pairs, such as [p]elo and [p^h]elo. The VOT was then adjusted systematically for each item such that all [p]-initial items had a VOT of 10 ms (representative of a short-lag, Spanish-like VOT) and [p^h]-initial items had a VOT of 60 ms (representing a long-lag, English-like VOT). Trials were presented in random order, with all critical words, nonwords and fillers included in the same list.

2.3. Procedure

The experimental activities were completed online by participants in Mexico. Participants were recruited through Prolific [20], and the experiment was designed with and hosted by Gorilla [24]. While the quality of online data collection cannot be guaranteed, participants were encouraged to participate from a quiet and distraction-free location and were required to use headphones, which was confirmed through a headphone screening task [25]. After completing a brief demographic screening and passing the headphone check, participants completed an auditory lexical decision task, followed by the Bilingual Language Profile questionnaire [19].

Each trial of the auditory lexical decision task began with the presentation of a red fixation cross in the center of the computer screen for 1 s. When the fixation cross disappeared, an auditory stimulus was presented, and participants responded by keypress. Participants were instructed to decide, for each stimulus, whether they believed the item was a real Spanish word. They indicated their decision by pressing the 'j' key for 'yes' or the 'f' key for 'no' and were instructed to respond as quickly and accurately as possible. Each trial ended upon the key press, or after 3 s if no response was made, and the fixation cross would reappear, signalling the start of a new trial. There were 5 practice trials, followed by 240 trials presented in 2 equal blocks, with one 30 s break in the middle.

3. RESULTS

All participants heard each critical word twice during the lexical decision task, once with short-lag VOT and once with long-lag VOT. Because hearing a word a second time results in speeded lexical access due to priming [26], only responses to the first encounter with each lexical item are discussed in this paper. Additionally, only data from participants with accuracy greater than 80% on the filler items (both words and pseudowords) was included in the following analyses to ensure participant engagement.

3.1. Word Acceptance Rates

An analysis of the proportions of 'yes' responses to real words with differing phonetic detail revealed that, overall, word acceptance rates were low for critical words. English-accented, [p^h]-initial words were judged to be real Spanish words less than half the time, with a mean word acceptance rate of 43.1%, 95% CI [37%, 48%]. Participants rejected words with English-accented VOT more often than they accepted them. For short-lag VOT words, mean acceptance was higher, at 60.8% [55%, 66%], but still rather low—filler words were accepted at a rate of 90.3% [89%, 92%]. A paired *t*-test on logit-transformed word acceptance rates yielded a significant difference between the mean acceptance rate for word tokens presented with short-lag VOT and those presented with long-lag VOT: $M_{diff} = -0.861$, 95% CI [−1.22, −0.5], $t(44) = -4.80$, $p < 0.001$, Cohen's $d_{avg} = -0.85$, 95% CI [−1.25, −0.49], $r = 0.28$. Participants reliably rejected words with English-accented VOT more often than they rejected words with Spanish-accented VOT. Note, however, that word acceptance for all critical words, including Spanish-accented tokens, was surprisingly low relative to filler words.

3.2. RT as a function of VOT

Reaction time measured from stimulus onset was analyzed for 'yes' responses to see if VOT variation affected the speed of lexical access. Responses to short-lag VOT words were faster (1279 ms, 95% CI [1236, 1323]) than responses to words with long-lag VOT (1348 ms, 95% CI [1290, 1406]). Mean RT to filler words was 1216 ms [1175, 1257]. A paired *t*-test revealed that the mean log RT to words with short-lag VOT was significantly faster than mean RT to those with long-lag VOT: $M_{diff} = 0.054$, 95% CI [0.016, 0.091], $t(44) = 2.9$, $p = 0.005$, Cohen's $d_{avg} = 0.42$, 95% CI [0.14, 0.72], $r = 0.75$. Thus, when participants heard Spanish words with long-lag VOT, they took relatively longer to find a match in their lexicon. Note, however, that RTs to filler words were on average even faster ($M = 7.08$, 95% CI [7.05, 7.12]) than to short-lag VOT words.

3.3. RT to long VOT words by word judgment

When words with long-lag VOT were accepted, listeners experienced a relative delay in lexical retrieval. However, words presented with long-lag VOT were rejected nearly 60% of the time as nonwords. Because we are interested in the full picture, it is also important to examine the speed of rejection for these words. Did listeners take just as long to reject [p^h]-initial words as they did to accept them?

Mean RT in trials where long-lag VOT words were accepted as words was slower ($M = 1358$ ms, 95% CI [1288, 1428]) than that in trials where those words were rejected as nonwords ($M = 1258$ ms, 95%

CI [1200, 1316]). A paired *t*-test revealed that mean log RT to words with long VOT was significantly slower for trials in which participants responded ‘yes’ than for trials in which participants responded ‘no’: $M_{diff} = -0.075$, 95% CI [−0.126, −0.023], $t(44) = -2.92$, $p = 0.005$, Cohen’s $d_{\text{avg}} = -0.49$, 95% CI [−0.85, −0.16], $r = 0.35$. Thus, participants were faster when rejecting long-lag VOT words than when accepting them. Not only were participants more likely to reject [pʰ]-initial Spanish words than accept them—recall the acceptance rate for long VOT words was only 43%—but they also took less time to reject these items than they did to accept them.

4. DISCUSSION AND CONCLUSION

This study examined to what extent English-like VOT in Spanish words affects lexical access for L1 Spanish L2 English speakers. It was hypothesized that these listeners would show sensitivity to the differences in phonetic detail between the English-accented and Spanish-accented stimuli, and that the processing of English-accented words would result in a delay when compared with the processing of Spanish-accented words. Nevertheless, since these participants are Spanish-English bilinguals, it was possible that cross-language phonetic interactions would modulate the speed of word processing even in their L1.

An analysis of word acceptance rates revealed that items with long-lag VOT were accepted as words only 43% of the time, indicating that the presence of long-lag VOT interfered with lexical access for these listeners—most English-accented words were not accepted as words at all. In these cases, the mismatch between the phonetic detail in English-accented Spanish words and listeners’ lexical representations was great enough to prevent lexical retrieval. L1 Spanish L2 English bilinguals from monolingual Spanish backgrounds who have never lived amongst L1 English L2 Spanish speakers in a bilingual environment did not easily recognize English-accented Spanish words as real words, at least not in a context like the one in our research study. Additionally, when these English-accented words were recognized as words, they were accepted far slower than their Spanish-accented counterparts, strengthening evidence of interference with lexical access.

The speed of lexical access is modulated by the phonetic detail in the acoustic signal, and these listeners do not appear to have had sufficient experience with English-accented Spanish to have formed functioning representations of Spanish words beginning with [pʰ] that would have allowed for more rapid lexical access of the base word. This suggests that phonetic detail is stored in lexical entries, as processing words with unfamiliar detail delayed lexical access in comparison to processing words with familiar detail. Additionally, an analysis of responses to long-lag VOT word tokens indicated that RT in trials where participants responded ‘yes’ was significantly slower than RT in trials where participants responded ‘no’, revealing that participants were quick to reject English-accented words and slow to accept them. Thus, there was no evidence that L2 phonetic detail was associated with the phonolexical representation of L1 words for these participants, suggesting that phonetic detail is mostly language-specific for these bilinguals.

Evidence from the lexical decision task indicates that phonetic detail is indeed stored in the lexicon, results that align with previous studies [1-6]. For these listeners, the phonetic detail in the English-accented Spanish words delayed lexical access or even prevented it completely. This is not surprising, given that these listeners are L1 Spanish speakers raised and living in Spanish-dominant environments. Previous research points to the role of experience with a given accent, and while these listeners may be implicitly familiar with some phonetic characteristics of English due to their experience as L2 learners, they do not show evidence of having had enough experience with English-accented Spanish to form “accented” lexical representations for L1 words.

This raises the question of whether Spanish-English bilinguals living in bilingual environments would behave differently. L1 English L2 Spanish late bilinguals living in the US are likely to be exposed to English-accented Spanish frequently, both in the speech of their classmates and in their own productions [27]. It is reasonable to hypothesize that they might show more acceptance of English-accented Spanish words if they have had the opportunity to form representations of this type of phonetic variants. There are also many Spanish-English early bilinguals living in the US who were raised bilingual and have had ample experience with the phonetic features of both Spanish and English, with frequent code-switching. Would this population have additional representations of English-accented Spanish words and accept them? Or would they show processing delays? The present study has outlined a background against which to compare the behaviour of these bilingual populations.

5. ACKNOWLEDGMENTS

This work was funded in part through a University of Arizona *Graduate and Professional Student Council Research and Project* (ReaP) *Grant* awarded to the first author.

6. REFERENCES

[1] Bürki, A., Ernestus, M., Frauenfelder, U. H. 2010. Is there only one 'fenêtre' in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *J. Mem. Lang*.

[2] Connine, C., and Pinnow, E. 2006. Phonological variation in spoken word recognition: Episodes and abstractions. *The Linguistic Review*. 23(3), 235–245.

[3] Connine, C., and Ranbom, L. 2004. Production frequency effects in perception of phonological variation. *J. Acoust. Soc. Am*.

[4] Pinnow, E., Connine, C., Ranbom, L. 2017. Processing pronunciation variants: the role of probabilistic knowledge about lexical form and segmental co-occurrence. *J. of Cognitive Psychology.*

[5] Ranbom, L., Connine, C. 2007. Lexical representation of phonological variation in spoken word recognition. *J. Mem. Lang*. 57(2), 273–298.

[6] Sumner, M., Samuel, A.G. 2009. The effect of experience on the perception and representation of dialect variants. *J. Mem. Lang*. 60(4), 487–501.

[7] Dahan, D., Magnuson, J.S., Tanenhaus, M.K. 2001. Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language*.

[8] Ju, M., Luce, P.A. 2006. Representational specificity of within-category phonetic variation in the long-term mental lexicon. *J. Exp. Psychol. Hum. Percept. Perform*. 32(1), 120–138.

[9] McMurray, B., Tanenhaus, M.K., Aslin, R.N. 2002. Gradient effects of within-category phonetic variation on lexical access. *Cognition*. 86(2), B33-42.

[10] Clarke, C.M., and Garrett, M.F. 2004. Rapid adaptation to foreign-accented English. *J. Acoust. Soc. Am*. 116(6), 3647–3658.

[11] Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., Floccia, C. 2012. Linguistic processing of accented speech across the lifespan. *Front. Psychol.* 3(479).

[12] Floccia, C., Goslin, J., Girard, F., Konopczynski, G. 2006. Does a regional accent perturb speech processing? *J. Exp. Psychol. Hum. Percept. Perform.* 32(5), 1276–1293.

[13] Bradlow, A., Bent, T. 2008. Perceptual adaptation to non-native speech. *Cognition*. 106(2), 707–729.

[14] Maye, J., Aslin, R.N., Tanenhaus, M.K., The weckud wetch of the wast: lexical adaptation to a novel accent. *Cogn. Sci.* 32(3), 543–562.

[15] Larraza, S., Samuel, A.G., Oñederra, M.L. 2016. Listening to accented speech in a second language: First language and age of acquisition effects. *J. Exp. Psychol. Learn. Mem. Cogn*. 42(11), 1774–1797.

[16] Larraza, S., Best, C.T. 2018. Differences in phonetic-to-lexical perceptual mapping of L1 and L2 regional accents. *Bilingualism: Language and Cognition*. 21(4), 805–825.

[17] Weber, A., Broersma, M., Aoyagi, M. 2011. Spoken-word recognition in foreign-accented speech by L2 listeners. *J. Phon.* 39(4), 479–491.

[18] Shea, C. 2017. L1 English / L2 Spanish: Orthography–phonology activation without contrasts. *Second language research*. 33(2), 207–232.

[19] Birdsong, D., Gertken, L.M., Amengual, M. 2012. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. *COERLL, University of Texas at Austin*.

[20] Prolific. www.prolific.co.

[21] Guasch, M., Boada, R., Ferré, P., Sánchez-Casas, R. 2013. NIM: A Web-based Swiss Army knife to select stimuli for psycholinguistic studies. *Behav. Res. Methods*. 45, 765–771.

[22] Boersma, P., Weenink, D. 2020. Praat: doing phonetics by computer.

[23] Winn, M.B. 2020. Manipulation of voice onset time in speech stimuli: A tutorial and flexible Praat script. *J. Acoust. Soc. Am.*, 147(2), 852-865.

[24] Anwyl-Irvine, A.L., Massonnié, J., Flitton, A., Kirkham, N. Evershed, J.K. 2020. Gorilla in our midst: An online behavioral experiment builder. *Behav. Res. Methods*. 52(1), 388–407.

[25] Milne, A.E. *et al.* 2021. An online headphone screening test based on dichotic pitch. *Behav. Res. Methods*. 53(4), 1551–1562.

[26] Forster, K., Davis, C. 1984. Repetition priming and frequency attenuation in lexical access. *J. Exp. Psychol. Learn. Mem. Cogn*. 10(4), 680–698.

[27] Tyler, M.D. 2019. PAM-L2 and phonological category acquisition in the foreign language classroom. In: Nyvad, A.M., Hejná, M., Højen, A., Jespersen, A.B., Sørensen, M.H. 2019. *A Sound Approach to Language Matters- In honor of Ocke-Schwen Bohn*. Aarhus: Aarhus University, 607–630.