Bilingualism: Language and Cognition 22 (5), 2019, 1101–1122 © Cambridge University Press 2018 doi:10.1017/S136672891800086X

Now you hear it, now you don't: Malleable illusory vowel effects in Spanish–English bilinguals*

MATTHEW T. CARLSON
The Pennsylvania State University

(Received: December 02, 2017; final revision received: June 20, 2018; accepted: July 01, 2018)

Spanish speakers tend to perceive an illusory [e] preceding word-initial [s]-consonant sequences, e.g., perceiving [stið] as [estið] (Cuetos, Hallé, Domínguez & Segui, 2011), but this illusion is weaker for Spanish speakers who know English, which lacks the illusion (Carlson, Goldrick, Blasingame & Fink, 2016). The present study aimed to shed light on why this occurs by assessing how a brief interval spent using English impacts performance in Spanish auditory discrimination and lexical decision. Late Spanish–English bilinguals' pattern of responses largely matched that of monolinguals, but their response times revealed significant differences between monolinguals and bilinguals, and between bilinguals who had just completed tasks in English vs. Spanish. These results suggest that late bilinguals do not simply learn to perceive initial [s]-consonant sequences veridically, but that elements of both their phonotactic systems interact dynamically during speech perception, as listeners work to identify what it was they just heard.

Keywords: bilingualism, phonotactics, speech perception, illusory vowel effects

Introduction

Bilinguals' extensive experience with two systems for organizing acoustic and articulatory space presents a uniquely complex instantiation of a general problem in speech perception. Namely, the same acoustic material may be compatible with many interpretations, and the same phonological entities may vary acoustically depending on an enormous range of factors. For bilinguals, this many-to-many mapping between acoustic material and phonological representations may play out in different, partially overlapping, or directly conflicting ways in their two language systems.

There is, of course, ample evidence that what may correspond to two systems in a bilingual's experience is not entirely separate in the mind. Interactions in phonetic category structure (Flege, 2003, 2007), prosody (Dupoux, Pallier, Sebastián-Gallés & Mehler, 1997), and phonotactics (Altenberg & Cairns, 1983; Anisfeld, Anisfeld & Semogas, 1969; Cohen, Tucker & Lambert, 1967; Ernestus, Kouwenhoven & van Mulken, 2017) countenance substantial permeability of sound systems. This can lead to interference, but also to enhanced perception of phonetic contrasts, accented speech (Bent & Bradlow, 2003; Chang, 2016; Chang & Mishler, 2012; Hanulíková & Weber, 2012), and word segmentation (Hanulíková, Mitterer & McQueen, 2011; Weber & Cutler, 2006) compared to monolinguals. Most of this work has focused on how the L1 influences the L2, but there is also evidence for L2 effects on the L1 (Caramazza, Yeni-Komshian & Zuriv, 1974; Flege & Eefting, 1987; Namjoshi, Tremblay, Spinelli, Broersma, Martínez-García, Connell, Cho & Kim, 2015; Tice & Woodley, 2012). Nonetheless, bilinguals are also remarkably good at maintaining the distinctness of language-specific sound patterns (Fowler, Sramko, Ostry, Rowland & Hallé, 2008; Gonzales & Lotto, 2013).

What, then, is the nature of bilinguals' phonological systems, and how do they navigate both the overlap and distinctness of the two source systems? The evidence for both permeability and separateness suggests that bilinguals possess neither a single system nor two separate systems, but rather a hybrid system (Cook & Wei, 2016; Goldrick, Putnam & Schwarz, 2016; Hall, Cheng & Carlson, 2006; Otheguy, García & Reid, 2015) in which corresponding structures across languages may be separate but linked, although how this plays out across specific structures and contexts is not yet clear. The present study seeks to shed light on this linkage by exploring its consequences for bilinguals' speech perception when abstract constraints on sound sequences lead to conflicting interpretations of the same acoustic material.

This question will be explored by examining Spanish–English bilinguals' perception of word-initial [s]-consonant clusters (#sC). These clusters are

* The research reported here was funded through a grant from the National Science Foundation (OISE-0968369). I also gratefully acknowledge the help of Alex McAllister and Manuel Pulido Azpíroz in the construction of stimuli, Teresa Bajo for providing lab space in Granada, and Giuli Dussias, Matt Goldrick, Mike Putnam, Katharina Schuhmann, and Frances Blanchette for their invaluable help and feedback.

Address for correspondence:
442 Burrowes Bldg., Dept. of Spanish, Italian, and Portuguese, The Pennsylvania State University, University Park, 16802
mtc173@psu.edu

impossible in Spanish, and foreign loanwords with #sC are obligatorily adapted via [e] prothesis, e.g., *snob* was adopted as *esnob*. In English, however, words beginning with #sC are abundant. This conflict in phonotactic patterning may lead bilinguals to perceive #sC clusters differently, compared to Spanish monolinguals.

Spanish monolinguals tend to perceive #sC sequences as if they were #esC. They report hearing [e], but not other vowels, preceding #sC (Cuetos et al., 2011), and they have greater difficulty discriminating #VsC stimuli with a short initial vowel from identical stimuli bearing a robust [e], vs. a robust [a] (Carlson et al., 2016). Moreover, in lexical decision they readily accept #esC words missing their initial vowel, e.g., [skwela] for *escuela* "school", but not words missing other initial vowels, e.g., [spiɾina] for *aspirina* "aspirin" (Hallé, Segui, Dominguez & Cuetos, 2013), a finding that extends even to visual word processing (Hallé, Dominguez, Cuetos & Segui, 2008). This mirrors perceptual illusions observed in other languages (Dupoux, Kakehi, Hirose, Pallier & Mehler, 1999; Pitt, 1998; Polivanov, 1931), where listeners perceive sound sequences that are unrepresentable in their phonology as the most likely representable alternative (Berent & Lennertz, 2010; Berent, Lennertz & Rosselli, 2012; Berent, Lennertz, Smolensky & Vaknin-Nusbaum, 2009; Berent, Steriade, Lennertz & Vaknin, 2007; Kabak & Idsardi, 2007). Careful manipulation of coarticulatory features (Dupoux, Parlato, Frota, Hirose & Peperkamp, 2011) and similarity to lexical items (Dupoux, Pallier, Kakehi & Mehler, 2001; Hallé et al., 2013) confirms the phonological nature of these illusions. However, the specific percept can also be shaped by acoustic details (Davidson, 2011; Davidson & Shaw, 2012), perceptual compensation (Gaskell, 2001), other phonological processes (Durvasula & Kahng, 2015), and similarity to known lexical items (Ganong, 1980; Samuel, 1981).

Unlike monolinguals, early Spanish–English bilinguals, especially those dominant in English, are unlikely to report hearing [e] when no vowel is present, and they are less susceptible than monolinguals to the effects of perceptual repair on auditory discrimination (Carlson et al., 2016). This apparent weakening of the perceptual illusion may reflect interacting phonotactic systems, but it could simply be that the early bilinguals never acquired the illusion. Evidence that conflicting phonotactic patterns can jointly influence perception is also found in Japanese–Brazilian Portuguese bilinguals, but only in explicit vowel identification, and not in a more implicit test of perception (Parlato-Oliveira, Christophe, Hirose & Dupoux, 2010), making the nature of any interaction between phonotactic systems unclear.

What happens in late Spanish–English bilinguals, who encounter English phonotactics after acquiring a stable Spanish system? Most saliently, their English strongly reflects filtering through Spanish [e]-prothesis, e.g., *school* is treated as [eskul], in both production (Abrahamsson, 1999; Carlisle, 1991, 1999; Daland & Norrmann-Vigil, 2015) and perception (Freeman, Blumenfeld & Marian, 2016). However, they nonetheless rapidly develop awareness that #sC clusters are possible in English (Altenberg, 2005), and can learn the relative well-formedness of different #sC clusters (Lentz & Kager, 2015). L1 filtering therefore does not altogether prevent learning of conflicting phonotactic patterns, allowing the possibility of changes in late bilinguals' L1 Spanish speech perception. This prediction is borne out by auditory discrimination data (Carlson, published online April 9, 2018). Late bilinguals, like early bilinguals, outperformed monolinguals slightly, but more substantial effects occurred in response times. Late bilinguals were slower when the two phonotactic systems conflicted directly, i.e., when the Spanish illusion was expected to render the stimuli maximally confusable, and this effect was larger for bilinguals currently immersed in English, but completing the experiment in Spanish.

It makes sense to attribute these changes in Spanish speech perception to knowledge of the English pattern, but how exactly would this work? On the one hand, experience hearing English #sC words may allow even late bilinguals to retune their perception of acoustic #sC sequences in favor of the veridical percept, during intense contact with English, e.g., immersion (Carlson, published online April 9, 2018). On the other hand, the now massive evidence for co-activation of two systems in bilinguals at all levels of linguistic structure (e.g., Hartsuiker, Pickering & Veltkamp, 2004; Kroll, Bobb & Wodniecka, 2006; Kroll & Gollan, 2014; Lauro & Schwartz, 2017) suggests that changes to L1 speech perception may reflect real-time interaction between the two systems. Given a view of speech perception that involves the evaluation of multiple candidates against both incoming information from the environment and the expectations of the perceiver (Apfelbaum, Bullock-Rest, Rhone, Jongman & McMurray, 2014; Apfelbaum & McMurray, 2014; Feldman, Griffiths & Morgan, 2009; Kleinschmidt & Jaeger, 2015; Norris, McQueen & Cutler, 2016), bilinguals' perception of #sC would reflect competition between #esC, consistent with Spanish phonotactics, and the veridical percept made available through English phonotactics.

We are now in a position to state a more specific hypothesis concerning late Spanish–English bilinguals' perception of #sC. Namely, while they learn to perceive #sC veridically, this does not displace the Spanish-like perceptual illusion. Rather it provides a new phonetic representation for the same acoustic material, affecting speech perception by introducing competition between alternative representations (rather than from relaxation or retuning of category boundaries, cf. Gonzales & Lotto, 2013; Llanos & Francis, 2017). That is, Spanish–English bilinguals may represent acoustic #sC as either #esC or as #sC, while only #esC is possible for Spanish monolinguals.

If this is true, then manipulating the degree of competition between the Spanish- and English-like competitors over the short term, by activating English, inhibiting Spanish, or both (Antoniou, Best, Tyler & Kroos, 2010, 2011; Dijkstra, van Jaarsveld & ten Brinke, 1998; Lauro & Schwartz, 2017), should lead to changes in the strength of the Spanish perceptual illusion (cf. Athanasopoulos, Bylund, Montero-Melis, Damjanovic, Schartner, Kibbe, Riches & Thierry, 2015; Chang, 2012, 2013; Grosjean, 1989; Grosjean & Miller, 1994; Olson, 2013; Sancier & Fowler, 1997). If not, then the illusion should be stable over the short term.

To manipulate the degree of competition between English and Spanish in the present experiment, late Spanish–English bilinguals performed tasks (which did not involve listening to speech) either in English (henceforth the English-switch group) or Spanish (the no-switch group) immediately before completing speech perception tasks in Spanish. The focus was on late bilinguals to ensure that a stable L1 system had been established prior to learning English, and the bilinguals were currently residing in a Spanish-dominant environment, to minimize baseline contact with English. Similar short-term manipulation of language context has been shown to modulate crosslinguistic competition in lexical processing (Elston-Güttler & Gunter, 2008; Elston-Güttler, Gunter & Kotz, 2005) and speech production (Balukas & Koops, 2015; Goldrick, Runnqvist & Costa, 2014; Olson, 2016a, 2016b; Simonet, 2014).

This manipulation was carried out in two experiments, using auditory discrimination (Carlson, published online April 9, 2018; Carlson et al., 2016) and auditory lexical decision (Dupoux et al., 2001; Hallé et al., 2013) to explore how bilinguals' perception of #sC plays out under different processing demands and decision criteria. In discrimination, stimuli are compared directly, whereas lexical decision tests the acceptability of single stimuli as tokens of known lexical items. This contrast in task conditions may further modulate the role and relevance of the veridical percept, shedding further light on how bilinguals navigate their interacting phonotactic systems.

Experiment 1: AX discrimination

Methods

Participants

Thirty-two late Spanish–English bilinguals (mean age 25.5 years sd = 6.8) were recruited at the University of Granada, Spain, through flyers posted on campus.[1] The data from the 14 Spanish monolinguals in Carlson et al. (2016), which were collected in the same location using the same materials, were used as the monolingual baseline. The monolinguals likely had taken English classes in school, but had little experience or ability to use it.[2] Sample size was determined on the basis of prior studies using similar methods (e.g., Carlson et al., 2016).

Language proficiency was assessed using cloze tests adapted from the DELE (Spanish; Ministry of Education, Culture, and Sport of Spain, 2006) and MELICET (English; English Language Institute, 2001), and verbal fluency tasks in both languages. For the latter, participants had 30s to produce as many words as possible in each of a series of categories (clothing, animals, furniture, and fruits), scored as the total number of unique, category-appropriate words produced. Participants also completed a language background questionnaire (Li, Zhang, Tsai & Puls, 2014).

The late bilinguals were first exposed to English in school around age 7, but this was probably also the case for the monolinguals, who grew up in the same school system. The age at which the bilinguals first began using English outside the classroom (around age 14) is therefore probably a better measure of when they learned to use English.

Proficiency scores, self-reported daily English usage, and age of acquisition statistics are shown in Table 1. No group differences emerged on any of these measures (all $p > .18$).

Materials

Two discrimination paradigms are commonly used to study perceptual illusions such as the one examined here, AX and ABX. AX was chosen here, for several reasons. First, it allows direct comparison with the results of Carlson et al. (2016) and Carlson (published online April 9, 2018), whose materials were adopted here. Second, AX can be configured to focus on fine acoustic differences (as described below). Although word-sized

[1] Three additional bilinguals were tested, but two were excluded because of high proficiency in an additional language, and one due to equipment failure. Three participants had knowledge of Catalan, Galician, or European Portuguese, which treat #sC clusters similarly to Spanish. These participants were retained, but repeating the analyses without them yielded an identical pattern of results.

[2] One reviewer suggested that including an English monolingual group might have been useful, because the availability of both #sC and #VsC in English might lead to competition, as predicted for the bilinguals. However, this was not done for two main reasons. First, the dominant representation in English (based on frequency) is #sC, and English speakers thus have no reason to perceive material that is not present. Second, the stimuli were Spanish-accented nonwords, which might lead English monolinguals to attend more to acoustic detail, obscuring any competition effects. It would therefore have been unclear how to interpret any differences between English monolinguals and the remaining groups.

Table 1. *Participant characteristics (mean (sd)) for bilinguals in Experiment 1.*

	No-switch (n = 17)	English-switch (n = 15)
Gender	11 female	10 female
DELE	.86 (.11)	.88 (.05)
Spanish verbal fluency	54.07 (7.40)	55.00 (8.10)
MELICET	.73 (.16)	.70 (.16)
English verbal fluency	36.40 (7.15)	37.18 (6.66)
Daily English use (hours)	3.70 (4.31)	2.48 (1.90)
Age of first exposure to English (years)	7 (4)	7 (2)
Age began using English outside classroom	16 (5)	13 (6)

stimuli such as in this study may increase difficulty, making participants more likely to encode acoustic detail via phonetic categories (Davidson & Shaw, 2012), this encoding is required in ABX, due to the memory demands of comparing three stimuli. AX thus provides a better test of how the availability of different phonetic representations changes the interplay of acoustic detail and phonetic representation, which is the primary focus here.

One criticism of AX is that listeners may apply different thresholds for detecting differences, making it hard to distinguish participants' ability to discriminate from their willingness to ignore the difference. However, the goal here is not to determine whether bilinguals can detect the relevant acoustic details (changes relative to the monolingual baseline would imply that they can), but rather how they balance acoustic correctness (favored by English), with linguistic correctness in Spanish (the language of the task), with the prediction that bilingualism and recent English use will boost the importance of acoustic correctness.[3]

Participants were told that they would hear pairs of possible Spanish words and asked to decide whether they were "identical" or not. The pseudowords had the form VsCid. The V was either [e] (the default prothetic vowel) or [a] (the next most common vowel in this position), and the C was one of those that follow #Vs in existing Spanish words, [b, d, g, p, t, k, f, m, n, l], resulting in 10 pairs of pseudowords differing only in their initial vowel, e.g., [aspið], [espið]. The pseudowords were recorded by a female native speaker of Mexican Spanish, with Spanish pronunciation (e.g., spirantization of final /d/ and final stress).[4] Stimuli from only one speaker were used in order to focus attention on acoustic detail as much as possible.

Two stimuli were created from each pseudoword by removing a portion of the initial vowel, leaving either 10 periods of phonation (about 40 ms, just under half its original duration), or 2.5 periods. Thus, in half the stimuli the initial vowel was long enough to be unambiguous, and in half it was too short to reliably determine its quality, or even to unambiguously distinguish it from a cough or other non-speech sound. Stimuli with the initial vowel completely removed were not included, because the bilinguals in Carlson et al. (2016) showed little evidence of perceiving an illusory vowel when no vocalic material was present, and to avoid making the task too long or complicating comparison with the earlier results.

This yielded 10 sets of 4 stimuli sharing the same consonant following [s]. On each trial two stimuli from the same set were presented, such that the only differences involved the quality ([a] or [e]) and/or duration of the initial vowel (longer vs. shorter).[5] Since there are 6 ways to select two different stimuli from each set, and two ways of ordering each pair, this resulted in 120 AX trials in which the two stimuli were acoustically different. In an additional 120 trials, each of the 40 stimuli was paired with itself 3 times, such that exactly half the trials contained acoustically identical stimuli. The order of trials was randomized by participant. This, together with the generally subtle differences between stimuli and a brief ISI (250 ms, cf. Davidson, 2011) was designed to focus attention on detecting fine acoustic differences. Table 2 provides an example of the trials generated from one of the sets.

The critical AX pairs were the 80 trials in which the duration of the initial vowels differed, because this

[3] The 4IAX would address this problem, but it is difficult to measure RT using this task.

[4] While the speaker's variety of Spanish differed from that of the participants, the stimuli did not include any salient dialect features. The production of coda /s/ as [s] differs from casual Andalusian Spanish, but is considered normative.

[5] Since pairs with different initial vowels were taken from different recorded tokens, rather than being cross-spliced, slight, subphonemic differences occurred following the vowel as well. This is accounted for in the analysis, but it did not impact the results.

Table 2. *Example of trials generated from one set of stimuli sharing the same post-[s] consonant. The forward slash,* /espid*, indicates the longer (half-length) vowel, and parentheses around the vowel,* (e)spid*, indicates the shorter vowel. A ≠ X pairs were presented in both orders, and A = X pairs were presented 3 times each, in order to achieve equal proportions of A = X and A ≠ X trials.*

Trial type		Pairs
Critical	A ≠ X: Same vowel	/espid - (e)spid /aspid - (a)spid
	A ≠ X: Different vowel	/espid - (a)spid /aspid - (e)spid
Control	A ≠ X: Longer vowels	/espid - /aspid
	A ≠ X: Shorter vowels	(e)spid - (a)spid
	A = X (repeated 3 times each)	/espid - /espid /aspid - /aspid (e)spid - (e)spid (a)spid - (a)spid

allows us to determine whether the shorter vowel is more difficult to discriminate from the longer [e] (the longer-[e] pairs) than from the longer [a] (the longer-[a] pairs), as predicted by Spanish phonotactics. The remaining trials served as controls, indicating baseline performance when discrimination was expected to be easy (the 20 pairs with longer vowels of contrasting quality), difficult (the 20 pairs with shorter vowels of contrasting quality), or impossible (the 120 acoustically identical pairs).

Each trial began with a centered fixation (+), with the first stimulus played concurrently on studio quality headphones at a comfortable level. A 250 ms pause preceded the second stimulus. Participants responded on a button box, followed by a 500 ms blank screen before the next trial. Response times (RTs) were measured from the onset of the second stimulus, i.e., the locus of the experimentally manipulated initial vocalic material. The first 16 trials were considered practice.

Procedure and prior language context manipulation

The experimenter was a near-native speaker of Castilian Spanish and native speaker of American English. Interaction occurred in the language of whichever task was current. Participants gave informed consent in Spanish. To manipulate prior language context, with half the bilinguals (the English-switch group, randomly assigned), the experimenter then switched to English, and participants completed the MELICET and English verbal fluency task. With the no-switch group, interaction remained in Spanish, and participants completed the DELE and Spanish verbal fluency. This took 10–15 minutes. All participants followed the experimenter's language choice for incidental interaction at all times.

After this, the experimenter signaled a switch back to Spanish (for the English-switch group) and the AX task began. Subsequently, participants completed the DELE/MELICET and verbal fluency in the other language, followed by the background questionnaire. Testing occurred in a quiet room at the University of Granada.

Predictions

Spanish phonotactics should cause the shorter vowel to be perceived as [e], leading to more "same" responses on longer-[e] pairs vs. longer-[a] pairs. Knowing English, however, is expected to allow bilinguals to perceive the shorter vowel items more veridically, facilitating discrimination of the longer-[e] pairs (Carlson et al., 2016). If this effect reflects competition between veridical (English-like) and repaired (Spanish-like) phonetic representations, and not merely the retuning of speech perception in an English-dominant environment, then activating English via the verbal fluency and cloze tasks should also lead to longer response times for longer-[e] pairs in the English-switch group, as it did for the bilinguals immersed in English (Carlson, published online April 9, 2018), and possibly greater accuracy for longer-[e] pairs. In addition, the effects of switching from English might decay during the task, as has been found for lexical access (Elston-Güttler & Gunter, 2008; Elston-Güttler et al., 2005), leading to changes in group differences as a function of trial number.

Results

Prior to analysis, 74 trials with RTs longer than 4000 ms were removed from the bilinguals' data (1% of the data). There were no trials with RT < 200 ms, so no lower cutoff was applied.

Accuracy

Since accuracy is a binary dependent variable, it was analyzed using mixed effects logistic regression, fit using the lme4 package in R (Bates, Maechler, Bolker & Walker, 2015; R Core Team, 2016). Significance was assessed via leave-one-out model comparison using likelihood-ratio tests.

The participant groups (monolingual, no-switch bilingual, English switch bilingual; henceforth language-Group) were first compared on the three subsets of control trials (see Table 2). Mean rates of *same* responses are shown in Table 3. All groups were near ceiling when the A and X were acoustically identical, and discrimination was uniformly accurate (few *same* responses) when both stimuli had longer vowels of different quality, where discrimination was expected to be easiest. Separate mixed-

Table 3. *Proportion of* same *responses for control AX trials, with 95% CIs computed by nonparametric bootstrap.*

	A = X	Diff. vowel both longer	Diff. vowel both shorter
Spanish Monolingual	.96 [.95, .98]	.19 [.10, .30]	.68 [.61, .76]
Bilingual, no-switch	.94 [.92, .96]	.11 [.07, .16]	.59 [.51, .66]
Bilingual, English-switch	.95 [.92, .97]	.13 [.07, .20]	.56 [.47, .64]

effects logistic regressions for these control trial types revealed no differences between groups (all $p > .15$).[6] For control trials with two shorter vowels of different quality, where discrimination was expected to be difficult, monolinguals responded *same* at a higher rate than chance, and bilinguals responded approximately at chance. This group difference was significant (in a mixed-effects logistic regression, $\beta = -0.50$, SE = 0.24, $\chi 2(1) = 4.23$, p = .040).

In the analysis of the critical trials the fixed effects of primary interest were longerV (longer-[a] vs. longer-[e] pairs) and languageGroup (monolingual, no-switch bilingual, English-switch bilingual). LongerV tested the discriminability of the shorter vowels from each longer vowel, i.e., the strength of Spanish perceptual repair, and its interaction with languageGroup tested whether bilingualism or recent English use modulated the strength of the perceptual repair. An additional factor (matchingV) was added to capture sensitivity to slight acoustic differences in AX pairs with differing initial vowel qualities (see Table 2), which led to better discrimination of these pairs (see Table A1), but its interactions with longerV and languageGroup were not significant (all p > .20), and as it does not bear directly on the perceptual illusion, it is not commented on further. To probe for short-term decay of the English-switch effect, trialNumber (as a continuous variable) and its interactions with longerV and languageGroup were explored but no significant effects or interactions emerged (all p > .12) and it was omitted from the final model. Effects of the consonant following the [s] were explored visually, but no salient trends emerged. The maximal random effects structure by-participant was included. Complete estimates of the fixed effects, and further details about the coding of predictors and the random effects structure are reported in Appendix A, Table A1.

Accuracy on the critical pairs is shown in Figure 1. Note that the y-axis shows the proportion of *same* responses, i.e., when participants failed to discriminate the pair. Discrimination accuracy was substantially worse (more *same* responses) when the longer vowel was [e], compared to [a], confirmed by the significant main effect

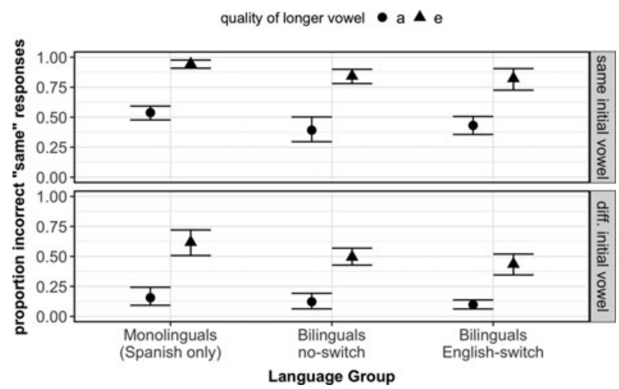

Figure 1. Proportion of incorrect *same* responses to critical AX pairs, with 95% CIs computed by nonparametric bootstrap. Top panel shows pairs with the same initial vowel quality, which are acoustically identical apart from the duration of the vowel, and the bottom panel shows pairs with a different initial vowel.

of longerV ($\beta = 2.38$, SE = 0.15, $\chi 2(1) = 93.37$, p < .0001). Bilinguals discriminated the critical pairs better overall than monolinguals, confirmed by the significant effect of bilingualCode ($\beta = -0.74$, SE = 0.22, $\chi 2(1) = 10.19$, p = .001), but while the size of the longerV effect was numerically smaller for bilinguals, the interaction of bilingualCode with longerV did not reach significance ($\beta = -0.49$, SE = 0.31, $\chi 2(1) = 2.41$, p = .12). No differences were found between the no-switch and English-switch groups (all p > .38).

Response time

RTs were analyzed using linear mixed effects regression, using the same model structure as the accuracy results. RTs were negative reciprocal transformed (x1000), which better approximated a normal distribution than a log transformation, as confirmed by inspection of q-q plots. Acoustically correct and incorrect responses were analyzed in separate models, but no effects or interactions of longerV were found for incorrect responses, and they will not be considered further. As above, effects of trialNumber were explored. A significant main effect of trialNumber emerged, whereby participants responded more quickly as the task progressed, but it did not interact

[6] The fixed effects factor was languageGroup, and random intercepts by-participant were included.

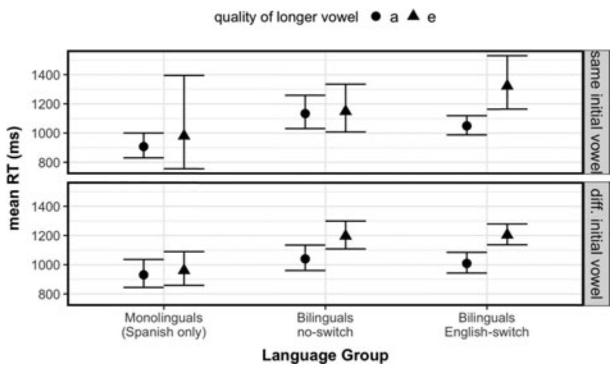

Figure 2. Mean RT (in milliseconds) for correct *different* responses, with 95% CIs. Top panel shows critical AX pairs with the same initial vowel quality, which are acoustically identical apart from the duration of the vowel, and bottom panel shows pairs with a different initial vowel.

with other predictors (all $p > .17$) and only the main effect was retained. Complete fixed effects for both models are given in Appendix A, Table A2.

Mean RTs for correct responses (where participants correctly discriminated the stimuli) are shown in Figure 2. A three-way interaction of longerV by bilingualCode by matchingV ($\beta = 0.24$, SE = 0.09, $\chi 2(1) = 6.62$, p = .010), and a two-way interaction of longerV by priorLangCode ($\beta = 0.08$, SE = 0.04, $\chi 2(1) = 3.90$, p = .048) were found. To unpack these interactions, consider first the monolingual vs. bilingual results. For monolinguals, the completely overlapping CIs reveal no influence of the longer vowel's quality on correct RTs, but bilinguals consistently responded more slowly when the longer vowel was [e], except for the no-switch group's responses when the initial vowels in the pair were the same. Separate models of the monolinguals' and bilinguals' data (with Bonferroni-corrected $\alpha = .025$ for two comparisons) confirmed the lack of a longerV effect for monolinguals ($\beta = 0.08$, SE = 0.07, $\chi 2(1) = 1.32$, p = .25), and a significant longerV effect for bilinguals ($\beta = 0.10$, SE = 0.02, $\chi 2(1) = 21.26$, $p < .0001$).

Turning to the comparison between the two bilingual groups, separate models (with Bonferroni-corrected $\alpha = .025$) for the English-switch and no-switch groups showed a highly significant longerV effect in the English-switch group ($\beta = 0.15$, SE = 0.03, $\chi 2(1) = 17.96$, $p < .0001$), but none in the no-switch group ($\beta = 0.06$, SE = 0.03, $\chi 2(1) = 4.72$, p = .030). However, despite the nonsignificant three-way interaction of longerV by priorLangCode by matchingV in the full model (p = .25), inspection of the CIs in Figure 2 suggests that the two bilingual groups differed from each other primarily on pairs with matching initial vowels.

To summarize, bilinguals responded more slowly overall in the AX task, and when they correctly discriminated the critical pairs, they were slower to discriminate longer-[e] pairs than longer-[a] pairs, with a larger effect being observed when the English proficiency tasks were given immediately prior to the AX task.

Discussion

The two dependent variables in Experiment 1 tell somewhat different stories about the influence of English on Spanish–English bilinguals' perception of #sC. The accuracy results suggest a stable illusory vowel effect across all groups, reflected in substantially poorer discrimination when the longer vowel was [e]. Bilinguals' small, generalized improvement in discrimination may simply reflect greater caution or attention to acoustics (consistent with their overall slower performance and superior discrimination of the control trials with two short vowels). In contrast, RTs revealed significant effects of bilingualism and immediate prior use of English on listeners' sensitivity to the longer vowel's quality. Interestingly, it was the bilinguals, primarily in the English-switch group, whose RTs depended on the longer vowel's quality, being slower when it was [e].[7]

Experimentally introducing a brief interval of L2 English use in an otherwise Spanish-speaking environment thus produced the same results that Carlson (published online April 9, 2018) found for bilinguals during English immersion. This provides strong experimental evidence for the hypothesis that these effects reflect coactivation of the English and Spanish systems in real-time processing, and not merely retuning of perception through long-term exposure to English.

Let us consider how the slower responses to longer-[e] pairs observed in the English-switch group support this conclusion. First, the presence of an effect shows that the bilinguals, particularly after using English, are sensitive to the subtle distinction in vowel length on these trials, even if it does not impact their final decisions. The lack of an RT effect in the monolinguals does not necessarily mean that they could not detect the difference, but, if they could, they appear to have ignored it completely.

Moreover, (as Carlson, published online April 9, 2018, points out) what makes the longer-[e] pairs special is that Spanish and English phonotactics point to conflicting responses in this task, suggesting that competition drives the present findings (cf. Altenberg & Cairns, 1983; Goldrick et al., 2014). To see how this plays out, consider how each phonotactic system contributes to perception of the stimuli with shorter initial vowels. In Spanish, it is impossible to represent the stimulus as #sC, and the

[7] To be fair, monolinguals produced few correct RTs to longer-[e] pairs with the same initial vowel quality. Nonetheless, even when they did – i.e. when the initial vowels were different – monolinguals' RTs showed no hint of sensitivity to the longer vowels' quality.

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

vowel is overwhelmingly more likely to be [e] than [a]. In English, however, #sC is also possible and, if anything, more likely, since #VsC is relatively infrequent in English (Carlson et al., 2016). When the longer vowel is [a], there is no conflict – neither #esC nor #sC matches #asC, and both languages support discrimination. When the longer vowel is [e], however, Spanish favors a *same* response, while English favors discrimination.

Bilinguals' slower responses when correctly discriminating longer-[e] vs. longer-[a] pairs can thus be interpreted as reflecting the need to resolve competition between the responses associated with language-specific representations of the short vowel stimulus, and not merely an improved ability to attend to acoustic detail.[8] This resembles results showing that bilinguals recognize interlingual homographs more slowly in unilingual lexical decision, when real but non-target language words must be rejected, but more quickly in bilingual lexical decision, where words in either language are to be accepted (Dijkstra et al., 1998). The parallel is that responses are fast when both languages point to the same response, and slow when they lead to conflicting responses.

Why, then, would this competition not also reduce the dependence of bilinguals' discrimination accuracy on the longer vowel's quality? The answer is that competition does not so much weaken the perceptual illusion as activate an alternative representation in parallel, and given the ambiguity of the AX task (where participants' thresholds for detecting difference may differ), and the ambiguity of the shorter-vowel items (where too little vocalic material was present to unambiguously indicate the presence of a vowel), either representation could be construed as correct. Nonetheless, in a task explicitly identified as being in Spanish, with phonetically Spanish stimuli, the linguistically appropriate alternative is #esC, and competition from #sC was apparently not enough to prevent the bilinguals from choosing #esC. This suggests an interesting prediction for future research. Namely, if participants were less certain that the stimuli were supposed to be Spanish, e.g., if they were recorded with an English accent, or with a mixture of Spanish- and English-like phonetic features (cf. Gonzales & Lotto, 2013), bilinguals' responses would depend on what language they thought they were listening to.

Experiment 2

If bilinguals represent the same acoustic material in two contrasting ways simultaneously, how might this affect performance in a task where the illusion is useful, e.g., when it can restore a missing or distorted sound (Spinelli & Gros-Balthazard, 2007)? Experiment 2 addresses this question using a lexical decision task (LDT) in which Spanish #VsC words are sometimes presented with their initial vowel removed. Monolinguals' behavior when the missing vowel matches vs. mismatches the illusory [e] will be taken to indicate the strength of prelexical repair of reduced speech tokens (see above; Dupoux et al., 2001; Hallé et al., 2013), providing a baseline against which to assess the effects of knowing and recently using English.

Methods

Participants

Fourteen functional Spanish monolinguals (7 female) and 40 Spanish–English bilinguals were recruited in Granada, Spain as in Experiment 1 (mean age 23.4 years, sd = 3.04).[9] Bilinguals were assigned randomly to the English-switch or no-switch groups, as in Experiment 1. Sample size was determined on the basis of prior studies using similar methods (e.g., Dupoux et al., 2001).

As in Experiment 1, the bilinguals completed the background questionnaire, DELE, MELICET, and English and Spanish verbal fluency tasks. Their mean scores, along with self-reported daily English usage are shown in Table 4. No differences were found between the two bilingual groups on any of these measures (all $p > .13$).

All participants had been exposed to English in school, but the functional monolinguals reported significantly lower skill in speaking, reading, and comprehending spoken English (all $p < .0002$), and none reported regular use of English. Participants reported no substantial knowledge of other languages except for 3 of the bilinguals, who had knowledge of Catalan or Gallego. They were retained, as above, but excluding their data did not change the pattern of results.

Materials

The critical words for the LDT were 48 Spanish words beginning with #VsC (listed in Appendix B). Half began with the Spanish default prothetic vowel /e/ (E-words) and half with /a, i, o/ (notE-words). Words beginning with #usC are rare in Spanish and were omitted. These words were randomly selected from among the 711 inflectional families with initial #VsC, occurring in SUBTLEX-ES (Cuetos Vega, Glez Nosti, Barbón Gutiérrez & Brysbaert, 2011), with the following criteria. Words with initial stress (uncommon in #VsC words) multiple derivational affixes, and English loanwords (e.g., *escáner*, "scanner") were excluded, but cognates and non-cognates were allowed and balanced as much as possible. Words with initial orthographic *h*, which is silent in Spanish, were allowed,

[8] We might also ask why this effect was absent when participants responded *same*, but this comparison is hard to make, because neither phonotactic system would favor this response for longer-[a] pairs.

[9] An additional bilingual was tested but excluded because she had substantial early education in English.

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

Table 4. *Mean (sd) proficiency and daily English usage for bilinguals in Experiment 2.*

	No-switch (n = 19)	English-switch (n = 21)
Gender	14 female	17 female
DELE	.89 (.05)	.89 (.07)
Spanish verbal fluency	52.16 (9.42)	53.90 (7.94)
MELICET	.74 (.13)	.73 (.16)
English verbal fluency	36.78 (9.83)	38.38 (6.18)
Daily English use (hours)	3.22 (3.04)	3.82 (3.78)
Age of first exposure to English (years)	7 (3)	6 (3)
Age began using English outside classroom	15(6)	12(6)

since the task was purely auditory. The few cases where two #VsC words were highly similar apart from the initial vowel (e.g., *escudo*, *oscuro*, "shield", "dark" differ only in the manner of articulation of one consonant) were also excluded. Upper and lower frequency cutoffs were fixed to ensure that the E and notE-words were all reasonably frequent and comparably distributed across the same overall frequency range. The consonants following /s/ were /p, t, k, f/, in similar proportion across E and notE-words, plus one word with /l/.

After testing, one item (*asbesto*, "asbestos") was excluded because participants rarely accepted its intact form. Of the 23 remaining notE-words, 12 began with /a/, 5 with /i/, and 6 with /o/.[10] This represents a compromise between the goal of balancing the number of E vs. notE-words, and reflecting the proportions in the SUBTLEX-ES corpus (Cuetos Vega et al., 2011), where 81% of #VsC words (collapsing across inflections) begin with /e/, 10% with /a/, and 4% each with /i, o/.

Table 5 summarizes the lexical properties of the critical words. Phonotactic probability was the average positional log biphone frequency (from Marian, Bartolotti, Chabal & Shook, 2012), excluding the first biphone. This was done because initial /es/ is considerably more frequent than other initial #Vs sequences. Rather than being a confound, however, this reflects, and may help explain, the status of [e] as the default repair vowel in Spanish. It was therefore deemed preferable to ensure that the E and notE-words were otherwise comparable in phonotactic probability, as opposed to offsetting the high probability of initial /es/ by allowing much lower probability biphones later in the E-words. The stimuli were all nouns save 4 adjectives with initial /e/, 2 with /a/, and one each with /i, o/. All items were recorded by a male native Spanish speaker from northern Spain, who produced coda [s] reliably. Regional differences between Northern and Andalusian Spanish are not expected to impact the results because university students in Granada are familiar with other varieties used in Spain (particularly socially prestigious ones), and often produce coda [s] in formal contexts.

To test the effects of perceptual repair, the critical words were presented either intact, with the initial vowel removed (noVowel), or with a different initial vowel (wrongVowel). The noVowel versions were created by removing all material preceding the offset of periodic noise and applying a 40 ms linear ramp. This contrasts with Experiment 1, where a fragment of the vowel remained, but the goal here was not to determine if listeners could identify tokens produced without an initial vowel, but whether they would accept them as valid tokens of real words. No attempt was made to avoid coarticulatory traces of the initial vowel quality remaining on the [s] in the noVowel stimuli, though others have done so (e.g., Dupoux et al., 2001). The linear ramp may attenuate any such traces, but it may also give the impression that a very quiet vowel may have preceded the [s]. Neither of these features, however, would lead to the predicted asymmetry of perceptual repair, whereby /e/ would be restored much more reliably than other vowels.

To create the wrongVowel versions, the entire #Vs sequence, up to the offset of aperiodic noise, was exchanged with one taken from a word with the opposite initial vowel and the same post-/s/ consonant, to maintain coarticulatory features on the /s/ associated with its neighboring segments. Thus, for example, *astuto* "astute" became [estuto], and *espejo* "mirror" became [ospexo].[11] To avoid unnatural differences in intensity between the

[10] The small number of items precluded vowel-specific comparisons.

[11] A reviewer inquired whether the wrongVowel versions might have been perceived as beginning with a reduced vowel, which might influence responses to words with English cognates having an initial unstressed vowel (e.g., *astuto* 'astute'). This seems unlikely, as the vowels were spliced from real Spanish words recorded by the same talker, and in any event, all participant groups treated these stimuli consistently as nonwords (see Results).

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

Table 5. *Descriptive statistics for critical lexical decision stimuli.*

	N	Mean (sd)	Min	Max
E				
log SUBTLEX frequency	24	2.61 (0.68)	1.53	4.01
n phonemes	24	8.00 (1.44)	6	10
n syllables	24	3.54 (0.66)	3	5
mean log biphone prob.	24	−5.03 (0.44)	−5.87	−4.28
Cognates	12			
notE				
log SUBTLEX frequency	23	2.31 (0.81)	1.04	4.17
n phonemes	23	8.00 (1.35)	6	11
n syllables	23	3.61 (0.58)	3	5
mean log biphone prob.	23	−5.20 (0.69)	−7.53	−4.36
cognates	15			

spliced portions of the wrongVowel stimuli, all items were normalized to a mean intensity of 71dB (the mean of the original recordings) before splicing, and the stimuli were re-normalized after splicing. All acoustic manipulations were carried out using Praat (Boersma & Weenink, 2018).

The three versions of each critical word (intact, noVowel, wrongVowel) were counterbalanced across participants, and presented along with 70 real and 70 pseudoword fillers (see Appendix B) in random order (always beginning with 6 fillers), for a total of 188 trials. Since the intact stimuli were all real words, and the wrongVowel stimuli were all pseudowords, there were equal proportions of unambiguously real and pseudowords. The noVowel stimuli are considered ambiguous. While technically pseudowords, the illusory [e] would accurately restore the original E-words, and render the notE-words as phonotactically legal pseudowords.

The 70 real fillers were similar to the critical items in frequency, length, phonotactic probability, syntactic category, and stress. The pseudowords were created by entering the real word fillers into Wuggy (Keuleers & Brysbaert, 2010), which creates pseudowords matched to the provided inputs. Pseudoword fillers were phonotactically legal in Spanish, so that phonotactics did not provide a strong cue to lexical status.

Participants were asked to decide as quickly and accurately as possible whether each stimulus was or was not a real word in Spanish, and responded on a button box. Each trial began with a 500 ms fixation, after which the stimulus was played on studio quality headphones at a comfortable level. After participants responded, or after 3000 ms, a 500 ms blank screen preceded the next trial. RTs were measured from stimulus onset.

Procedure

The same procedure was followed as in Experiment 1, with the LDT replacing the AX task.

Predictions

Monolinguals were expected to accept noVowel words more readily when the missing vowel matched the perceptual repair, [e], than when it did not, but the predictions for bilinguals are more complex. If perceptual repair is simply less robust for bilinguals, they may reject tokens such as [skwela] (for *escuela*) more often than monolinguals. The missing vowel might still be restored via comparison with lexical representations (e.g., Ganong, 1980; Samuel, 1981), but this should lead to slower and less reliable acceptance than perceptual restoration of the vowel, which occurs prelexically (Dupoux et al., 2001).

However, since Spanish–English bilinguals are aware that #sC can occur in English (Altenberg, 2005) and know that it is a frequent source of difficulty, they may treat #esC and #sC as phonotactic translation equivalents (cf. Hanulíková et al., 2011). Moreover, when Spanish #esC words have English cognates, these tend to begin in #sC, e.g., *especial*, *special*, but other Spanish #VsC words have English cognates beginning with a similar vowel, e.g., *aspirina*, *aspirin* (Amengual, 2012, provides evidence that pronunciation can be affected by the presence of cognates). This pattern is strong enough that it could not be tightly controlled here (see Appendix B). Activation of English cognates such as *special* upon hearing the noVowel version of *especial* may therefore lead bilinguals to accept these items as much or even more than monolinguals.

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

Results

Prior to analysis, 29 trials with no response and 7 trials with RT $<$ 400 ms were removed ($<$ 1% of the data), leaving 2522 valid critical trials. This lower cutoff was higher than in Experiment 1 to ensure that most of the item had been heard prior to the response (in contrast, the AX decisions could be made as soon as the [s] began). No upper cutoff was applied, as trials timed out at 3000 ms, and few RTs approached this duration. RTs were negative reciprocal transformed ($\times$1000), which approximated a normal distribution well based on inspection of q-q plots.

Word vs. nonword responses

The analysis of *word* vs. *nonword* responses focused on the noVowel trials (843 observations), because the central predictions concern the noVowel condition, and because performance was (unsurprisingly) at ceiling for the intact and floor for the wrongVowel stimuli, with mean accuracy in excess of 92%, with no apparent dependence on group or the initial vowel.

The likelihood of responding *word* was analyzed using mixed effects logistic regression. The fixed effects of interest were initialV (E vs. notE) and languageGroup (as above). To explore possible transient language switch effects, trialNumber and its interactions with initialV and languageGroup were included. To control for other lexical properties of the items, length (of the stimulus as presented, in phonemes), phonotactic probability, and log frequency (all z-scored), plus cognate status were explored, but only length and cognate status were retained based on likelihood ratio tests ($p < .1$, all other $p > .4$). Cognate status was allowed to interact with languageGroup, because it is expected to affect bilinguals more than monolinguals. Interactions involving cognate status and initialV were explored, but none emerged as significant (all $p > .25$), and they were discarded.

Random intercepts by participant and item were included, with by-participant slopes for initialV and trialNumber and by-item slopes for languageGroup and trialNumber. Random interaction terms yielded very large variance estimates, suggesting quasi-complete separation, and they were omitted. The estimates in the final model appeared to be reasonable, however, so we report the results here (complete fixed effects estimates and further details appear in Table C1, Appendix C), and pursue this issue below.

The proportions of *word* responses to the critical stimuli are shown in Figure 3 (including intact and wrongVowel items, for reference). Figure 3 shows that, in the noVowel condition, participants nonetheless readily accepted words missing their initial [e], e.g., [spexo] for *espejo*, but performed around chance when the missing vowel was of different quality, e.g., [spiɾina] for *aspirina*. This large difference is confirmed by the main effect

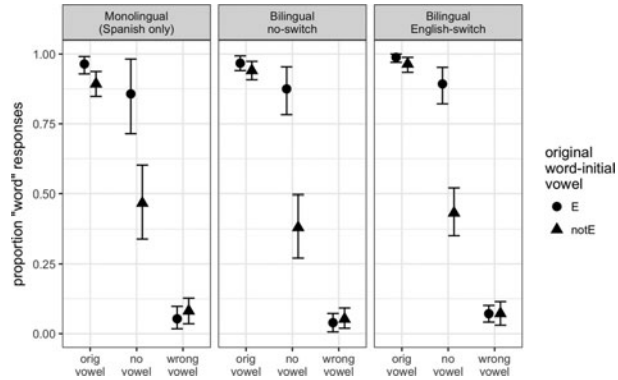

Figure 3. Proportion of *word* responses #VsC words presented intact, with the initial vowel removed, or with the initial vowel replaced with a different vowel, with 95% CIs computed by nonparametric bootstrap.

of initialV ($\beta = 4.50$, $SE = 0.74$, $\chi2(1) = 43.85, p < .0001$). The interactions of initialV with bilingualCode and priorLangCode were not significant (both $p > .7$), but a three-way interaction of initialV, priorLangCode and trialNumber emerged ($\beta = -4.42$, $SE = 2.27$, $\chi2(1) = 4.00$, $p = .046$), suggesting a possible transient switch effect.

To unpack this interaction, a model was first fit to the notE-words only, which yielded no significant effects or interactions of trialNumber (all $p > .13$). However, a similar model fit to the E-words yielded unreasonably large estimates. This likely reflects quasi-complete separation, since well over half the participants completely ignored the absence of a word-initial [e], responding *word* every time to these items (proportions across groups were similar, 9/14 monolinguals, 10/19 no-switch bilinguals, and 12/21 English-switch bilinguals; note that the *nonword* responses that did occur were distributed evenly over 22 of the 24 E-words).

The possible trialNumber effects on E-word responses are thus explored graphically in Figure 4. Of the participants who ever answered *nonword*, monolinguals and no-switch bilinguals tended to do so early in the task, continuing to do so throughout, whereas English-switch bilinguals rarely did so until the last third of the task. While this should be taken cautiously, it nonetheless fits the expectation that any transient effects of prior language context would occur after switching back from English, and suggests an elevated tendency to accept stimuli with a missing [e] immediately after using English, which faded later in the task.

Response times

To explore the consequences of removing the initial vowel on response times, the analysis focused on the intact and noVowel conditions. WrongVowel trials, which all three participant groups consistently rejected, were excluded

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

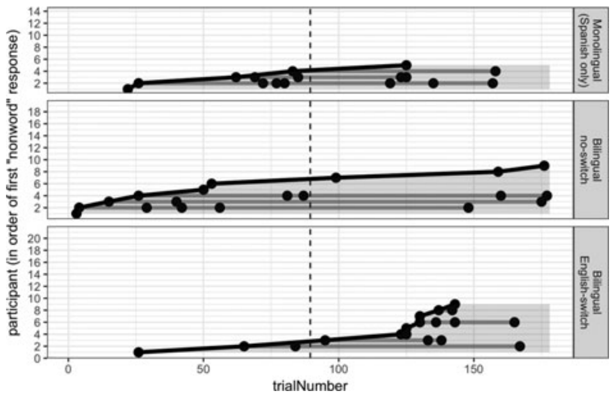

Figure 4. *Nonword* responses to #VsC words presented with no initial vowel, as a function of trialNumber, separated by participant and group. Each point is a *nonword* response; responses by the same participant are linked with a dark gray line. Height of shaded area shows the number of participants in each group to have responded *nonword* at least once, as a function of trialNumber. Height of each panel shows the total number of participants in each group. The dotted vertical line shows the midpoint of the task.

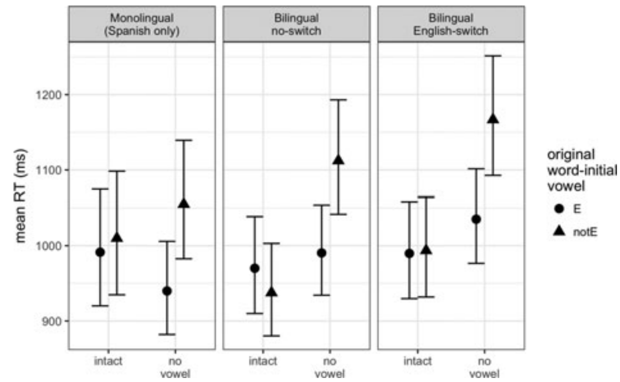

Figure 5. Estimated RTs and 95% CIs (back-transformed to milliseconds) for intact and noVowel stimuli in Experiment 2, separated by the original vowel and languageGroup.

(though rejections were faster for E-words, e.g., [ospexo], from *espejo* "mirror", than notE-words).[12] All responses were included (n = 1685), due to the ambiguous lexical status of the noVowel stimuli, but the highly uneven distribution of *nonword* responses across conditions (cf. Figure 3) prevented including Response as a predictor. However, refitting the model using only *word* responses yielded the same results.

The model structure was the same as for *word* vs. *nonword* responses, except for the following changes. To capture differences between the intact and noVowel conditions, a two-level factor (Vpresent) was added and allowed to interact with initialV and languageGroup. The three-way interaction of initialVowel, priorLangCode, and trialNumber approached significance ($\beta = -0.129$, SE = 0.07, $\chi 2(1) = 3.17$, p = .075) and was retained (with its component two-way interactions), but interactions of trialNumber and Vpresent were not (all $p > .23$). Log word frequency was included as a covariate, and stimulus duration (z-scored) was substituted for length in phonemes, as a more direct way to control for the effects of stimulus length on RT (the same results are obtained either way). Complete fixed effects estimates and further details appear in Table C2 in Appendix C.

Figure 5 shows the estimated RTs to intact and noVowel stimuli. Crucially, removing the initial vowel affected monolinguals and bilinguals differently overall, reflected in the two-way interaction of Vpresent with bilingualCode ($\beta = 0.09$, SE = 0.02, $\chi 2(1) = 14.26$, p = .0002). Bilinguals were slower than monolinguals when the initial vowel was missing, as confirmed via post-hoc tests ($\beta = 0.08$, SE = 0.04, $\chi 2(1) = 5.32$, p = .021; Bonferroni-corrected $\alpha = .025$) but bilinguals' and monolinguals' RTs to intact words did not differ ($\chi 2 < 1$)

When words were presented intact, no dependence of RTs on the quality of the initial vowels was found, but in the noVowel condition, all groups responded faster to stimuli with a missing [e] than to those missing a different vowel. This is confirmed by a significant interaction between initialV and Vpresent ($\beta = -0.12$, SE = 0.02, $\chi 2(1) = 19.50$, $p < .0001$), and by a main effect of initialV in a model fit to only the noVowel results ($\beta = -0.14$, SE = 0.02, $\chi 2(1) = 32.75$, $p < .0001$). Additional post-hoc tests (Bonferroni-corrected $\alpha = .0125$) were performed to confirm that removing the initial vowel led to slower responses in bilinguals for both E-words ($\beta = 0.05$, SE = 0.02, $\chi 2(1) = 8.23$, p = .004) and notE-words ($\beta = 0.14$, SE = 0.02, $\chi 2(1) = 37.18$, $p < .0001$). Removing the initial vowel did not change RTs for monolinguals, for either E or notE-words (both $\chi 2 < 1$).

Discussion

As in Experiment 1, bilinguals and monolinguals differed only slightly in their responses, but salient group differences were observed in RTs. Taking the similarities first, participants in all three groups largely ignored the absence of initial [e], accepting tokens like [skwela] (from *escuela*) nearly as frequently as [eskwela], but responding to [spirina] (from *aspirina*) at chance, and more slowly than to [skwela]. This, together with the large longerV effect on discrimination observed across groups in Experiment 1, shows that repair via [e]-prothesis is alive and well for Spanish–English bilinguals.

[12] Dupoux et al. (2001) employed an alternative strategy, using all "expected" responses, i.e. all *word* responses to intact words and noVowel E-words, and *nonword* responses to wrongVowel words and noVowel notE-words. Applying this strategy here yielded the same results.

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

What, then, of the group differences? First, there appeared to be a transient enhancement of perceptual repair after switching to Spanish from English, whereby bilinguals were initially more likely to accept words missing their initial [e], but not other initial vowels. Though this occurred only for some participants, it bears out the prediction that transient effects would appear only in the English-switch group (Elston-Güttler & Gunter, 2008; Elston-Güttler et al., 2005; Goldrick et al., 2014; Olson, 2013). Second, removing the initial vowel slowed responses for bilinguals only, regardless of the original vowels' quality. This strongly suggests that they detected the absence of the vowel even though they, like the monolinguals, accepted items missing an initial [e] consistently, and more than items missing a different initial vowel.

While bilinguals may have noticed the initial #sC sequences, they nonetheless accepted the noVowel items at a similar rate to monolinguals, suggesting that bilinguals found them to be both recognizable and acceptable as valid tokens of their intact counterparts. This may be due to a greater role for lexical repair, consistent with bilinguals' overall slower responses to noVowel items, although the dependence of bilinguals' responses and RTs on the quality of the missing vowel shows that bilinguals' performance is also shaped by (prelexical) perceptual repair.

General discussion

In both experiments, then, bilinguals' responses revealed robust, monolingual-like repair of #sC, but they responded more slowly in those conditions where Spanish and English phonotactics were in conflict. The acoustically accurate parse, #sC, thus appears to play a greater role in bilinguals' processing than it does in monolinguals, although they ultimately reach the same decisions. There were, however, some differences across experiments. In discrimination, only bilinguals' RTs were sensitive to the longer vowel's quality, with a stronger effect found after switching from English. In lexical decision, on the other hand, all groups were equally sensitive to the missing vowel's quality, and RTs were not affected by switching languages.

These differences could be methodological (e.g., missing vs. very short vowels), but they may also reflect how English interacts with varying task conditions. In lexical decision, the English-like representation would indicate a (phonotactically illicit) nonword regardless of the original vowel's quality, leading to a uniform processing cost for bilinguals and leaving the effect of the missing vowel's quality stable across groups. In discrimination, however, the impact of the English-like representation (#sC) would lead to competition between the acoustically and linguistically correct responses only for the longer-[e] pairs (see above), and thus to asymmetrical RT effects.

The difference in switching effects across tasks may also reflect the greater relevance of acoustic information in the discrimination task. Using English before this discrimination task may have made this information more accessible or salient. In lexical decision, on the other hand, the absence of other illicit items in the stimulus list may have downplayed the importance of acoustic detail. This may have suppressed switch effects, but it would not necessarily prevent them. This leads to the prediction that repeating the experiment with the addition of other illicit pseudowords should create an expectation that some of the nonwords will be phonotactically impossible (cf. Altenberg & Cairns, 1983), forcing bilinguals to decide whether to count an item like [skwela] as one of the illicit ones or not. This would enhance the competition between *word* and *nonword* responses, allowing switch effects to emerge. Conversely, competition could be weakened in discrimination by reducing response bias, e.g., in the ABX paradigm. Monolinguals should perform poorly on this task, but if bilinguals can at all detect the difference between #esC and #sC when required, then they should perform well, regardless of the preceding language context.

The present findings thus support the conclusion that hearing an acoustic #sC sequence appears to activate a wider range of candidate representations for bilinguals than for monolinguals. However, the important thing is not whether bilinguals are sensitive to the presence/absence of an initial vowel (the RT effects provide clear evidence that they are), but rather what listeners do with the acoustic variability they can detect, and how different linguistic and task conditions influence the activation and sifting of candidates during processing (Apfelbaum et al., 2014; Apfelbaum & McMurray, 2014; Kleinschmidt & Jaeger, 2015; Norris et al., 2016).

How might this work in the present case? Monolinguals and Spanish–English bilinguals alike know that Spanish words do not begin with #sC, and they also know that the most probable valid parse in Spanish is #esC, with the remaining probability mass divided primarily among the other vowel-initial possibilities. After all, Spanish was the bilinguals' first and clearly dominant language, and they were currently residing in a more or less monolingual Spanish environment. The difference is in the treatment of #sC. For monolinguals, the veridical parse is not even representable in their phonotactic system, and it is thus unlikely to impact processing much, even if they notice the absence of the vowel. It is, however, representable for bilinguals, albeit in the non-target language, allowing it to impact processing in a way that it can't for monolinguals. One interesting consequence of this reasoning is that knowledge of the impossibility of #sC in Spanish may be different for bilinguals and monolinguals. In a sense,

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

monolinguals may not actually know this constraint at all, they simply have no way of representing the unattested sequences. Bilinguals, however, can represent them (even at early stages of learning, Altenberg, 2005), and can therefore identify them as structures that happen to be impossible in Spanish.

What happens next depends on the interplay of available representations and task demands. With no reason to expect that #sC sequences would be relevant to, or even occur in, lexical decision, bilinguals behaved like monolinguals; only a bit slower because they detected the vowel's absence. After encountering a few words with missing initial vowels, bilinguals might have become suspicious and revised their decision strategy, but while there was a hint of this in the English-switch group, more overt cues (e.g., other illicit nonwords) may be required to produce reliable changes in strategy. In discrimination, however, the robustness (duration) of the initial vowel was in focus from the beginning, and clearly relevant for discrimination. While the vowel was never completely absent, its brevity allowed room for doubt, supported by the representability of #sC. Bilinguals were slower where the ultimate response depended on how that doubt was resolved, and using English prior to the AX appears to have made the alternative resolution more salient, even though they reached similar decisions in the end.

Conclusion

The emerging story here is that experience with L2 English allows Spanish–English bilinguals to perceive #sC sequences veridically. The consequence, however, is not to prevent perception of the illusory vowel but rather to add an alternative to it, broadening the range of alternatives that participate in processing. This changes the interplay of acoustic information and phonotactic knowledge during speech perception and subsequent decision making, and this interplay evolves dynamically depending on the bilinguals' current linguistic context.

This conclusion finds a natural explanation in the recent application of Bayesian reasoning to speech perception (Apfelbaum et al., 2014; Apfelbaum & McMurray, 2014; Feldman et al., 2009; Kleinschmidt & Jaeger, 2015; Norris et al., 2016), where speech perception is thought of as an active cognitive process (Heald & Nusbaum, 2014), in which different sources of information can be re-weighted, and expectations based on prior knowledge adjusted to match current conditions, leading to an interesting possibility. While the present study did not directly test the role of listener expectations in the observed effects, a Bayesian approach suggests that this would be a fruitful direction in which to continue this research. Indeed there is already evidence that listeners adapt their perception of phonetic categories based on knowledge of a talker's bilingualism (Molnar, Ibáñez-Molina & Carreiras, 2015; Samuel & Larraza, 2015), or dialect background (Dahan, Drucker & Scarborough, 2008; Hay, Drager & Warren, 2010; Hay, Nolan & Drager, 2006; Kraljic, Brennan & Samuel, 2008).

This points ultimately to the proposal that bilinguals not only know how to use the structures involved in each of their languages, but they also know how the two systems are related to each other, and this counterpoint allows for ways of using their languages that are not available to monolinguals (Cook, 1992; Cook & Wei, 2016; Hall et al., 2006; Otheguy et al., 2015). In a sense, it may be better to think of bilinguals not as balancing two systems, at the mercy of extra representations that compete for activation, but rather as having a measure of control over the porosity of their language systems, allowing them to process the same acoustic material adaptively based on current conditions. Looking forward, this may add crucial nuance to the holistic view of bilingualism advanced by researchers like Grosjean (1989), and to the idea that both languages are not only always active, but gradiently integrated (Goldrick et al., 2016) when bilinguals use language.

Appendix A

Model details, Experiment 1, AX Discrimination

The fixed effects factors were coded as follows. LongerV, a binary factor encoding the quality of the longer vowel in each pair, was sum-coded as −0.5 for [a]

Table A1. *Fixed effects for response accuracy in Experiment 1. Positive values indicate higher likelihood of responding same. Significance assessed via likelihood ratio tests of individual coefficients.*

	Dependent variable: Response = word
	β (SE(β))
Intercept	0.007 (0.103)
longerV	2.386*** (0.147)
bilingualCode	−0.740*** (0.221)
priorLangCode	−0.085 (0.238)
matchingV	−2.118*** (0.132)
longerV:bilingualCode	−0.487 (0.311)
longerV:priorLangCode	−0.184 (0.325)
longerV:matchingV	−0.284 (0.248)
bilingualCode:matchingV	0.351 (0.278)
priorLangCode:matchingV	−0.250 (0.286)
longerV:bilingualCode:matchingV	0.529 (0.517)
longerV:priorLangCode:matchingV	0.181 (0.528)
Observations	3,282

Note: . $p < .1$; *$p < .05$; **$p < .01$; ***$p < .0001$

Table A2. *Fixed effects for RT in Experiment 1. Units are in negative reciprocal milliseconds (x1000). Significance was assessed via likelihood ratio tests of individual coefficients.*

	Dependent variable: −1000/RT(ms)	
	Correct responses	Incorrect responses
Intercept	−0.940*** (0.022)	−0.875*** (0.023)
longerV	0.110*** (0.021)	−0.030. (0.018)
bilingualCode	0.137** (0.048)	0.157** (0.049)
priorLangCode	−0.012 (0.051)	−0.037 (0.055)
matchingV	−0.039* (0.018)	0.063*** (0.017)
trialNumber	−0.037*** (0.006)	−0.029*** (0.006)
longerV:bilingualCode	−0.009 (0.049)	0.029 (0.037)
longerV:priorLangCode	0.084* (0.043)	0.034 (0.045)
longerV:matchingV	−0.017 (0.036)	−0.001 (0.035)
bilingualCode:matchingV	0.025 (0.044)	0.013 (0.036)
priorLangCode:matchingV	0.003 (0.036)	−0.013 (0.044)
longerV:bilingualCode:matchingV	0.237* (0.087)	−0.031 (0.071)
longerV:priorLangCode:matchingV	−0.081 (0.072)	−0.008 (0.087)
Observations	1,682	1,600

Note: .p < .1; *p < .05; **p < .01; ***p < .001

and 0.5 for [e]. LanguageGroup, a 3-level factor, was coded using orthogonal contrasts. The first contrast (bilingualCode) compared monolinguals with bilinguals, and the second (priorLangCode) compared the English-switch and no-switch groups. MatchingV was coded as –0.5 (the quality of the initial vowels in each pair was the same) and 0.5 (different). TrialNumber (RT model only) was a continuous variable, scaled to a range of 1 and centered.

The maximal random effects structure by-participants was used, i.e., random intercepts and slopes for longerV, matchingV, and their interaction. By-item effects were omitted because the items did not constitute a random sample, and the most important differences between items are captured in the fixed effects.

Both models were bootstrap validated, confirming the results reported here, and the residuals were checked, confirming normality, homoskedasticity, and independence.

Appendix B

Stimuli used in Experiment 2, Lexical Decision

Table B1. *Critical stimuli used in the lexical decision task.*

Word	Gloss	Initial vowel	Wrong vowel	Cognate
asbesto	asbestos	a	e	1
asfalto	asphalt	a	e	1
asfixia	asphyxia	a	e	1
aspavientos	fuss	a	e	0
aspecto	aspect	a	e	1
aspereza	roughness	a	e	0
aspiradora	vacuum cleaner	a	e	0
aspirante	aspirant	a	e	1
aspirina	aspirin	a	e	1
asqueroso	disgusting	a	e	0

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

Table B1. *Continued.*

Word	Gloss	Initial vowel	Wrong vowel	Cognate
asterisco	asterisk	a	e	1
astronauta	astronaut	a	e	1
astuto	astute	a	e	1
escoba	broom	e	a	0
escorpión	scorpion	e	a	1
escritor	writer	e	o	0
escuela	school	e	a	1
escultor	sculptor	e	o	1
esfera	sphere	e	a	1
esfuerzo	effort	e	a	0
espalda	back	e	i	0
espantoso	frightening	e	o	0
espátula	spatula	e	i	1
específico	specific	e	a	1
espejo	mirror	e	o	0
esperanza	hope	e	a	0
espinaca	spinach	e	a	1
esponja	sponge	e	a	1
espuela	spur	e	a	0
espuma	foam	e	a	0
estimado	esteemed	e	i	0
estornudo	sneeze	e	a	0
estrategia	strategy	e	a	1
estropeado	broken	e	o	0
estructura	structure	e	i	1
estudiante	student	e	a	1
estufa	stove	e	o	1
hispano	hispanic	i	e	1
histeria	hysteria	i	e	1
histérico	hysterical	i	e	1
historia	history	i	e	1
isleño	islander	i	e	0
hospedaje	lodging	o	e	0
hospicio	hostel/hospice	o	e	1
hospital	hospital	o	e	1
hostilidad	hostility	o	e	1
oscuridad	darkness	o	e	0
ostentoso	ostentatious	o	e	0

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

Table B2. *Filler stimuli used in the lexical decision task.*

Real word (gloss)		Pseudowords	
abogado (lawyer)	luchador (fighter)	acasto	gruceno
accidente (accident)	lucidez (lucidity)	aciesto	insarsivo
alergia (allergy)	lustroso (lustrous)	ademente	jupentiz
aluminio (aluminum)	moderno (modern)	alarcio	macalera
asunto (issue)	monitor (monitor)	amiñada	mergación
bajista (bassist)	natación (swimming)	amulicia	mermelaca
bandera (flag)	necesario (necessary)	baltista	mestidor
cacharro (junk)	orquesta (orchestra)	benidario	mituento
cajero (cashier)	palidez (pallor)	bentilia	mochega
camioneta (pickup truck)	panadero (baker)	bicrinia	modisco
caracol (snail)	perjuicio (prejudice)	cabricho	nadura
carretera (highway)	pescador (fisherman)	cachilero	niduna
cocinero (cook)	pimiento (pepper)	cacuileto	nupidor
colector (collector)	planeta (planet)	canvina	onviesta
columna (column)	pobreza (poverty)	carasán	paraduz
conciencia (conscience)	porcelana (porcelain)	cavena	pasguicio
cortina (curtain)	prodigio (marvel)	ceoste	pematija
dentista (dentist)	programa (program)	chadena	pomator
decisión (decision)	relativo (relative)	coraltor	poparno
demanda (demand)	remedio (remedy)	coresna	posator
dentista (dentist)	rosario (rosary)	cortuencia	prafioso
deporte (sport)	salvación (salvation)	cotelera	propagia
desierto (desert)	sistema (system)	debalte	propliga
difusión (diffusion)	soldado (soldier)	demista	ractera
discoteca (disco)	soltera (single)	demuelto	retemión
emigrante (emigrant)	tontería (nonsense)	devedia	romiria
famoso (famous)	travieso (naughty)	dimporena	sulbido
figura (figure)	tremendo (tremendous)	dosteza	tariante
frutero (fruit bowl)	unicornio (unicorn)	ecefrinte	taroco
garaje (garage)	utensilio (utensil)	farime	uneculsio
granjero (farmer)	vacuna (vaccine)	fidrana	uternivia
intensiva (intensive)	valiente (valient)	flonero	vallesta
jugador (player)	virtuoso (virtuous)	forente	valmera
juventud (youth)	vitrina (display cabinet)	fubridor	virtosno
librería (bookstore)	volante (steering wheel)	fumbroso	vuradez

Appendix C

Model details, Experiment 2, Lexical Decision

The quality of the original initial vowel, initialV, was sum-coded as −0.5 for notE and 0.5 for E. The factor languageGroup was coded as in Experiment 1 using orthogonal contrasts for bilingualCode and priorLangCode. TrialNumber was scaled to a range of 1 and centered, and Cognate was sum-coded as −0.5 for noncognates and 0.5 for cognates. Length (in phonemes) was z-scored. In the analysis of response times, Vpresent encoded the difference between intact (–0.5) and noVowel (0.5) items, Duration (in ms, z-scored) was substituted for Length, and log word frequency (z-scored) was added as a covariate.

In the analysis of *word* vs. *nonword* responses, random intercepts by participant and item were included, with by-participant slopes for initialV and trialNumber and by-item slopes for languageGroup and trialNumber. In the analysis of RTs, the by-item random effects referred

https://doi.org/10.1017/S136672891800086X Published online by Cambridge University Press

Table C1. *Fixed effects for word vs. nonword responses in Experiment 2 (noVowel trials only). Positive coefficients indicate increasing likelihood of a word response. Significance was assessed via likelihood ratio tests of individual coefficients.*

	Dependent variable: Response = word
	β (SE(β))
Intercept	2.092*** (0.469)
initialV	4.501*** (0.740)
bilingualCode	−0.314 (0.674)
priorLangCode	0.497 (0.720)
trialNumber	−1.787*** (0.623)
Cognate	−0.300 (0.516)
length (phonemes)	0.557* (0.289)
initialV:bilingualCode	0.222 (0.924)
initialV:priorLangCode	0.313 (0.954)
initialV:trialNumber	−1.223 (1.283)
bilingualCode:trialNumber	0.843 (1.153)
priorLangCode:trialNumber	−1.722 (1.189)
bilingualCode:Cognate	0.013 (0.612)
priorLangCode:Cognate	−1.289** (0.589)
initialV:bilingualCode:trialNumber	−0.250 (2.182)
initialV:priorLangCode:trialNumber	−4.420* (2.271)
Observations	843

*Note: .p < .1; *p < .05; **p < .01; ***p < .001*

Table C2. *Fixed effects for RT in Experiment 2, word responses to intact and noVowel stimuli only. Significance was assessed via likelihood ratio tests of individual coefficients. Units are in negative reciprocal milliseconds (x1000).*

	Dependent variable: Response = *−1000/RT(ms)*
	β SE(β)
Intercept	−0.973*** (0.017)
initialV	−0.073*** (0.017)
bilingualCode	0.036 (0.032)
priorLangCode	0.042 (0.037)
Vpresent	0.062*** (0.015)
trialNumber	0.037 (0.024)
Cognate	−0.016 (0.015)
log word frequency	−0.029*** (0.007)
stimulus duration	0.029*** (0.008)
initialV:bilingualCode	0.014 (0.024)
initialV:priorLangCode	−0.014 (0.028)
initialV:Vpresent	−0.116*** (0.024)
initialV:trialNumber	0.020 (0.032)
bilingualCode:Vpresent	0.092*** (0.023)
priorLangCode:Vpresent	0.002 (0.025)
bilingualCode:trialNumber	−0.004 (0.050)
priorLangCode:trialNumber	−0.007 (0.055)
bilingualCode:Cognate	0.028 (0.020)
priorLangCode:Cognate	0.048** (0.023)
initialV:bilingualCode:Vpresent	−0.024 (0.045)
initialV:priorLangCode:Vpresent	0.041 (0.049)
initialV:bilingualCode:trialNumber	−0.017 (0.066)
initialV:priorLangCode:trialNumber	−0.129* (0.072)
Observations	1,376

*Note: . p < .1; *p < .05; **p < .01; ***p < .001*

to the original lexical item, regardless of the version presented, such that differences between the intact and noVowel versions are encoded via Vpresent in the fixed effects. By-participant and by-item slopes for Vpresent, initialV, languageGroup, and trialNumber, including their interactions, were included as justified by the design.

Both models were bootstrap validated, confirming the results reported here, and the residuals were checked, confirming normality, homoskedasticity, and independence.

References

Abrahamsson, N. (1999). Vowel Epenthesis of /sC(C)/ Onsets in Spanish/Swedish Interphonology: A Longitudinal Case Study. *Language Learning, 49,* 473–508. doi: 10.1111/0023-8333.00097

Altenberg, E. P. (2005). Judgment, perception, and production of consonant clusters in a second language. *International Review of Applied Linguistics in Language Teaching, 43,* 53–80. doi: 10.1191/0267658305sr250oa

Altenberg, E. P., & Cairns, H. S. (1983). The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior, 22,* 174–188. doi: 10.1016/S0022-5371(83)90134-2

Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition, 15,* 517–530. doi: 10.1017/S1366728911000460

Anisfeld, M., Anisfeld, E., & Semogas, R. (1969). Cross-influences between the phonological systems of Lithuanian-English bilinguals. *Journal of Verbal Learning and Verbal Behavior, 8,* 257–261. doi: 10.1016/S0022-5371(69)80072-1

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics, 38,* 640–653. doi: 10.1016/j.wocn.2010.09.005

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics, 39,* 558–570. doi: 10.1016/j.wocn.2011.03.001

Apfelbaum, K. S., Bullock-Rest, N., Rhone, A. E., Jongman, A., & McMurray, B. (2014). Contingent categorisation in speech perception. *Language, Cognition and Neuroscience, 29,* 1070–1082. doi: 10.1080/01690965.2013.824995

Apfelbaum, K. S., & McMurray, B. (2014). Relative cue encoding in the context of sophisticated models of categorization: Separating information from categorization. *Psychonomic Bulletin & Review, 22,* 916–943. doi: 10.3758/s13423-014-0783-2

Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., Riches, N., & Thierry, G. (2015). Two Languages, Two Minds Flexible Cognitive Processing Driven by Language of Operation. *Psychological Science, 26,* 518–526. doi: 10.1177/0956797614567509

Balukas, C., & Koops, C. (2015). Spanish–English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism, 19,* 423–443. doi: 10.1177/1367006913516035

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67,* 1–48. doi: 10.18637/jss.v067.i01

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America, 114,* 1600–1610. doi: 10.1121/1.1603234

Berent, I., & Lennertz, T. (2010). Universal constraints on the sound structure of language: Phonological or acoustic? *Journal of Experimental Psychology: Human Perception and Performance, 36,* 212–223. doi: 10.1037/a0017638

Berent, I., Lennertz, T., & Rosselli, M. (2012). Universal linguistic pressures and their solutions: Evidence from Spanish. *The Mental Lexicon, 7,* 275–305. doi: 10.1075/ml.7.3.02ber

Berent, I., Lennertz, T., Smolensky, P., & Vaknin-Nusbaum, V. (2009). Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology, 26,* 75. doi: 10.1017/S0952675709001729

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition, 104,* 591–630. doi: 10.1016/j.cognition.2006.05.015

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.40, retrieved 30 June 2018 from http://www.praat.org/

Caramazza, A., Yeni-Komshian, G., & Zurif, E. B. (1974). Bilingual switching: The phonological level. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *28,* 310–318.

Carlisle, R. S. (1991). The Influence of Environment on Vowel Epenthesis in Spanish/English Interphonology. *Applied Linguistics, 12,* 76–95. doi: 10.1093/applin/12.1.76

Carlisle, R. S. (1999). The Modification of Onsets in a Markedness Relationship: Testing the Interlanguage Structural Conformity Hypothesis. *Language Learning, 49,* 59–93. doi: 10.1046/j.0266-8254.2003.03701.x-i1

Carlson, M. T. (2018). Making room for L2 phonotactics: Effects of L2 learning and environment on L1 speech perception. *Language and Speech*, doi: 10.1177/0023830918767208. Published online by Sage, April 9, 2018

Carlson, M. T., Goldrick, M., Blasingame, M., & Fink, A. (2016). Navigating conflicting phonotactic constraints in bilingual speech perception. *Bilingualism: Language and Cognition, 19,* 939–954. doi: 10.1017/S1366728915000334

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics, 40,* 249–268. doi: 10.1016/j.wocn.2011.10.007

Chang, C. B. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics, 41,* 520–533. doi: 10.1016/j.wocn.2013.09.006

Chang, C. B. (2016). Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition, 19,* 791–809. doi: 10.1017/S1366728914000261

Chang, C. B., & Mishler, A. (2012). Evidence for language transfer leading to a perceptual advantage for non-native listeners. *The Journal of the Acoustical Society of America, 132,* 2700–2710. doi: 10.1121/1.4747615

Cohen, S. P., Tucker, G. R., & Lambert, W. E. (1967). The comparative skills of monolinguals and bilinguals in perceiving phoneme sequences. *Lanugage and Speech, 10,* 159–168.

Cook, V. J. (1992). Evidence for Multicompetence. *Language Learning, 42,* 557–591. In

Cook, V. J., & Wei, L. (eds.). (2016). *The Cambridge Handbook of Linguistic Multi-competence*. Cambridge: Cambridge University Press.

Cuetos, F., Hallé, P. A., Domínguez, A., & Segui, J. (2011). Perception of prothetic /e/ in #sC utterances: Gating data. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 540–543).

Cuetos Vega, F., Glez Nosti, M., Barbón Gutiérrez, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica: Revista de metodología y psicología experimental, 32,* 133–143.

Dahan, D., Drucker, S. J., & Scarborough, R. A. (2008). Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition, 108,* 710–718. doi: 10.1016/j.cognition.2008.06.003

Daland, R., & Norrmann-Vigil, I. (2015). Toward a generative theory of language transfer: Experiment and modeling of sC prothesis in L2 Spanish. *Open Linguistics, 1*. doi: 10.1515/opli-2015-0024

Davidson, L. (2011). Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. *Journal of Experimental Psychology: Human Perception & Performance, 37,* 270–282. doi: 10.1037/a0020988

Davidson, L., & Shaw, J. A. (2012). Sources of illusion in consonant cluster perception. *Journal of Phonetics, 40,* 234–248. doi: 10.1016/j.wocn.2011.11.005

Dijkstra, T., van Jaarsveld, H., & ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1,* 51–66. doi: 10.1017/S1366728998000121

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. doi: https://doi.org/10.1037/0096-1523.25.6.1568

Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New Evidence for Prelexical Phonological Processing in Word Recognition. *Language and Cognitive Processes, 16,* 491–505. doi: 10.1080/01690960143000191

Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A Destressing "Deafness" in French? *Journal of Memory and Language, 36,* 406–421. doi: 10.1006/jmla.1996.2500

Dupoux, E., Parlato, E., Frota, S., Hirose, Y., & Peperkamp, S. (2011). Where do illusory vowels come from? *Journal of Memory and Language, 64,* 199–210.

Durvasula, K., & Kahng, J. (2015). Illusory vowels in perceptual epenthesis: The role of phonological alternations. *Phonology, 32,* 385–416. doi: 10.1017/S0952675715000263

Elston-Güttler, K. E., & Gunter, T. C. (2008). Fine-tuned: Phonology and Semantics Affect First- to Second-language Zooming In. *Journal of Cognitive Neuroscience, 21,* 180–196. doi: 10.1162/jocn.2009.21015

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research, 25,* 57–70. doi: 10.1016/j.cogbrainres.2005.04.007

English Language Institute. (2001). *MELICET-GCVR user's manual*. Ann Arbor, MI: English Language Institute, The University of Michigan.

Ernestus, M., Kouwenhoven, H., & van Mulken, M. (2017). The direct and indirect effects of the phonotactic constraints in the listener's native language on the comprehension of reduced and unreduced word pronunciation variants in a foreign language. *Journal of Phonetics, 62,* 50–64. doi: 10.1016/j.wocn.2017.02.003

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review, 116,* 752–782. doi: 10.1037/a0017196

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. O. Schiller & A. S. Meyer (eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities*, pp. 319–355. Berlin: Mouton de Gruyter.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory phonology, 9,* 353–382.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15,* 67–83.

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics, 36,* 649–663. doi: 10.1016/j.wocn.2008.04.001

Freeman, M. R., Blumenfeld, H. K., & Marian, V. (2016). Phonotactic Constraints Are Activated across Languages in Bilinguals. *Frontiers in Psychology, 7*. doi: 10.3389/fpsyg.2016.00702

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 110–125.

Gaskell, M. G. (2001). Phonological variation and its consequences for the word recognition system. *Language and Cognitive Processes, 16,* 723–729. doi: 10.1080/01690960143000128

Goldrick, M., Putnam, M., & Schwarz, L. (2016). Coactivation in bilingual grammars: A computational account of code mixing. *Bilingualism: Language and Cognition, 19,* 857–876. doi: 10.1017/S1366728915000802

Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science, 25,* 1031–1036. doi: 10.1177/0956797613520014

Gonzales, K., & Lotto, A. J. (2013). A *bafri*, un *pafri*: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science, 24,* 2135–2142. doi: 10.1177/0956797613486485

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36,* 3–15. doi: 10.1016/0093-934X(89)90048-5

Grosjean, F., & Miller, J. L. (1994). Going in and out of languages: An example of bilingual flexibility. *Psychological Science, 5,* 201–206.

Hall, J. K., Cheng, A., & Carlson, M. T. (2006). Reconceptualizing multicompetence as a theory of language knowledge. *Applied Linguistics, 27,* 220–240. doi: 10.1093/applin/aml013

Hallé, P. A., Dominguez, A., Cuetos, F., & Segui, J. (2008). Phonological mediation in visual masked priming: Evidence from phonotactic repair. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 177–192. doi: 10.1037/0096-1523.34.1.177

Hallé, P. A., Segui, J., Dominguez, A., & Cuetos, F. (2013). *Special* es *especial* pero *stuto* no es *astuto*: La percepción de una /e/ protética en el habla y la escritura de hablantes de español. In V. Jaichenco & Y. Sevilla (eds.), *Psicolingüística en español: Homenaje a Juan Segui*, pp. 31–47. Buenos Aires: Editorial de la Facultad de Filosofía y Letras, Universidad de Buenos Aires.

Hanulíková, A., Mitterer, H., & McQueen, J. M. (2011). Effects of first and second language on segmentation of non-native speech. *Bilingualism: Language and Cognition, 14,* 506–521.

Hanulíková, A., & Weber, A. (2012). *Sink* positive: Linguistic experience with *th* substitutions influences nonnative word recognition. *Attention, Perception, & Psychophysics, 74,* 613–629. doi: 10.3758/s13414-011-0259-7

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is Syntax Separate or Shared Between Languages? *Psychological Science, 15,* 409–414. doi: 10.1111/j.0956-7976.2004.00693.x

Hay, J., Drager, K., & Warren, P. (2010). Short-term Exposure to One Dialect Affects Processing of Another. *Language and Speech, 53,* 447–471. doi: 10.1177/0023830910372489

Hay, J., Nolan, A., & Drager, K. (2006). From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review, 23,* 351–379. doi: 10.1515/TLR.2006.014

Heald, S. L. M., & Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience, 8*. doi: 10.3389/fnsys.2014.00035

Kabak, B., & Idsardi, W. J. (2007). Perceptual Distortions in the Adaptation of English Consonant Clusters: Syllable Structure or Consonantal Contact Constraints? *Language and Speech, 50,* 23–52. doi: 10.1177/00238309070500010201

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42,* 627–633. doi: 10.3758/BRM.42.3.627

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122,* 148–203. doi: 10.1037/a0038695

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition, 107,* 54–81. doi: 10.1016/j.cognition.2007.07.013

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition, 9,* 119–135. doi: 10.1017/S1366728906002483

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In M. Goldrick, V. Ferreira, & M. Miozzo (eds.), *The Oxford handbook of language production*, pp. 165–181. Oxford: Oxford University Press.

Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language, 92,* 217–233. doi: 10.1016/j.jml.2016.06.010

Lentz, T. O., & Kager, R. W. J. (2015). Categorical phonotactic knowledge filters second language input, but probabilistic phonotactic knowledge can still be acquired. *Language and Speech, 58,* 387–413. doi: 10.1177/0023830914559572

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition, 17,* 673–680. doi: 10.1017/S1366728913000606

Llanos, F., & Francis, A. L. (2017). The Effects of Language Experience and Speech Context on the Phonetic Accommodation of English-accented Spanish Voicing. *Language and Speech*, *60,* 3–26.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE, 7,* e43230. doi: 10.1371/journal.pone.0043230

Ministry of Education, Culture, and Sport of Spain. (2006). *Diploma de español como lengua extranjera (Diploma of Spanish as a Second Language) Examination*.

Molnar, M., Ibáñez-Molina, A., & Carreiras, M. (2015). Interlocutor identity affects language activation in bilinguals. *Journal of Memory and Language, 81,* 91–104. doi: 10.1016/j.jml.2015.01.002

Namjoshi, J., Tremblay, A., Spinelli, E., Broersma, M., Martínez-García, M. T., Connell, K., Cho, T., & Kim, S. (2015). Speech segmentation is adaptive even in adulthood: Role of the linguistic environment. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences* (paper no. 0676). Glasgow, UK: University of Glasgow.

Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience, 31,* 4–18. doi: 10.1080/23273798.2015.1081703

Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics, 41,* 407–420. doi: 10.1016/j.wocn.2013.07.005

Olson, D. J. (2016a). The gradient effect of context on language switching and lexical access in bilingual production. *Applied Psycholinguistics, 37,* 725–756. doi: 10.1017/S0142716415000223

Olson, D. J. (2016b). The role of code-switching and language context in bilingual phonetic transfer. *Journal of the International Phonetic Association, 46,* 263–285. doi: 10.1017/S0025100315000468

Otheguy, R., García, O., & Reid, W. (2015). Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review, 6,* 281–307. doi: 10.1515/applirev-2015-0014

Parlato-Oliveira, E., Christophe, A., Hirose, Y., & Dupoux, E. (2010). Plasticity of illusory vowel perception in Brazilian-Japanese bilinguals. *Journal of the Acoustical Society of America, 127,* 3738–3848. doi: 10.1121/1.3327792

Pitt, M. A. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception & Psychophysics, 60,* 941–951. doi: 10.3758/BF03211930

Polivanov, E. (1931). La perception des sons d'une langue étrangère [Perception of sounds of a foreign language]. *Travaux du cercle linguistique de Prague, 4,* 79–96.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110,* 474–494.

Samuel, A. G., & Larraza, S. (2015). Does listening to non-native speech impair speech perception? *Journal of Memory and Language, 81,* 51–71. doi: 10.1016/j.jml.2015.01.003

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25,* 421–436. doi: 10.1006/jpho.1997.0051

Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics, 43,* 26–37. doi: 10.1016/j.wocn.2014.01.004

Spinelli, E., & Gros-Balthazard, F. (2007). Phonotactic constraints help to overcome effects of schwa deletion in French. *Cognition, 104,* 397–406. doi: 10.1016/j.cognition.2006.07.002

Tice, M., & Woodley, M. (2012). Paguettes & bastries: Novice French learners show shifts in native phoneme boundaries. *UC Berkeley Phonology Lab Annual Report*, 72–75.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America, 119,* 597–607. doi: 10.1121/1.2141003