

First-language influence on second language speech perception depends on task demands

Language and Speech

1–24

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0023830920983368

journals.sagepub.com/home/las**Max R. Freeman** 

St. John's University, USA

Henrike K. Blumenfeld

San Diego State University, USA

Matthew T. Carlson

The Pennsylvania State University, USA

Viorica Marian

Northwestern University, USA

Abstract

While listening to non-native speech, second language users filter the auditory input through their native language. We examined how bilinguals perceived second language (L2 English) sound sequences that conflicted with native-language (L1 Spanish) constraints across three experiments with different task demands. We used the L1 Spanish phonotactic constraint (i.e., rule for combining speech sounds) that vowels must precede s + consonant clusters (e.g., Spanish: *estricto*, “strict”). This L1 Spanish constraint may influence Spanish-English bilinguals’ processing of L2 English words such as *strict* because of a missing initial vowel, as in *estrict*. We found that the extent to which bilinguals were influenced by the L1 during L2 processing depended on task demands. When metalinguistic awareness demands were low, as in the AX word discrimination task (Experiment 1), cross-linguistic effects were not observed. When metalinguistic awareness demands were high, as in the vowel detection (Experiment 2) and lexical decision (Experiment 3) tasks, response times demonstrated that bilinguals were influenced by the L1 constraint when processing L2 words beginning with an s + consonant. We conclude that bilinguals are cross-linguistically influenced by L1 phonotactic constraints during L2 processing when metalinguistic demands are higher, suggesting that L2 input may be mapped onto L1 sub-lexical representations

Corresponding author:

Dr. Max R. Freeman, Department of Communication Sciences and Disorders, St. John's University, 8000 Utopia Parkway, Jamaica, New York 11439, USA.

Email: freemanm@stjohns.edu

during perception. These results extend previous research on language co-activation and speech perception by providing a more fine-grained understanding of task demands and elucidating when and where cross-linguistic phonotactic access is present during bilingual comprehension.

Keywords

Bilingualism, second language acquisition, phonotactic constraints, phonology, lexical access, metalinguistic demands

Introduction

When hearing non-native sounds or sequences of sounds, second language (L2) listeners may interpret the auditory input through the filter of their native language (L1). Illicit input is often perceptually repaired whereby the input is “fixed” to conform to the L1 (Carlson et al., 2016; Darcy et al., 2007; Dupoux et al., 2011; Hallé et al., 2008; Weber & Cutler, 2006). Perceptual repair is especially apparent when a sound sequence is not present in the L1. For example, Spanish speakers encounter conflict when producing an English word such as *strict*, which frequently becomes *estric*, as the syllable structure for s+consonant cluster (s+c) onset words does not exist in Spanish. English permits s+c word onsets, as in *strict*, as well as vowel+s+c (v+s+c) word onsets, as in *estimate*. However, Spanish requires the presence of a vowel at the beginning of all s+c sequences, as in *estric*. Adding a vowel, specifically, “e,” is the process whereby the impossible s+c cluster is adapted to Spanish phonotactics, also converting loanwords like English *snob* to Spanish *esnob*, yielding a permissible v+s+c sequence. In addition, during production, L1 Spanish speakers often add a vowel to the onset of English s+c words, as in *estric* (Yavas & Someillan, 2005). The question then arises, how do bilinguals *perceive* L2 words that conflict with L1 phonotactic constraints? We address this question in the current study by testing Spanish-English bilinguals and English monolingual controls across three measures of speech perception that vary in task demands: AX word discrimination, vowel detection, and lexical decision.

1.1 Parallel processing in bilinguals

Processing through an L1 filter implies that the L1 may be accessed when speaking or listening to the L2. Evidence for parallel processing in bilinguals comes from Marian and Spivey (2003), who found that L1 (Russian) was activated while Russian-English bilinguals listened to L2 English. Participants viewed four pictures on a visual display. Two of the pictures overlapped phonologically across Russian and English (e.g., a picture of a *marker* and a *stamp*, Russian: “*marka*”), while two filler items contained no phonological overlap with the target and competitor. The bilinguals fixated the target *marker* and the between-language phonological competitor *stamp* more than the filler items, suggesting parallel processing. Importantly, Marian and Spivey were among the first to demonstrate that bilinguals process both languages simultaneously during single-language comprehension. The present investigation examines how bilinguals mapped L2 acoustic signals that conflict with L1 phonological representations.

1.2 Perceptual repair in monolinguals and bilinguals

Previous investigations suggest that monolinguals perceptually repair non-native spoken input to conform to the phonotactic constraints of their language. For example, Japanese monolinguals in

Dupoux et al. (1999) repaired an illegal VCCV non-word sequence (e.g., *ebzo*) to a legal VCVCV non-word sequence (e.g., *ebuzo*), leading to difficulty distinguishing between the two in an ABX discrimination task. In an ABX discrimination task, two stimuli are presented sequentially (A and B) and then a third stimulus (X) is played that is either the same as A or B. Participants decide if the X stimulus is the same as the A or B stimulus. Findings suggest that *ebzo* was perceptually repaired to *ebuzo* using epenthesis (addition of a vowel).

Monolinguals perceptually repair input that conflicts with their language; however, does knowing a second language affect perceptual repair? Indeed, Carlson et al. (2016) found that bilingual status altered perception of an illusory vowel. Carlson et al. aurally presented Spanish speakers of English with Spanish-like non-words that conflicted with Spanish phonotactic constraints, such as *stid* (/stid/, “estid”) using vowel detection and AX discrimination tasks. Findings suggest that Spanish-dominant, but not English-dominant bilinguals, heard the illusory vowel onset in the s+c non-words, favoring “e.” Dominance in English was associated with a weakened Spanish vowel illusion. The results from Carlson et al. suggest that knowledge of English, a language in which the constraint does not exist, affects whether bilinguals experience an illusory vowel onset, where L1-like sounds conform to the L1. In the current study, we seek to examine whether Spanish-English bilinguals are cross-linguistically influenced by L1 constraints (i.e., adding a vowel). This process may result in two potential outcomes: 1) L2 auditory input being processed through an L1 filter, therefore potentially mapping L2 input onto existing L1 representations; and 2) perceptual repair of L2 words to L1 representations.

1.3 Perceptual repair and metalinguistic demands

In addition to language dominance, metalinguistic demands induced by a task may also influence the presence of perceptual repair in bilinguals. Parlato-Oliveira et al. (2010) tested early and late Japanese-Portuguese bilinguals across explicit (vowel detection) and implicit (forced-choice recall) measures of perceptual repair containing Japanese- and Portuguese-like stimuli. The stimuli exploited the same Japanese and Portuguese phonotactic VCCV constraint as in Dupoux et al. (1999), where the repair vowel is “u” in Japanese, but “i” in Brazilian Portuguese. Both participant groups demonstrated some evidence of both the Japanese (L1) and Brazilian Portuguese (L2) vowel perceptual repairs in vowel detection, but only the L1 Japanese repair affected performance in the sequence recall task. In the sequence recall task, participants were not directly cued into the presence of the L1 Japanese repair vowel as they were in the vowel detection task. The authors suggested that listeners could apply L2-specific perceptual patterns in metalinguistic tasks, in this case, involving participants explicitly reflecting on what was heard. However, the L2-specific patterns were not applied in the more implicit task of remembering a sequence of items, where the same phonetic information was needed, but explicit reflection on this information was not. Thus, the metalinguistic demands of the task affected the extent to which perceptual repair occurred. To confirm and extend these initial findings, in the present study, we manipulated task demands across three measures of speech perception in early Spanish-English bilinguals using the Spanish v+s+c constraint. We therefore examine how metalinguistic load influences bilinguals’ L2 (English) perception of input that conflicts with the L1 (Spanish) constraint.

1.4 L1 filtering during L2 processing

An alternative account to perceptual repair in bilinguals is that they *filter* the L2 acoustic signal through existing L1 representations. In Spanish-English bilinguals, if this erroneous mapping (filtering) of English s+c onsets occurs during spoken word comprehension, then L1 Spanish

speakers would process words such as *strict* cross-linguistically, for example accessing the “e” onset (e.g., *estric*). Given previous work on perceptual repair in monolinguals and bilinguals, the more bilinguals have exposure to and experience with their second language, the less likely it is that perceptual repair of L1-conflicting input during L2 processing occurs (e.g., Carlson et al., 2016). However, the literature on parallel processing in bilinguals suggests that a bilingual’s two languages are simultaneously active during comprehension even as the L2 becomes the more proficient language (e.g., Marian & Spivey, 2003), and therefore, the L1 filter is online during L2 processing.

Empirical evidence indeed suggests that bilinguals process L2-specific speech sound sequences using an L1 filter (Lentz & Kager, 2015; Weber & Cutler, 2006). Weber and Cutler (2006) gave L1 German speakers of English, as well as English monolinguals, English nonsense sequences in which participants detected when they heard an English word. The sound sequences spanning the onset of the real word, however, were required to be interpreted as a syllable boundary according to either L1 German phonotactics, L2 English phonotactics, both or neither. While German-English bilinguals and English monolinguals were almost equally successful in identifying the embedded word when its onset coincided with English syllabification, the German-English bilinguals were also aided when the word’s onset coincided with a syllable boundary only according to the German constraints. The results suggest that bilinguals become sensitive to L2 phonotactics but that they are still influenced by L1 phonotactic constraints during L2 processing. Weber and Cutler used nonsense sequences with embedded real words. In the current study, we included single words to examine real-world generalizability, and to investigate the extent to which top-down lexical processes affect mapping of the acoustic signal.

Further evidence for influence of L1 phonotactic constraints during L2 processing comes from Lentz and Kager (2015). Dutch monolinguals, L1 Japanese speakers of Dutch, and L1 Spanish speakers of Dutch performed a cross-modal lexical decision task. Participants were first presented with auditory *v+s+c* primes that were legal across Dutch, Japanese and Spanish. After the auditory primes, visual lexical decision *s+c* targets were presented, which were legal in Dutch but illegal in Japanese and Spanish. Results suggest that the native-Spanish group accessed the vowel “e” onset with the illicit *s+c* onsets. These findings demonstrate that the native-Spanish group used an L1 filter when processing L2 sounds. In Weber and Cutler (2006), as well as in Lentz and Kager, implicit measures of vowel perception were used. In the current study, as in Carlson et al. (2016), we employed implicit and explicit measures of perception. In addition, we administered perceptual tasks that varied in metalinguistic load, since Parlato-Oliveira et al. (2010) demonstrated that task demands influenced how input was perceived.

1.5 The present study

The current investigation tests the extent to which L2 perceptual representations are influenced by the L1. Building on findings by Carlson et al. (2016), Freeman et al. (2016), Lentz and Kager (2015), Parlato-Oliveira et al. (2010), and Weber and Cutler (2006), we examine whether bilinguals are influenced by L1 phonotactic constraints during L2 comprehension across tasks that vary in metalinguistic demands. Our main objective is to determine whether bilinguals are guided cross-linguistically by L1 phonotactics when *perceiving* L2 sound sequences. Understanding whether and when the L2 is processed through an L1 filter would provide insight into the extent to which cross-linguistic influences impact bilingual speech and language processing. Moreover, L1 phonological representations affect L2 perception when a task’s metalinguistic demands are high (e.g., Parlato-Oliveira et al., 2010), such as when participants are explicitly segmenting phonemes from

the speech stream and/or detecting the presence of a vowel. Therefore, if metalinguistic demands influence the extent to which bilinguals perceive English stimuli according to Spanish-like phonotactic patterns, then results would contribute novel evidence to the involvement of top-down processes during perception (e.g., Dupoux et al., 2001). The hypothesis is that, when metalinguistic load is high, Spanish-English bilinguals rely on L1 phonological representations to perceive L2 English and English-like sounds as Spanish-like.

We tested our hypothesis across three tasks that varied in metalinguistic demands. First, in Experiment 1 (AX word discrimination), implicit perception of English sounds that conflicted with Spanish phonotactic constraints was examined. This discrimination task tapped into lower-level perceptual processes as only same/different judgments were made on English word pairs. Participants compared memory traces of two, newly presented items. Second, in Experiment 2 (vowel detection), metalinguistic demands were higher than in Experiment 1 as knowledge of and attention to vowels and consonants was required and participants compared English words and English-like non-words with already-established representations. The use of non-words *and* words in the same experiment was an additional extension of previous research (e.g., Carlson et al., 2016; Cuetos et al., 2011; Weber & Cutler, 2006), permitting the investigation into whether phonotactic knowledge was influenced by top-down lexical processes. As words exist within the lexicon, it was likely that participants recruited top-down lexical information in addition to top-down phonotactic information associated with non-words. Third, in Experiment 3 (lexical decision), participants matched English words and English-like non-words with existing representations, as in Experiment 2; however, this task required the most abstract level of metalinguistic awareness, as participants also had to search their lexicon for word and non-word matches. Lexical decision tasks are particularly difficult for bilinguals, as they search lexical entries across both languages to find a match (Soares & Grosjean, 1984). The use of three tasks that differ in metalinguistic demands tests the extent to which metalinguistic awareness influences cross-linguistic effects, specifically how L2 representations are mapped, or shaped by the L1.

2 Experiment 1: AX word discrimination in L2

The purpose of Experiment 1 was to examine whether bilinguals were cross-linguistically influenced by L1 phonotactic constraints during L2 comprehension in a task with low metalinguistic demands. In the AX word discrimination task, participants simply decided if two consecutive auditory stimuli (A and X) were the same or different. Previous studies employing the AX word discrimination task have used stimuli that were nearly acoustically identical, except for a small amount of acoustic material, such as an additional vowel, and in all cases the A and X stimuli were the same non-word (e.g., A = *esnid* and X = *snid*) (e.g., Carlson, 2018a; Carlson et al., 2016; Dupoux et al., 1999). In the current study, our goal was to identify whether Spanish-English bilinguals differ in response time and accuracy when the A word represents a conflict with the L1 Spanish vowel constraint and the X word probes for presence of such cross-linguistic conflict (e.g., A = *strict* and X = *egg*), as opposed to when the A word does not conflict with the constraint (e.g., A = *work* and X = *egg*). In the conditions of interest, the s+c word (A = *strict*) always preceded the “e” onset word (X = *egg*), as we expected the phonotactic effect to occur in this order and not necessarily in the reverse. The reasoning behind the chosen order was that the L1 perceptual filter or perceptual repair does not come online until the listener knows that a consonant follows the /s/. An accurate “different” judgment could thus be made based on the first few milliseconds of hearing the s+c word, if it were presented second. This would not occur when the s+c word is presented first because if

it is held in memory as beginning with a vowel onset (e.g., “e”), it would temporarily match the actual “e” at the onset of the second word (e.g., *egg*). Therefore, the L1 filter/perceptual repair would be turned on due to the violation of the L1 v+s+c constraint, resulting in cross-linguistic competition. Then, by hearing the “e” onset word second, its identification would be temporarily hindered due to the competition resulting from the s+c word. This order of stimulus presentation is also supported by the findings of Freeman et al. (2016), who used a cross-modal lexical decision task in which an L2 s+c auditory prime (e.g., *stable*) preceded an “e” onset lexical decision target (e.g., *esteriors*). Results demonstrated significant effects of cross-linguistic activation of the L1 v+s+c phonotactic constraint during L2 processing through slower response times to these conflicting prime/target pairs, relative to non-conflicting controls. Accuracy rates would not likely be affected, since the second phoneme completely disambiguated the words. Alternatively, if exploiting memory traces of two newly presented items was not sufficiently sensitive to tap into speech perception (i.e., the AX word discrimination task’s low task demands), then bilinguals would not show differences in reaction times and accuracy rates to s+c onset words followed by “e” onset words, relative to controls. This outcome would suggest that L1 cross-linguistic influences from phonotactic constraints might not emerge during L2 processing when metalinguistic demands are low, in line with Parlato-Oliveira et al. (2010).

2.1 Experiment 1 methods

2.1.1 Participants. Twenty-five English monolinguals (five males) and 26 Spanish-English bilinguals (nine males) were tested in Evanston, Illinois and in San Diego, California. Monolinguals and bilinguals were tested in both locations. All participants had normal or corrected-to-normal vision and no history of a neurological impairment. Participants completed various cognitive and linguistic assessments to create equivalence across monolingual and bilingual groups. The Language Experience and Proficiency Questionnaire (LEAP-Q) (Kaushanskaya et al., 2019; Marian et al., 2007) was given at the start of the study and served as an index of history and proficiency for the monolinguals and bilinguals. English monolinguals were tested who reported Spanish speaking proficiency of less than or equal to 3 (1-10 scale) or another foreign language speaking proficiency of less than or equal to four on the LEAP-Q. Bilinguals were native-Spanish speakers. Bilinguals’ daily exposure to Spanish was around 30% and they had acquired English upon entering primary school (around age five). Monolinguals and bilinguals differed on English age of acquisition ($p < .001$), current English exposure ($p < .001$), foreign accent in English ($p < .01$), and self-reported English proficiency ($p < .001$). The remaining linguistic and cognitive measures were given after the conclusion of the main experimental tasks. Vocabulary knowledge was indexed with the Peabody Picture Vocabulary Test-3 (PPVT-3) (Dunn & Dunn, 1997) for monolinguals and bilinguals, and the Vocabulario en Imágenes Peabody (TVIP) (Dunn et al., 1997) for the bilinguals only. The Wechsler Abbreviated Scale of Intelligence (WASI) (PsychCorp, 1999) examined non-verbal cognitive reasoning. Working memory was assessed through the backward digit span task (numbers reversed, Woodcock et al., 2001; 2007). A non-linguistic Stroop task indexed competition resolution abilities (adapted from Blumenfeld & Marian, 2014). See Table 1 for participant information. Monolinguals and bilinguals were matched on age, English receptive vocabulary (Dunn & Dunn, 1997), non-verbal cognitive reasoning (WASI; PsychCorp, 1999), working memory (backward digit span) (Woodcock et al., 2001; 2007), and the Stroop Effect (Blumenfeld & Marian, 2014).

Table 1. Linguistic and cognitive background of Spanish-English bilingual (n = 26) and English monolingual (n = 25) participants.

	Bilinguals' Mean (SE)	Monolinguals' Mean (SE)	P-value
Age	22.42 (0.84)	22.08 (0.34)	0.78
Age of Spanish acquisition	0	–	–
Age of English acquisition	5.42 (0.54)	0	< 0.01
Current exposure to Spanish	34.88% (4.10)	1.47% (0.96)	
Current exposure to English	60.81% (4.56)	98.65 (0.69)	< 0.01
Foreign accent in Spanish (1–10 scale)	1.69 (0.35)	–	
Foreign accent in English (1–10 scale)	3.15 (0.59)	1.56 (0.64)	0.07
Self-reported Spanish proficiency (1–10 scale)	9.09 (0.71)	–	
Self-reported English proficiency (1–10 scale)	9.01 (0.11)	9.61 (0.08)	< 0.01
Spanish receptive vocabulary (TVIP, standard score)	112.35 (1.79) Range = 33.00 Max Score = 124	–	
English receptive vocabulary (PPVT, standard score)	107.39 (2.12) Range = 41 Max Score = 129	108.04 (2.44) Range = 66 Max Score = 129	0.49
WASI, matrix reasoning	28.50 (0.99) Range = 20 Max Score = 33	28.89 (0.61) Range = 14 Max Score = 33	0.88
Backward digit span	8.46 (0.95) Range = 17 Max Score = 18	11.39 (0.81) Range = 15 Max Score = 17	0.11
Stroop Effect (Incongruent – Congruent Trials) Smaller Stroop Effect = better performance	85.50ms (0.01) Range = 192.13ms Max Score = 14.09ms	88.44ms (0.01) Range = 180.54ms Max Score = 30.36ms	0.78

2.1.2 Materials. The AX word discrimination task examined implicit perception of the Spanish “e” onset in English s+c onset words. The “st” and “sp” consonant clusters were employed in Experiment 1 as these are illegal consonant clusters in Spanish without the obligatory “e” at the word onset. These two consonant clusters were previously used in Freeman et al. (2016), which demonstrated significant effects of parallel processing of L1 phonotactic constraints during L2 comprehension. To maintain a strictly English testing environment, only L2 English stimuli were used throughout the study. The stimulus set included s+c onset words, control onset words, and an additional set of “e” onset words, which was used across the A and X lists to form the AX word pairings. Trials consisted of match and mismatch (same/different) pairs. Three mismatch conditions of interest were included for analyses:

- 1) Spanish conflicting: s+c word followed by e-onset word (e.g., *strict* → *egg*) (**S→E**);
- 2) Control: control onset followed by e-onset (e.g., *work* → *egg*) (**C→E**);
- 3) ‘E’ onset control: e-onset followed by e-onset (e.g., *effort* → *egg*) (**E→E**).

The C→E and E→E conditions served as controls for the S→E condition. The C→E condition did not provide the potential for bilinguals to access the L1 phonotactic constraint since there was no

violation of the L1 v+s+c rule. The E→E condition served as a positive control, as it revealed whether the “e” onset stimuli were responsible for potential effects relative to the S→E trials. Additional filler (also considered non-critical) conditions included match and mismatch trials so that participants could not 1) guess the purpose of the experiment and 2) learn to predict that the answer was “different” each time they heard an s+c word.

The three types of words (s+c onset, control onset, and “e” onset) were controlled for the following lexical characteristics ($ps > .05$): number of letters in English and in Spanish (translation), English and Spanish (translation) lexical frequency, English and Spanish orthographic neighborhood density (CLEARPOND) (Marian et al., 2012). See Appendix 1 for stimuli from Experiment 1.

A total of 120 stimuli were created, comprised of 24 s+c words, 24 “e” onset words, and 72 control words. Within the AX word pairs, 44% of trials consisted of match responses, while 56% of trials included mismatch responses. Only 28% of all trials contained an “e” onset word as the X stimulus. Experiment 1 contained a total of 228 trials (12 practice and 216 experimental) and the experiment was divided into two blocks. The trials were pseudo-randomized so that no more than two consecutive trials contained s+c onsets. Trial order was counterbalanced across participants by reversing the order of presentation.

2.1.3 Procedure. Monolinguals were tested by a male proficient in English. Bilinguals were tested by a male proficient in English and Spanish. Participants were seated in a quiet room in front of an iMac computer and were first administered the Language Experience and Proficiency Questionnaire (LEAP-Q) (Kaushanskaya et al., 2019; Marian et al., 2007) to obtain linguistic background information. Participants next performed all experimental measures (in order: AX word discrimination, vowel detection and lexical decision). This order was chosen so that the least metalinguistically explicit task would come first. In addition, it was likely that the purpose of the studies would be further disguised by administering the AX word discrimination task before the vowel detection and lexical decision tasks. AX word discrimination implicitly measured perception through low-level same/different judgments, while vowel detection and lexical decision explicitly asked participants about specific characteristics of the auditory stimuli. The AX word discrimination task was programmed in MatLab (Psychtoolbox add-on). Reaction time and accuracy data were collected with keyboard button presses. The task was controlled by an iMac 3.3 GHz Intel Core i5 running MatLab 2011a, and the display included a 27-inch monitor with a screen resolution of 5120x2880. The stimuli were recorded in a sound attenuated room (44,100 Hz, 16 bits) by a male native speaker of English. The audio recording was normalized (via audio compression), split into individual audio files in Praat (Boersma & Weenink, 2013), and exported into MatLab (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

Participants were instructed in English to listen to two consecutive English words and then indicate if the two words were the same or different as quickly and as accurately as possible. After the instructions and 12 practice trials, participants performed the experimental task in which they first heard the A stimulus followed by the X stimulus. There was a 250ms inter-stimulus interval between the two words. During presentation of the A stimulus and the X stimulus, participants viewed a central fixation crosshair on the computer screen. Participants were then asked if the two words they heard were the same or different, see Figure 1. Reaction times were measured from the onset of the second auditory (X) stimulus. Presentation of the question about whether the words were the same or different lasted until the participant made a response. The left/right Shift keys represented same/different responses.

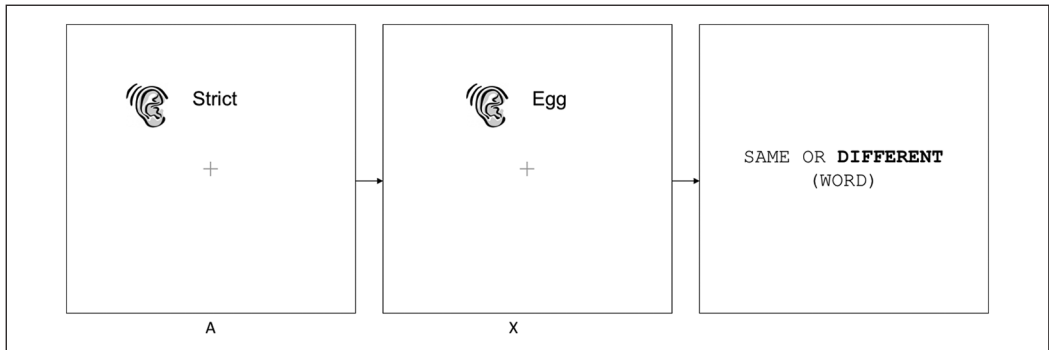


Figure 1. Sample trial from the AX word discrimination task in Experiment 1. In this example, participants heard *strict* followed by *egg* and had to decide if the words they heard were the same or different.

Participants were given one short, but untimed, break halfway through the experiment. The total time to complete this task was approximately 12 minutes.

2.1.4 Coding and analyses. For the AX word discrimination task, the within-subjects independent variable was critical trial type (Spanish-conflicting $S \rightarrow E$, $C \rightarrow E$ control, and $E \rightarrow E$ control) and the between-subjects independent variable was Language Group (monolingual and bilingual), resulting in a 3×2 mixed-factorial design. Incorrect trials and trials 2.5 standard deviations above and below the mean reaction time for each participant were excluded from the analyses (2% of the data). Accuracy and reaction time data for critical comparisons, including the Spanish-conflicting ($S \rightarrow E$), control ($C \rightarrow E$), and “e” onset control ($E \rightarrow E$) conditions, were fit using two mixed effects models with the lme4 package in R (Bates et al., 2015; R Core Team, 2016). Accuracy was fit using a generalized linear mixed effects logistic regression model and reaction time was fit using a linear mixed effects model. The following factors were included: Trial type (orthogonal contrasts, centered), Language Group (monolingual and bilingual; centered), Items, and Participants. Both models included fixed effects of Condition by Language Group, random intercepts and a random slope for Condition by Participant, and by-item random intercepts. The models therefore accounted for participant and item variability. The reaction time model included log transformation of reaction times. The accuracy model failed to converge due to ceiling effects; therefore, we only report raw means for accuracy data.

2.2 Experiment 1 results

2.2.1 Accuracy on the AX word discrimination task (see Table 2)

Table 2. Raw accuracy mean percentages and (Standard Error) across language groups and conditions on Experiment 1, AX Word Discrimination.

Accuracy	$S \rightarrow E$	$C \rightarrow E$	$E \rightarrow E$
Bilinguals	99.27 (0.57)	99.27 (0.34)	98.91 (0.37)
Monolinguals	99.46 (0.30)	99.09 (0.36)	99.63 (0.25)

2.2.2 Reaction time on the AX word discrimination task. Log transformed reaction times for bilinguals and monolinguals were analyzed using a linear mixed effects regression model. The model revealed only a significant main effect of Language Group, $\beta = 0.09$, $SE = 0.03$, $t = 2.60$, $p = 0.01$. Bilinguals ($M = 1096.37\text{ms}$, $SE = 22.99$) responded more slowly overall than monolinguals ($M = 991.312\text{ms}$, $SE = 12.03$). See Figure 2 for differences across bilinguals and monolinguals, and conditions of interest. Overall, these results suggest that there are no low-level effects of cross-linguistic interaction or perceptual repair in bilinguals in this task, with the only main effect being Language Group.

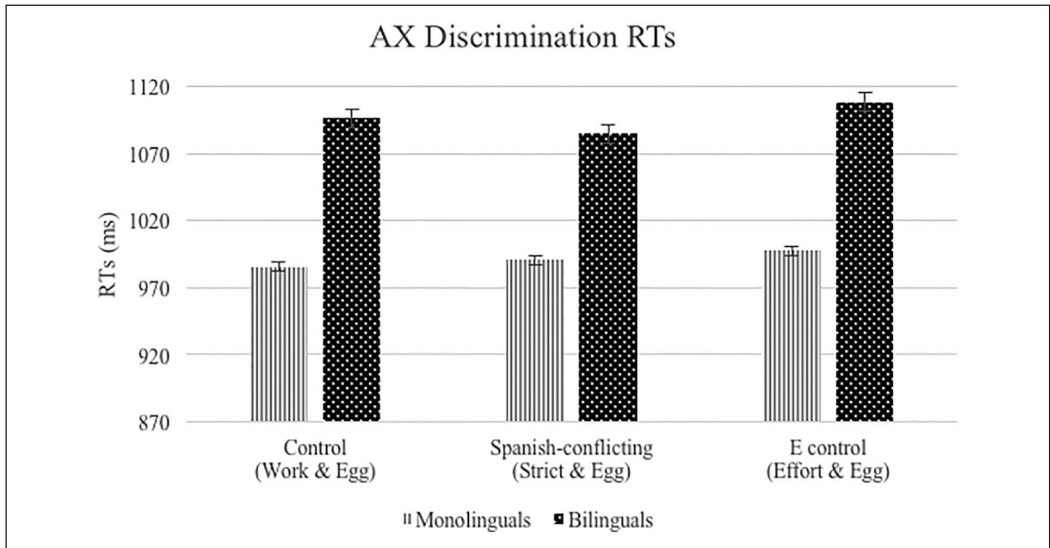


Figure 2. Monolingual and bilingual reaction times (RTs) on the AX word discrimination task. Group effects indicated that bilinguals were slower than monolinguals overall. There were no between-condition differences within bilinguals and monolinguals.

2.3 Experiment 1 discussion

The prediction for Experiment 1 was that if L1-driven filtering occurred, bilinguals would respond differently to Spanish-conflicting trials (e.g., *strict* followed by *egg*) relative to control trials (e.g., *work* followed by *egg*). Alternatively, if a lower metalinguistic load muted influence of cross-linguistic processing, then such effects would be absent. Indeed, reaction time rates were similar across all conditions within monolinguals and bilinguals, suggesting the latter prediction that cross-linguistic influences were not present in AX word discrimination. Moreover, there were no effects of age of acquisition, dominance, or proficiency on bilinguals' response times. A tentative explanation for this pattern of results is found in the low-level perceptual nature of this task. Participants paid attention only to the combination of sounds within the A and X stimuli, without the need to focus on a specific aspect of the stimuli (vowel detection) or to access the lexicon (lexical decision); therefore, metalinguistic demands in AX word discrimination were lower (see Parlato-Oliveira et al., 2010 for further discussion). Moreover, an illusory vowel did not sufficiently interfere with a decision of "different," even for the *strict* followed by *egg* trials. This could be because, from the onset of *strict*, listeners can recognize the presence of the "s" in the word's

onset, even if they access the illusory vowel, and even if they recognize the first syllable as [es], that would be enough to differentiate *strict* from *egg*. There was a lack of L1 mapping in the AX word discrimination task even though ordering of the A (e.g., *strict*) and X (e.g., *egg*) stimuli was designed to create ambiguity at the onset of the X for participants who perceived the L1. Furthermore, participants did not show a difference between E→E (e.g., *effort*→*egg*) trials versus C→E (e.g., *work*→*egg*) trials, suggesting that any potential response time effects that are related to brief acoustic features at the onset of the stimuli are obscured.

3 Experiment 2: Vowel detection in L2

In Experiment 2, we examined whether bilinguals perceived a non-native sound sequence in line with L1 phonological representations, specifically the English s+c onset, in a vowel detection task. The vowel detection task's metalinguistic demands were higher, relative to AX word discrimination (Experiment 1), as the explicit focus was on the stimulus onset, with participants identifying if a vowel was present at the word onset, with consonant (e.g., *strict*) and vowel (e.g., *issue*) initial phonemes. Previous studies have used the task as an index of perceptual processing of illicit sound sequences (e.g., Carlson et al., 2016; Cuetos et al., 2011). English words and English-like non-words that conflicted with the Spanish v+s+c constraint were included to examine whether top-down lexical influences affected cross-linguistic phonological processing. If bilinguals used the Spanish vowel filter during English auditory comprehension, then they would demonstrate differences in response times and accuracy rates to Spanish-conflicting s+c stimuli (words and non-words) relative to control onset stimuli (words and non-words). Further, if differential effects were found for Spanish-conflicting words relative to non-words, then top-down lexical influences may affect the presence of the L1 filter.

3.1 Experiment 2 methods

3.1.1 Participants. The same participants were tested in Experiment 2 as in Experiment 1.

3.1.2 Materials. The vowel detection task examined explicit L1 speech perception in an L2 context by measuring perception of the Spanish “e” onset constraint in English words and English-like non-words with s+c onsets. Stimuli for the vowel detection task contained two types of words (s+c and control), adapted from Experiment 1. Control words contained consonant or vowel onsets. All word stimuli were controlled for the following lexical characteristics ($ps > .05$): number of letters in English and in Spanish (translation: e.g., *strict/estricto* and *strong/fuerte*), English and Spanish (translation) lexical frequency, and English and Spanish orthographic neighborhood density (CLEARPOND) (Marian et al., 2012). Stimuli also included two types of non-words (s+c and control). Control non-words contained consonant or vowel onsets. All non-word stimuli were controlled for the following lexical characteristics ($ps > .05$): number of letters and neighborhood density (CLEARPOND) (Marian et al., 2012). See Appendix 2 for stimuli.

A total of 192 stimuli were created, consisting of 24 s+c words, 24 s+c non-words, 72 control onset words (24 consonant and 48 vowel onset), and 72 control onset non-words (24 consonant and 48 vowel onset). The ratio of words to non-words and of vowel to consonant onsets was 1:1. The experiment was divided into two intermixed blocks and items were pseudo-randomized such that no more than two consecutive trials contained s+c stimuli. The trial order was counterbalanced across participants by reversing the order of presentation.

3.1.3 Procedure. The vowel detection task preceded the lexical decision task (Experiment 3) since the vowel detection task was the primary method of explicitly observing cross-linguistic phonotactic effects. The vowel detection task was programmed in MatLab (Psychtoolbox add-on) (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Button presses to the keyboard allowed for the collection of accuracy and reaction time data. Experiment 2 was administered on the same computer as Experiment 1, using the same process for recording and splitting the audio files.

Participants were instructed in English to pay attention to the beginning sound of the word or non-word they heard. Participants were asked to respond to the stimulus they heard as quickly and as accurately as possible. A “yes” response indicated that the participant detected a vowel at the stimulus onset, while a “no” response signified that a consonant was identified at the onset. After the instructions and 12 practice trials, participants performed the experimental task in which they heard a word (s+c onset and control onset) or non-word (s+c onset and control onset). During presentation of the stimulus, participants viewed a central fixation crosshair on the computer screen for the duration of the stimulus, immediately followed by a prompt on the visual display asking if a vowel was heard (see Figure 3). Reaction times were measured from the onset of the auditory stimulus. The crosshair and subsequent prompt were presented in the center of a white screen in black, size 16 font, Courier, and the left/right Shift keys represented yes/no responses. Presentation of the prompt lasted until the participant made a response. Accuracy and reaction times during vowel detection were measured. Participants were given one short break halfway through the experiment. The total time to complete this task was approximately 10 minutes.

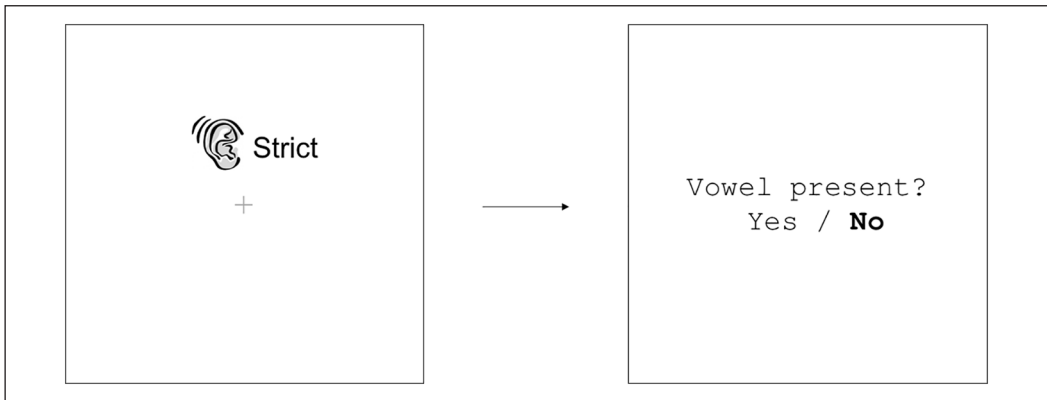


Figure 3. Sample trial in the vowel detection task in Experiment 2. Participants heard *strict* and decided if a vowel was present at the word onset (Yes response= vowel present, No response= no vowel present).

3.1.4 Coding and analyses. The within-subjects independent variables included Lexical Status (word and non-word) and Onset Type (s+c onset and consonant-onset control), and the between-subjects independent variable was Language Group (monolingual and bilingual), yielding a 2x2x2 mixed-factorial design. To examine how irrelevant-language phonotactic constraints affected perception in the vowel detection task, two mixed effects logistic regression models were used to analyze accuracy and reaction time data. The same cleaning and coding procedure was used in Experiment 2 as in Experiment 1, resulting in the removal of 3% of the data. The models included the following factors: Lexical Status (word and non-word; centered); Onset Type (s+c and consonant-onset control; centered), Language Group (monolingual and bilingual; centered); Items, and

Participants. Reaction times were log transformed for the reaction time model. The two models included fixed effects of Lexical Status by Onset Type by Language Group, random intercepts of Lexical Status by Onset Type with a random slope of Participants, and by-item random intercepts. As in Experiment 1, the accuracy model failed to converge due to ceiling effects; therefore, we only report raw means for accuracy data.

3.2 Experiment 2 results

3.2.1 Accuracy on the vowel detection task (see Table 3).

Table 3. Raw accuracy mean percentages (and Standard Error) across language groups and conditions on Experiment 2, vowel detection.

Accuracy	S+C Word	Control Word	S+C Non-word	Control Non-word
Bilinguals	96.38 (0.96)	97.58 (0.54)	96.74 (1.25)	95.77 (0.86)
Monolinguals	97.28 (0.67)	97.28 (0.45)	99.27 (0.42)	95.83 (0.85)

3.2.2 Reaction times on the vowel detection task. We examined reaction times in the vowel detection task with consonant-onset only stimuli. We expected that, if bilinguals perceived an “e” onset, their decisions to s+c words and non-words would be different, relative to controls. The mixed effects model revealed a main effect Language Group, $\beta = 0.11$, $SE = 0.03$, $t = 3.30$, $p < 0.01$, with bilinguals ($M = 1104.57\text{ms}$, $SE = 19.29$) responding slower than monolinguals ($M = 958.16\text{ms}$, $SE = 10.34$). There was a main effect of Lexical Status, $\beta = 0.04$, $SE = 0.02$, $t = 2.20$, $p = 0.04$, with participants responding slower to non-words ($M = 1049.14\text{ms}$, $SE = 18.96$) relative to words ($M = 1013.59\text{ms}$, $SE = 18.36$). There were also interactions between Lexical Status and Language Group, $\beta = 0.04$, $SE = 0.02$, $t = 2.23$, $p = 0.02$; and Onset Type and Language Group, $\beta = -0.04$, $SE = 0.02$, $t = -2.15$, $p = 0.04$. No other main effects or interactions were found.

Following up on the main effects and interactions, new models were added with the same factors within each Language Group. Results demonstrated within the bilingual group only, a main effect of Lexical Status, $\beta = 0.06$, $SE = 0.02$, $t = 3.00$, $p < 0.01$; Onset Type, $\beta = -0.04$, $SE = 0.02$, $t = -2.00$, $p < 0.01$; and a marginal interaction between Lexical Status and Onset Type, $\beta = 0.06$, $SE = 0.02$, $t = 1.50$, $p = 0.07$. Bilinguals responded more slowly to non-words ($M = 1133.11\text{ms}$, $SE = 30.76$) than to words ($M = 1076.03\text{ms}$, $SE = 22.88$), $\beta = 0.05$, $SE = 0.02$, $t = 2.40$, $p = 0.02$. Bilinguals showed slower response times to s+c words ($M = 1112.85\text{ms}$, $SE = 30.29$) than to control words ($M = 1039.09\text{ms}$, $SE = 28.98$), $\beta = -0.07$, $SE = 0.03$, $t = -2.4$, $p = 0.03$. Bilinguals’ response times to s+c non-words ($M = 1136.03\text{ms}$, $SE = 35.45$) relative to control non-words ($M = 1130.17\text{ms}$, $SE = 47.47$) were similar, $\beta = -0.01$, $SE = 0.03$, $t = -0.42$, $p = 0.10$. The monolingual model did not reveal any main effects or interactions. For comparison, monolinguals did not show a large difference in response times to s+c words ($M = 959.65\text{ms}$, $SE = 18.86$) relative to control words ($M = 942.63\text{ms}$, $SE = 22.80$). Monolinguals also did not show a large difference between s+c non-words ($M = 960.11$, $SE = 18.84$) and control non-words ($M = 970.22\text{ms}$, $SE = 18.79$).

Difference scores were employed to graphically illustrate the magnitude of slowing caused by the Spanish-conflicting onset for critical words relative to control words, and for critical non-words relative to control non-words. See Figure 4.

Thus, reaction time results revealed that 1) monolinguals responded more quickly overall in the vowel detection task; 2) bilinguals responded more slowly to non-words relative to words; 3) onset type and lexical status affected participants’ responses to real words, wherein Spanish-conflicting

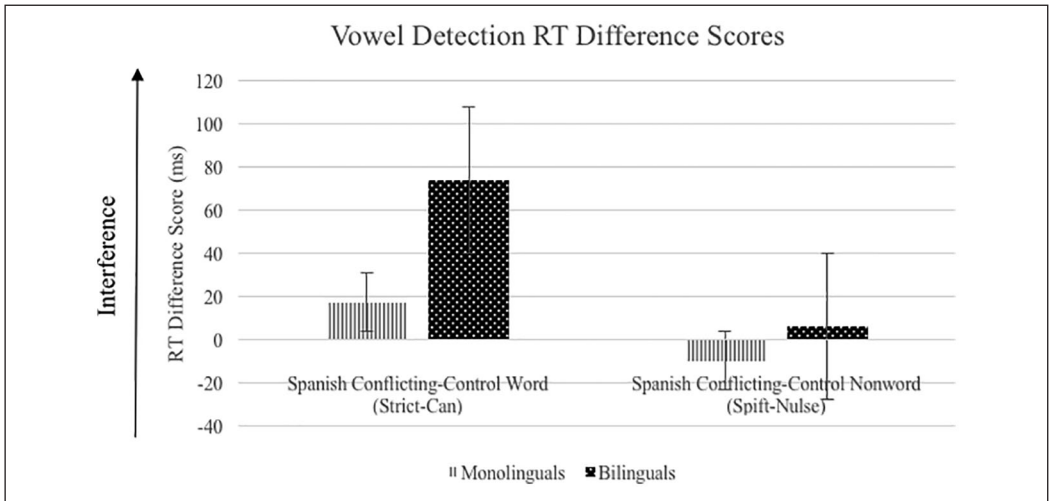


Figure 4. Monolingual and bilingual RT difference scores for Spanish- conflicting minus control trails in word and non - word condition on the vowel detection task. Error bars represent 1 standard error. Bilinguals demonstrated marginally greater interference from Spanish- conflicting words than monolinguals. Within bilinguals, Spanish- conflicting words resulted in increased interference relative to controls, as indicated by a significant RT difference score.

words resulted in slower response times; 4) this effect was driven by bilinguals' slower response times to Spanish-conflicting words relative to control words, suggesting L1 cross-linguistic influence on L2 s+c words.

3.3 Experiment 2 discussion

The results from the vowel detection task contrast with the findings from the AX word discrimination task (Experiment 1). Current findings partially confirm our hypothesis that bilinguals are influenced by the L1 “e” onset filter, mapping L2 English s+c onset words using L1 knowledge. Although bilinguals and monolinguals demonstrated accuracy rates across all conditions at ceiling, only bilinguals exhibited a significant reaction time difference across Spanish-conflicting words relative to control words. Since a purely phonotactic source of the illusory vowel would predict the same results for words and non-words (Dupoux et al., 2001), this latter finding also suggests that top-down lexical processes modulated perceptual processing. Interestingly, the result for words but not non-words suggests that Spanish-like phonotactic processing may have affected word learning in earlier stages of acquiring English, but it may no longer affect the learning of new s+c words in English, for this population (Darcy & Thomas, 2019). Moreover, cross-linguistic influences from the Spanish L1 constraint may have affected the L2 phonological representation of s+c words, as indexed by reaction times only. Given that bilinguals in the current study were tested in their L2 in which s+c is permissible, reaction times, rather than accuracy, were more sensitive to the observed effects of perceptual processing (see General Discussion). No effects of age of acquisition, proficiency, or language dominance were observed in Experiment 2.

4 Experiment 3: Lexical decision in L2

To further examine the top-down lexical processes, as well as the role of metalinguistic demands involved in speech perception, a follow-up lexical decision task was included with the same stimulus set as in Experiment 2. Participants listened to the English words and English-like non-words, which either conflicted with the Spanish *v+s+c* constraint (e.g., *strict*) or did not conflict (e.g., control: *issue*), and decided if the stimulus formed a word or non-word in English. Carlson (2018b) included a lexical decision task to examine the extent to which Spanish-English bilinguals perceived the illusory “e” vowel during *Spanish (L1)* processing. The current experiment builds on Carlson’s findings by investigating whether bilinguals are cross-linguistically influenced by the L1 Spanish “e” onset rule during *English (L2)* comprehension. Moreover, Experiment 3 also expands upon Parlato-Oliveira et al. (2010) by examining speech perception in a task where metalinguistic demands are higher than in Experiment 2 (vowel detection), due to participants’ need to search the lexicon for a match. If bilinguals were influenced by the Spanish vowel rule with Spanish-conflicting words (e.g., *strict*) and non-words (e.g., *spelg*) in English, then they would respond differently with respect to response times and accuracy rates to conflicting stimuli, relative to non-conflicting (e.g., *issue*) stimuli. This prediction is in line with Freeman et al. (2016) in which Spanish-conflicting stimuli slowed bilinguals’ responses, relative to non-conflicting stimuli.

4.1 Experiment 3 methods

4.1.1 Participants. The same participants were tested in Experiment 3 as in Experiments 1 and 2.

4.1.2 Materials. The materials from Experiment 2 were also employed in Experiment 3, adapted for use in a lexical decision task (see Appendix 2 for stimuli).

4.1.3 Procedure. See Experiment 2 (vowel detection) for programming, response collection and stimulus recording procedures.

The lexical decision task was administered after the AX word discrimination and vowel detection tasks, and before the remaining cognitive and linguistic measures. Participants were instructed in English to pay attention to the English word or English-like non-word they heard. Participants were then asked to respond to the stimulus they heard as quickly and as accurately as possible, indicating whether they had heard a real English word or a non-word. After the instructions and 12 practice trials, participants performed the experimental task in which they heard a word (*s+c* onset and control onset) or non-word (*s+c* onset and control onset). During presentation of the stimulus, participants viewed a central fixation crosshair on the computer screen for the duration of the stimulus, immediately followed by a prompt on the visual display, asking if what they heard was a word or a non-word, see Figure 5. The crosshair and preceding prompt were presented in the center of a white screen in black, size 16 font, Courier, and the left/right shift keys represented word/non-word responses. Reaction times were recorded from the onset of the auditory stimulus and presentation of the prompt lasted until the participant made a response.

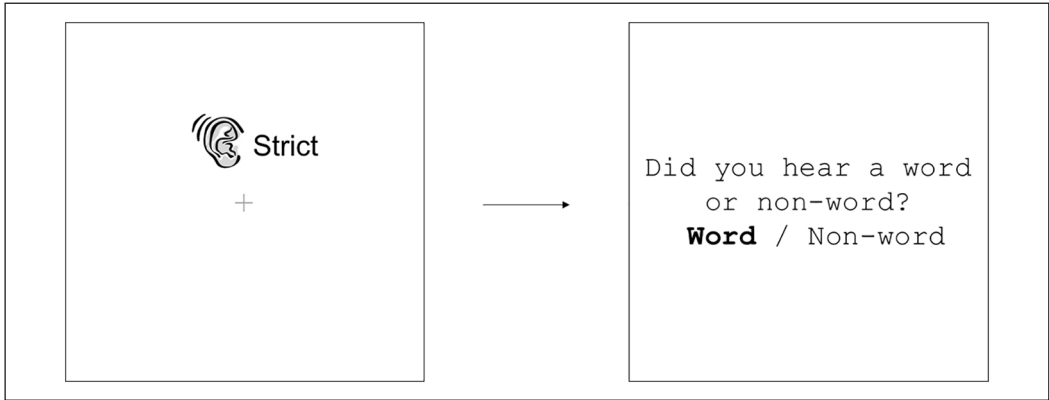


Figure 5. Sample trail from the lexical decision task in Experiment 3. In this example, participants heard *strict* and decide if it was a word or non-word (Word = real English word, Non-word = non-word in English).

The total time to complete this task was approximately 10 minutes. Participants were then debriefed about the study and compensated. The total study duration (Experiments 1, 2, 3 and cognitive and linguistic measures) was approximately two hours.

4.1.4 Coding and analyses. To examine perception of irrelevant-language phonotactic constraints in the lexical decision task, mixed effects models were employed for accuracy rates and reaction times, analyzing words and non-words in separate models. The same coding and analysis procedures were used in Experiment 3 as in Experiment 2, resulting in 10% of the data being omitted. The word and non-word models included the following factors: Onset Type (s+c and control; centered), Language Group (monolingual and bilingual; centered); Items, and Participants. Reaction times were log transformed for the reaction time models. The four models (two accuracy and two reaction time models) included fixed effects of Onset Type by Language Group, random intercepts with a random slope of Onset Type by Participants, and by-item random intercepts. The accuracy model for words failed to converge due to ceiling effects; therefore, we report the accuracy model for non-words, along with raw accuracy percentages across conditions and language groups.

4.2 Experiment 3 results

4.2.1 Accuracy on the lexical decision task. See Table 4 for accuracy data.

4.2.1.1 Non-words. There was a main effect of Language Group, $\beta = -1.26$, $SE = 0.25$, $z = -5.01$, $p < 0.01$, with bilinguals ($M = 85.73\%$, $SE = 1.55$) responding less accurately to non-words than monolinguals ($M = 95.44\%$, $SE = 1.58$). There was an interaction between Onset Type and Language Group, $\beta = 0.65$, $SE = 0.28$, $z = 2.36$, $p = 0.02$. In follow-up models within each Language Group, there was a main effect of Onset Type for bilinguals, $\beta = 0.54$, $SE = 0.25$, $z = 2.21$, $p = 0.03$, while no further effects were observed within the monolingual model ($p > 0.1$). Note: the main effect of Onset Type within the bilingual model should be interpreted with caution: bilinguals were less accurate to respond to s+c non-words ($M = 83.09\%$, $SE = 2.82$) than to control non-words ($M = 86.47\%$, $SE = 2.24$), a difference of only ~3%.

Table 4. Raw accuracy mean percentages and (Standard Error) across language groups and conditions on Experiment 3, lexical decision.

Accuracy	S+C Word	Control Word	S+C Non-word	Control Non-word
Bilinguals	91.49 (2.56)	93.48 (1.72)	83.09 (2.82)	86.47 (2.24)
Monolinguals	94.20 (1.25)	97.70 (0.35)	95.47 (1.04)	94.98 (0.56)

4.2.2 Reaction time effects on the lexical decision task

4.2.2.1 Words. Similar to the reaction time results on the vowel detection task (Experiment 2), there was a main effect of Language Group, $\beta = 0.16$, $SE = 0.04$, $t = 4.00$, $p < 0.01$, with bilinguals ($M = 1131.89\text{ms}$, $SE = 30.64$) responding more slowly than monolinguals ($M = 958.47$, $SE = 16.69$). There was a main effect of Onset Type, $\beta = -0.06$, $SE = 0.02$, $t = 2.90$, $p = 0.04$, with participants responding slower to s+c words ($M = 1063.86\text{ms}$, $SE = 28.34$) relative to control words ($M = 1026.50$, $SE = 27.02$). There was also an interaction between Onset Type and Language Group, $\beta = -0.06$, $SE = 0.03$, $t = -2.00$, $p = 0.043$.

Follow-up models revealed that the main effect of Onset Type and interaction between Onset Type and Language Group was driven by bilinguals who showed longer response times to s+c words ($M = 1166.14\text{ms}$, $SE = 44.34$) than to control words ($M = 1097.65\text{ms}$, $SE = 42.03$), $\beta = -0.07$, $SE = 0.03$, $t = -2.58$, $p = 0.01$. Monolinguals did not show a difference in response times to s+c words ($M = 961.58\text{ms}$, $SE = 17.87$) and control words ($M = 955.35\text{ms}$, $SE = 25.33$), $\beta = -0.02$, $SE = 0.03$, $t = -0.89$, $p = 0.37$. Overall, bilinguals were slower to respond to Spanish-conflicting words (e.g., *strict*) than to control words (e.g., *can*) suggesting interference from the L1 due to violation of the L1 Spanish constraint, while monolinguals did not show this pattern of slowing.

4.2.2.2 Non-words. The same main effects and a marginal interaction for words were observed for non-words. There was a main effect of Language Group, $\beta = 0.15$, $SE = 0.04$, $t = 3.70$, $p < 0.01$, with bilinguals ($M = 1190.11\text{ms}$, $SE = 30.50$) responding more slowly than monolinguals ($M = 973.22\text{ms}$, $SE = 22.42$). There was a main effect of Onset Type, $\beta = 0.15$, $SE = 0.04$, $t = -2.57$, $p = 0.01$, with participants responding more slowly to s+c non-words ($M = 1111.13\text{ms}$, $SE = 30.78$) relative to control non-words ($M = 1052.20\text{ms}$, $SE = 31.14$); and a marginal interaction between onset and Language Group, $\beta = -0.07$, $SE = 0.04$, $t = -1.91$, $p = 0.06$.

Once again, these slowing effects were led by the bilingual group, with follow-up models revealing that bilinguals showed longer response times to s+c non-words ($M = 1242.61\text{ms}$, $SE = 36.04$) than to control non-words ($M = 1137.61\text{ms}$, $SE = 38.32$), $\beta = -0.10$, $SE = 0.02$, $t = -4.16$, $p < 0.01$. Monolinguals did not show a difference in response times to s+c non-words ($M = 979.65\text{ms}$, $SE = 23.16$) and control non-words ($M = 966.79\text{ms}$, $SE = 34.57$), $\beta = -0.03$, $SE = 0.04$, $t = -0.73$, $p = 0.47$. Response times across Spanish-conflicting and control conditions (words and non-words) indicate that bilinguals experienced increased L1 interference from Spanish-conflicting stimuli than with control stimuli. See Figure 6 for difference scores across conditions and Language Group.

Overall, results suggest that bilinguals are slower to respond to L2 s+c words and non-words that are phonotactically illicit in the L1. This task was more difficult than vowel detection, as bilinguals did not perform near ceiling with accuracy rates. Moreover, the lexical decision required a lexical search, and for bilinguals, across two languages. Increased task and metalinguistic demands, and explicit access to the lexicon enhanced the word/non-word effect, relative to vowel detection.

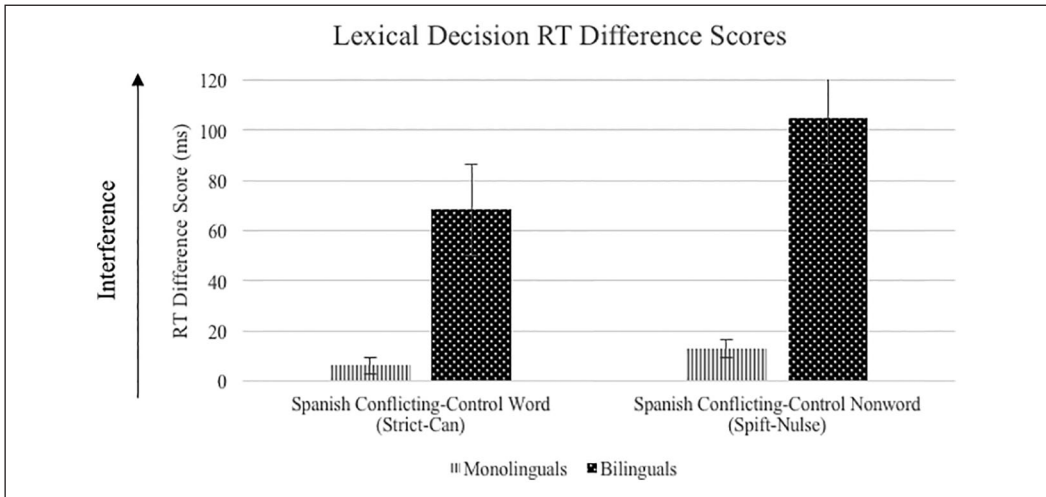


Figure 6. Monolingual and bilingual RT difference scores for Spanish-conflicting minus control trails across word and non-word condition on the lexical detection task. Error bar represent 1 standard error. Bilinguals demonstrated greater interference from Spanish-conflicting words and non-words than monolinguals. Within bilinguals, Spanish-conflicting words and non-words resulted in increased interference relative to controls.

4.3 Experiment 3 discussion

A lexical decision task was employed to further examine how top-down lexical processes affect L2 perceptual processing. L1 mapping was observed for L1-conflicting L2 words and non-words in lexical decision, whereas L1 mapping was observed only for L1-conflicting L2 words in vowel detection (Experiment 2). Top-down phonotactic and lexical knowledge influenced bilinguals' perception; however, the explicit step (metalinguistic demand) of deciding on stimulus lexicality in the L2 for bilinguals further revealed perceptual effects, over and above vowel detection. The slower response times for s+c items suggest that the L2 input is being mapped onto L1-like phonological representations, in this case, the "e" onset (i.e., word: *strict* and *estric*, non-word: *spelg* and *espelg*), and that Spanish-English bilinguals generally have difficulty resolving the illicit s+c sequence. Also, given the slowing in reaction times that was observed for bilinguals and not monolinguals across s+c words and non-words, relative to controls, it is clear that stimulus frequency throughout the experiment did not affect the results (i.e., fewer s+c words and non-words than control words and non-words). There was an even greater slowing effect for non-words, not only because representations were competing, but also because non-words were more difficult to reject in the L2 (e.g., Dijkstra et al., 1999), as bilinguals were searching across languages to verify if there was a lexical match (e.g., *spelg* and *espelg*). As in Experiments 1 and 2, no effects of age of acquisition, proficiency, or language dominance were observed in Experiment 3.

5 General discussion

5.1 Summary of findings

We examined the extent to which bilinguals employed a native-language (L1) filter when processing second language (L2) input. To do so, we exploited the L1 Spanish

vowel+s+consonant cluster (v+s+c) constraint (e.g., *estricto*) by aurally presenting English monolinguals and Spanish-English bilinguals with English words (e.g., *strict*; Experiment 1, AX word discrimination; Experiment 2, vowel detection; and Experiment 3, lexical decision) and non-words (e.g., *spelg*; Experiment 2, vowel detection; Experiment 3, lexical decision) that conflicted with the Spanish constraint. In AX word discrimination, cross-linguistic influence from the L1 v+s+c phonotactic constraint during L2 processing was not observed, likely due to the task's relatively low metalinguistic demands (see Parlato-Oliveira et al., 2010 for further support of this argument). Mapping L2 input onto L1 acoustic categories may be inherent to the increased metalinguistic demands of vowel detection and lexical decision, as knowledge of a vowel or a lexical search was required. In vowel detection, an explicit measure of vowel perception, Spanish-English bilinguals were influenced by the L1 v+s+c rule, an "e" onset, when listening to L1-conflicting L2 words (e.g., *strict*). In lexical decision, cross-linguistic influences were more robust in that L1 phonotactic constraints affected perceptual mapping with L2 s+c words *and* non-words. The results extend previous findings and support the initial hypothesis that bilinguals employ an L1 filter when listening to L2 sound sequences, exploiting L1 phonotactic constraints. However, the L1 perceptual effects on bilinguals' L2 processing were influenced by task difficulty and stimulus lexicality.

5.2 Accounting for differential perceptual outcomes across experiments and previous studies

The current investigation 1) used real words, 2) included three tasks that varied in metalinguistic demands, and 3) tested cross-linguistic influence of L1 perception during L2 auditory input. The current study extends previous investigations examining perceptual processing in bilinguals by using real words. Previous studies have used non-words (Carlson et al., 2016) and nonsense sound sequences (Weber & Cutler, 2006), which limit real-world generalizability. Stimulus lexicality resulted in differential perceptual effects. During vowel detection, with words, bilinguals recruited top-down perceptual knowledge of phonotactic constraints, as well as top-down lexical knowledge. It thus appears that when the onset sound of a stimulus is the explicit focus, a stronger perceptual representation exists for words than for non-words. The contribution of top-down phonotactic and lexical information has previously been examined by Dupoux et al. (2001) who found that perceptual repair through vowel epenthesis ("u" in illicit CVCCV clusters) in Japanese monolinguals occurred through top-down phonotactic and not lexical influences. Monolinguals repaired conflicting input into Japanese-like non-words with the preferential "u" (e.g., *mikdo* repaired to *mikudo*, non-word). In the current study, top-down phonotactic and lexical influences affected perception, likely not through perceptual repair (see reasons below), but through an L1 filter resulting in parallel processing of L1- illicit sound sequences during L2 processing.

The tasks in the present investigation differed from those in Dupoux et al. (2001) in terms of demands on metalinguistic awareness. As metalinguistic demands increased, so did the perceptual effects. Lexical decision is a more difficult task, especially for bilinguals, as individuals search the lexicon for a match (Soares & Grosjean, 1984). Therefore, perceptual effects were more apparent in lexical decision, as evidenced by cross-linguistic processing of Spanish-conflicting words and non-words. It appears that bilinguals first resolve which phonemes are present in the string of letters they hear in order to make a lexical decision. This process is faster for English control stimuli since bilinguals do not experience cross-linguistic conflict with Spanish constraints.

The motivation for the current investigation was to uncover the cross-linguistic influences involved during speech perception. Previous investigations have examined perceptual processing violations across languages, but not across tasks that vary in metalinguistic demands (Lentz & Kager, 2015; Weber & Cutler, 2006). These investigations have used reaction times to measure the presence of the L1 filter during L2 processing, while other studies have used accuracy rates, or “false alarms.” Accuracy directly indicates perceptual repair, for example, whether the participant hears a vowel (e.g., Carlson et al., 2016; Cuetos et al., 2011). The Spanish-dominant bilinguals in Carlson et al. (2016) reported hearing an initial “e” 22% of the time in s+c non-words and the monolinguals in Cuetos et al. (2011) reported 56% during within-language perceptual repair. In the present study, it was less likely that L1 Spanish speakers would employ perceptual repair, reporting explicitly hearing a vowel (2–4% of the time), as they were tested in their L2. Reaction times were a more sensitive measure of cross-linguistic influences, such as parallel processing during L2 perception (Lentz & Kager, 2015; Weber & Cutler, 2006).

To account for the discrepancy between reaction times and accuracy rates, bilinguals map Spanish-conflicting English stimuli onto existing Spanish phonological representations, which result in differences with response rates relative to non-conflicting stimuli. Bilinguals did not appear to explicitly “fix” the conflicting input to align with L1 rules, as bilinguals had access to the unrepaired representation in their accurate response patterns, likely due to their knowledge of and experience with English. Therefore, perceptual repair might not have been the mechanism responsible for the observed results. Carlson (2018a) suggested that the representation of the s+c onsets was inherently different between Spanish monolinguals and Spanish-English bilinguals, since Spanish monolinguals have not been exposed to s+c, while Spanish-English bilinguals have been exposed to permissible s+c onsets through their L2 knowledge and experience. However, the results obtained here, in Lentz and Kager (2015), and in Weber and Cutler (2006) support that bilinguals map L2 input using an L1 filter.

5.3 Empirical and theoretical implications

Overall, when hearing English words, L1 Spanish speakers were cross-linguistically influenced by Spanish constraints, especially when the task’s metalinguistic demands were high. The current investigation contributes to a framework for further studying the mechanisms involved in speech perception in bilinguals. Combined with the literature on parallel language activation (e.g., Marian & Spivey, 2003), the mechanism that accounts for the cross-linguistic effects observed in the present investigation is likely parallel processing, whereby the bilinguals’ two languages were online during auditory perception. Moreover, during speech perception, monolinguals and bilinguals may process auditory input in a top-down way. For example, when an individual hears an L2 word or a sound sequence that conflicts with L1 constraints in a processing context where metalinguistic demands are high (e.g., *strict* /*strikt*/), top-down processes are likely initiated where activation of the non-conflicting representation comes online (*estricto*, “strict”). Then, the constraint comes online as well (“e” onset constraint), and dictates how the input should be mapped. This top-down, rule-based approach is supported in previous investigations on perceptual processing (e.g., Carlson et al., 2016; Dupoux et al., 1999; Lentz & Kager, 2015; Parlato-Oliveira et al., 2010; Weber & Cutler, 2006) and within the Perceptual Assimilation Model (Best, 1994).

These findings have implications not only for processing L2 sounds and words that conflict with L1 phonotactic constraints, but also for L2 speech perception more broadly. We suggest, along with Lentz and Kager (2015), that bilinguals process L2 speech through an L1 phonotactic filter. We draw links between this idea and what has been described in previous research as “L1 transfer” (Gass & Selinker, 1994), “new wine in old bottles” (Flege, 1998), and “L1 assimilation” (Best,

1994). L2 speech sounds and sound sequences are represented in terms of L1 categories and patterns through the transfer or assimilation of these categories and patterns. In the current investigation, the subtler effects seen in response times, but not as much in the responses themselves, show that these transferred/assimilated categories have either evolved, or that more L2-like categories have been added, but nonetheless traces have been left.

Notably, the perceptual effects in the present investigation emerged with increased metalinguistic demands, when the focus of the task was on a specific aspect of the auditory stimulus (e.g., word onset or lexicality). Miyawaki et al. (1975) also found that perceptual effects emerge when the task's focus is on speech input. Miyawaki et al. used the /r/-/l/ phonemic categorical contrast and asked L1 English and L1 Japanese speakers if they perceived the difference in /ra/- /la/-onset syllables. In a "speech" task, L1 Japanese speakers discriminated equally well at or across the /r-/l/ category boundary as they did farther from the boundary, whereas English L1 speakers showed poor discrimination away from the boundary, and good discrimination at it. The differences across Japanese and English speakers reflect the fact that Japanese does not distinguish between /r/ and /l/, while English does. However, when participants performed the same categorical discrimination task with "nonspeech" stimuli that contained the /ra/-/la/ syllables synthesized to a chirp-like (glissando) sound, English and Japanese monolinguals performed similarly (i.e., Japanese participants discriminated these syllables as well). Crucially, the "speech" and "nonspeech" stimuli involved identical F3 contours. The difference was that in the "speech" condition, stimuli were combined with F1 and F2 contours, making them more speech-like. In the "nonspeech" condition, stimuli contained F3 contours played in isolation, which sounded like high chirps or glissandos. Therefore, the identical acoustic material could be discriminated well, but when it was construed to represent speech sounds, it was "filtered" through linguistic representations, allowing the structure of each listener's L1 to shape discrimination performance. Combined with the current investigation, these findings suggest that when metalinguistic demands are high, or when there is a specific focus on "speech" sounds within a task, perceptual effects are likely to emerge. The difference between Miyawaki et al. and the current study was that Miyawaki et al. used a task comparing two nearly identical, very brief sounds, whereas the AX word discrimination task contrasted two clearly different full words. The AX word discrimination task was not sufficient to tap into perceptual representations, as were the other tasks that focused on a specific aspect of the speech input (i.e., vowel detection and lexical decision).

5.4 Limitations and future directions

A limitation of the current investigation is potential residual task effects. In lexical decision, it is possible that participants previously focusing on initial vowels (vowel detection) in the presence of s+c stimuli might have triggered metalinguistic knowledge of the constraint. However, given that participants were not cued into the stimulus onset in the lexical decision task, but rather the whole stimulus for lexical search, it is likely that the results are due to differences in metalinguistic load specific to each task and not task order effects. Overall, the current findings support the literature on parallel language activation (e.g., Marian & Spivey, 2003). Results confirm that L1 influences during L2 processing are robust, even affecting speech perception, and suggest that top-down rules dictate how the input is mapped in certain perceptual contexts. Future work should further characterize and examine how parallel processing, along with bottom-up and top-down influences, operate in tandem as the mechanisms involved during speech perception in bilinguals.

6 Conclusion

While listening in their L2, bilinguals may process L2 input through an L1 filter, especially when metalinguistic task demands are higher. For native-Spanish speakers, English words such as *strict* may result in ambiguity for mapping onto either L1 or L2 phonological representations, since the Spanish L1 phonotactic constraint of adding an “e” to the onset of s+consonant cluster words is violated. The current investigation employed three tasks that differed in metalinguistic demands to examine perceptual processing in bilinguals. In the case of phonotactic constraints, perceptual parallel processing may be affected by top-down information and metalinguistic demands, and limited to tasks where knowledge of a vowel matters and/or lexical access is required. Parallel processing is influenced by task demands and lexical status. Findings suggest that bilinguals process their languages in parallel at the sub-lexical level. Importantly, these results shed light on when and where co-activation and perception may be present during bilingual language comprehension. We conclude that a bilingual’s two languages mutually affect each other, with the extent of the interaction influenced by task demands and the nature of linguistic processing.

Acknowledgements

We would like to thank members of the *Bilingualism and Psycholinguistics Research Group* at Northwestern University and the *Bilingualism & Cognition Lab* at San Diego State University.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD059858 to Viorica Marian. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ORCID iD

Max R. Freeman  <https://orcid.org/0000-0003-0730-1685>

Supplemental material

The supplemental files are available at <http://journals.sagepub.com/doi/suppl/10.1177/0023830920983368>

References

- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39(3), 445–459.
- 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.
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The Development of Speech Perception: The Transition From Speech Sounds to Spoken Words*, 167(224), 233–277.
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [Computer program].
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436.
- Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in stimulus–stimulus inhibition. *Bilingualism: Language and Cognition*, 17(3), 610–629.

- Carlson, M. T. (2018a). Making room for second language phonotactics: Effects of L2 learning and environment on first language speech perception. *Language and Speech*, 61(4), 598–614. doi: 10.1177/0023830918767208
- Carlson, M. T. (2018b). Now you hear it, now you don't: Malleable illusory vowel effects in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 22(5), 1101–1122. doi:10.1017/S136672891800086X
- Carlson, M. T., Goldrick, M., Blasingame, M., & Fink, A. (2016). Navigating conflicting phonotactic constraints in bilingual speech perception. *Bilingualism: Language and Cognition*, 19(5), 939–954. doi:10.1017/S1366728915000334
- Cuetos, F., Hallé, P. A., Dominguez, A., & Segui, J. (2011, August). Perception of prothetic /e/ in #sC utterances: Gating data. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 540–543).
- Darcy, I., Peperkamp, S., & Dupoux, E. (2007). Bilinguals play by the rules: Perceptual compensation for assimilation in late L2-learners. *Laboratory Phonology*, 9(9), 411–442.
- Darcy, I., & Thomas, T. (2019). When blue is a disyllabic word: Perceptual epenthesis in the mental lexicon of second language learners. *Bilingualism: Language and Cognition*, 22(5), 1141–1159. doi:10.1017/S1366728918001050
- Dijkstra, T., Grainger, J., & Van Heuven, W. J. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41(4), 496–518.
- Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. American Guidance Service.
- Dunn, L. M., Lugo, P., & Dunn, L. M. (1997). *Vocabulario en imágenes Peabody (TVIP)*. American Guidance Service.
- 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(6), 1568–1578.
- Dupoux, E., Pallier, C., Kakehi, Y., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes*, 16(5/6), 491–505
- Dupoux, E., Parlato, E., Frota, S., Hirose, Y., & Peperkamp, S. (2011). Where do illusory vowels come from? *Journal of Memory and Language*, 64(3), 199–210.
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: Stress deafness in simultaneous French-Spanish bilinguals. *Cognition*, 114(2), 266–275.
- Flege, J. E. (1998). The phonetic study of bilingualism. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies*, (35), 017–026.
- Freeman, M. R., Blumenfeld, H. K., & Marian, V. (2016). Phonotactic constraints are activated across languages in bilinguals. *Frontiers in Psychology*, 7(702). doi: 10.3389/fpsyg.2016.00702 PMID: PMC4870387
- Gass, S., & Selinker, L. (1994). *Second language acquisition: An introductory course*. Lawrence Erlbaum Associates.
- 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(1), 177–192.
- Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2019). The Language Experience and Proficiency Questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*, 1–6. doi: <https://doi.org/10.1017/S1366728919000038>
- Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3? Perception 36 ECVF Abstract Supplement*. Max Planck Institute for Biological Cybernetics.
- Lentz, T. O., & Kager, R. W. (2015). Categorical phonotactic knowledge filters second language input, but probabilistic phonotactic knowledge can still be acquired. *Language and Speech*, 58(3), 387–413.
- Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE*, 7(8), e43230. doi: 10.1371/journal.pone.0043230

- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research, 50*(4), 940–967.
- Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition, 6*(2), 97–115.
- Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*(5), 331–340.
- Parlato-Oliveira, E., Christophe, A., Hirose, Y., & Dupoux, E. (2010). Plasticity of illusory vowel perception in Brazilian-Japanese bilinguals. *The Journal of the Acoustical Society of America, 127*(6), 3738–3748. doi: <http://dx.doi.org/10.1121/1.3327792>
- Pelli, D. G. (1997). The Video Toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.
- PsychCorp. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. Harcourt Assessment.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. URL: <http://www.r-project.org>.
- Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition, 16*(2), 304–324.
- Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition, 12*(4), 380–386.
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*(3), 281–284.
- Strange, W. (1999). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research contains the contributions to a Workshop in Cross-Language Perception held at the University of South Florida, Tampa, Fla. Phonetica, 56*, 105–107.
- Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America, 119*(1), 597–607.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001, 2007). *Woodcock Johnson III Tests of Cognitive Abilities*. Riverside Publishing.
- Yavas, M., & Someillan, M. (2005). Patterns of acquisition of f/s/-clusters in Spanish-English bilinguals. *Journal of Multilingual Communication Disorders, 3*(1), 50–55.