Research Article

Code-Switching in Highly Proficient Spanish/English Bilingual Adults: Impact on Masked Word Recognition

Paula B. García,^{a,b} Lori Leibold,^b Emily Buss,^c Lauren Calandruccio,^d and Barbara Rodriguez^e

Purpose: The purpose of this study was to evaluate the impact of code-switching on Spanish/English bilingual listeners' speech recognition of English and Spanish words in the presence of competing speech-shaped noise.

Method: Participants were Spanish/English bilingual adults ($N = 27$) who were highly proficient in both languages. Target stimuli were English and Spanish words presented in speech-shaped noise at a -14 -dB signal-to-noise ratio. There were 4 target conditions: (a) English only, (b) Spanish only, (c) mixed English, and (d) mixed Spanish. In the mixed-English condition, 75% of the words were in English, whereas 25% of the words were in Spanish. The percentages were reversed in the mixed-Spanish condition.

Results: Accuracy was poorer for the majority (75%) and minority (25%) languages in both mixed-language conditions

compared with the corresponding single-language conditions. Results of a follow-up experiment suggest that this finding cannot be explained in terms of an increase in the number of possible response alternatives for each picture in the mixed-language condition relative to the single-language condition.

Conclusions: Results suggest a cost of language mixing on speech perception when bilingual listeners alternate between languages in noisy environments. In addition, the cost of code-switching on speech recognition in noise was similar for both languages in this group of highly proficient Spanish/English bilingual speakers. Differences in response-set size could not account for the poorer results in the mixed-language conditions.

The purpose of this study was to evaluate the impact of code-switching on highly proficient Spanish/English bilingual listeners' speech recognition of English and Spanish words in the presence of competing speech-shaped noise. Bilingual speakers often switch between their first language (L1) and second language (L2) during daily life conversations with other bilinguals; this has been referred to as *code-switching* (Barker, 1947; Cooper, 2013). Code-switching occurs when bilingual speakers switch languages within sentences, phrases, or between words. Code-switching requires cognitive flexibility and language control skills for speakers, listeners, or readers who are processing the mixed-language messages (van Hell & Witteman, 2009).

Code-switching also slows comprehension. This has been demonstrated with both behavioral measures (Dijkstra, 2005; Olson, 2017; Spivey & Marian, 1999) and neurophysiological measures (Abutalebi et al., 2007; Alvarez, Holcomb, & Grainger, 2003; Chauncey, Grainger, & Holcomb, 2008; Moreno, Federmeier, & Kutas, 2002). The majority of studies evaluating effects of code-switching on spoken language processing have presented speech in quiet. It is unclear, therefore, whether the processing costs associated with code-switching result in poorer masked speech recognition for mixed-language speech than for single-language speech. This is an important consideration because communication often occurs in noisy environments.

The cognitive and lexical selection control mechanisms associated with code-switching have been studied mostly by using picture-naming and number-naming paradigms. In these paradigms, the bilingual participant is instructed to name a picture/number presented on a screen on the basis of a cue that indicates which language to use (e.g., if the square is green, name the picture in English; if the square is red, name the picture in Spanish). The cost of alternating between languages has been measured by calculating the difference in latency for naming pictures or numbers in one

^aUniversidad de los Andes, Bogotá, Colombia

^bBoys Town National Research Hospital, Omaha, NE

^cThe University of North Carolina at Chapel Hill

^dCase Western Reserve University, Cleveland, OH

^eThe University of New Mexico, Albuquerque

Correspondence to Paula B. García: pb.garcia@uniandes.edu.co

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Steve Aiken

Received October 26, 2017

Revision received February 15, 2018

Accepted April 2, 2018

https://doi.org/10.1044/2018_JSLHR-H-17-0399

Disclosure: The authors have declared that no competing interests existed at the time of publication.

language versus two languages. On the other hand, the cost associated with switching languages between trials (switch cost) has been measured by calculating the naming latency difference between switch trials (when the participant is asked to name a picture using a different language from the previous trial) and nonswitch trials (when the language to name the object does not change in the next trial) within the same experimental condition.

In a seminal paper, Meuter and Allport (1999) reported code-switching cost asymmetries in a group of sequential bilinguals who acquired their L2 later in life and whose L1 remained their dominant language. The L1 and L2 languages for these bilinguals included English, French, German, Italian, Portuguese, and Spanish. Sixteen adult participants (ages 22–44 years) spoke English either as their L1 or L2. Subjects were required to verbally name single digits displayed in a background that was either blue or yellow; the color served as a cue indicating which language the listener was expected to use for their verbal response (L1 or L2). The color switched unpredictably from trial to trial, forcing the participant to switch between their L1 and L2. The probability of a language switch occurring on a given trial was 0.3. The naming latency of the responses in the switch trials was longer than in the nonswitch trials. The switching cost was asymmetric, with a greater cost observed when having to switch from their less proficient language (L2) to their dominant language (L1). The authors argued that spoken language in L2 required the suppression of the competing L1. To test the impact of relative language proficiency, participants were further divided into two groups ($n = 8$ per group) according to differences in reaction times (RTs) for naming numerals in their L1 versus in their L2. RTs were based on all nonswitch trials that were presented prior to the L1 switch. The mean RT in their L1 was then subtracted from the mean RT in their L2. Interestingly, the group of participants with the smallest naming RT differences between the L1 and the L2 (the more balanced bilinguals) did not show an asymmetry in switching cost, but the group with the largest naming RT differences (the bilinguals with larger differences in language proficiency between their L1 and L2) did. These results have been corroborated in subsequent behavioral studies involving naming (Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006) and reading (Filippi, Karaminis, & Thomas, 2014; Macizo, Bajo, & Paolieri, 2012).

Findings from electrophysiological studies support the idea that switching between languages also has a negative impact on word and sentence comprehension. Event-related potential (Alvarez et al., 2003; Chauncey et al., 2008) and functional magnetic resonance imaging (Abutalebi et al., 2007) methods have been implemented to study language-switching costs in comprehension tasks, such as go/no-go semantic categorization. In Alvarez et al. (2003), event-related potential responses of beginning learners of Spanish were recorded, whereas L1 or L2 words were presented in a mixed-language condition. Words were either preceded by the same word in the same language or by the translation of the word in the other language. On each trial,

participants were asked to indicate whether the presented word referred to a body part using a button press, regardless of the language of presentation. The results indicated that language switching modulated the amplitude of the N400 component, thought to reflect semantic violation, but only when the switch was from the L1 to the L2.

Most studies on code-switching have relied on visual stimuli; however, code-switching commonly occurs in verbal interactions, where the input is the auditory speech signal (Chauncey et al., 2008). A recent study by Olson (2017) examined code-switched auditory comprehension using an eye-tracking paradigm. The study included 25 Spanish/English–proficient bilingual listeners who heard target words in the context of sentences recorded by a highly proficient Spanish/English female bilingual speaker with no detectable foreign accent in either language. The sentences were presented auditorily with visual images of the target words in each of the two language modes. In the “monolingual mode,” most but not all of the words were drawn from a single language, for example, “El chico dijo que quiere ver **spiders** cuando anda en el bosque.” In the “bilingual mode,” words were drawn from both languages with similar frequency, for example, “Cuando era pequeña, mi hermana **loved spiders and other bugs**.” In these examples, English words are bolded, whereas Spanish words remain in plain text. In both presentation modes, there are two types of target words: stay (the target word is preceded by a speech produced in the same language) and switch (the target word is preceded by a speech produced in the other language). The study results showed a switch cost only in the monolingual mode and only when switching into the dominant language. The authors explained this finding within the context of the inhibitory framework (Green, 1998), suggesting that comprehension in the monolingual mode requires strong inhibition of the nontarget language, with stronger inhibition applied to the L1 than to the L2.

We are aware of only one behavioral study in which the effect of switching between languages on speech-in-noise perception was assessed. Piccini and Garellek (2014) tested eight Spanish/English bilingual adults. These adults lived in the United States, learned English before 2 years of age, and were heritage speakers of Mexican Spanish. All listeners reported English as their dominant language. The stimuli consisted of modified versions of the Bamford–Kowal–Bench (Bench, Kowal, & Bamford, 1979) sentences produced by a bilingual female speaker of American English and Mexican Spanish. The English sentences were translated and modified to include high-frequency words in Spanish. Listeners were presented with four lists of sentences in four conditions: (English only, Spanish only, code-switch English to Spanish, and code-switch Spanish to English). Code-switch location within a sentence (at specific keywords: nouns and verbs) was counterbalanced for a syntactic position, reflecting the grammatical rules of code-switching. Those sentences were presented in white noise at four signal-to-noise ratios (SNRs) –6, –3, +0, and +3 dB. Performance was better in the Spanish-only condition than in

the other three conditions, but the performance among the three remaining conditions did not differ; the English-only condition was not easier than the two code-switching conditions. The authors argued that the lack of code-switching cost was due to the stimuli used because, unlike words or pictures, the auditory speech signal contains prosodic cues that the bilingual listeners could use to rapidly identify the target language and reduce the costs associated with switching languages.

One consideration when evaluating effects of code-switching for masked speech recognition is the increased susceptibility to masking when listeners hear a speech in their nondominant language. Although understanding single-language speech in the presence of noise and/or reverberation is challenging for all listeners, it is particularly difficult for nonnative listeners (e.g., Bradlow & Alexander, 2007; Florentine, Buus, Scharf, & Canevet, 1984; Garcia Lecumberri, & Cooke, 2006). The detrimental effects of noise on nonnative speech recognition have been reported for recognition of consonants (Garcia Lecumberri & Cooke, 2006; Hazan & Simpson, 2000), words (Mendel & Widner, 2016), and sentences (Bradlow & Bent, 2002; Florentine et al., 1984; Mendel & Widner, 2016; van Wijngaarden, Bronkhorst, Houtgast, & Steeneken, 2004). The pronounced speech-in-noise difficulties experienced by nonnative listeners are evident even for individuals who learned the L2 early in life (Florentine et al., 1984; Garcia Lecumberri & Cooke, 2006; Mayo, Florentine, & Buus, 1997; Nábelek & Donahue, 1984; Rogers, Lister, Febo, Besing, & Abrams, 2006). These effects necessitate considering the listener's language history and dominance when evaluating masked speech recognition.

Purpose and Hypotheses

Bilingual speakers commonly code switch with other members of the same bilingual group (Myers-Scotton, 2006). There are some processing costs associated with changing languages when listening in quiet conditions (Abutalebi et al., 2007). However, quiet listening conditions are not typical; there is often background noise in natural communication environments, and it is unclear whether noise affects the cost of code-switching. Therefore, the purpose of this study was to determine whether highly proficient bilingual listeners are less accurate at recognizing words in noisy environments when they switch between their languages. We examined Spanish/English bilingual listeners' speech recognition of English and Spanish words in a background of speech-shaped noise. Although language dominance varied across listeners, all listeners were highly proficient speakers of both languages. It was hypothesized that masked speech recognition for highly proficient Spanish/English bilingual listeners is better when words are spoken in one language than when languages are mixed. Further, we predicted that the effect of code-switching would be more pronounced in noise than previously observed in quiet due to the degradation of acoustic/phonetic cues associated with masking—the greater processing load associated with switching languages could reduce the listener's ability to

recognize speech based on sparse cues. Based on previous findings showing that the switching cost is similar for both languages in proficient bilingual listeners (Costa et al., 2006; Costa & Santesteban, 2004; Meuter & Allport, 1999), we hypothesized that the cost of mixing the languages in the presence of noise is symmetrical for highly proficient Spanish/English bilingual adults. Our approach in the mixed-language conditions was to present words more frequently in the dominant language (75%) than in the non-dominant language (25%). We therefore predicted that word recognition would be poorer on trials associated with a switch from the majority language (75%) to the minority language (25%) than switches in the opposite direction.

Experiment 1: Spanish/English Word Recognition in Noise During Code-Switching

Method

Listeners

All listeners provided informed consent according to the institutional review board of Boys Town National Research Hospital regulations and were compensated for their participation in the study. Listeners were 26 Spanish/English bilingual adults with normal hearing (≤ 20 dB HL bilaterally at octave frequencies from 0.25 to 8 kHz). This sample, which included 21 women and six men, ranged in age from 19 to 40 years ($M = 28.7$, $SD = 7$). There were large individual differences in listeners' reports of when they learned English; the reported range was birth to age 28 years ($M = 9.3$, $SD = 7.3$). All listeners reported Spanish being the language spoken at home. Nineteen listeners reported being born in a Spanish-speaking country, whereas seven reported being born in the United States. See Table 1 for more detailed information on listeners' language backgrounds.

Listeners completed a demographic and language background questionnaire and the bilingual language profile (BLP; Birdsong, Gertken, & Amengual, 2012) before completing the experimental speech recognition testing. At the end of the second testing session, each listener completed the Versant Test (Pearson, 2011) of spoken language proficiency in both English and Spanish.

Self-Reported English and Spanish Proficiency From the BLP

Listeners provided information about their language proficiency, language learning experience, age of first exposure to the language, length of residency in the United States, and language used on a daily basis by filling out the BLP. Self-reported language proficiency appears to be strongly correlated with objective measures of proficiency (e.g., Marian, Blumenfeld, & Kaushanskaya, 2007), and together with reports of language dominance, it has been found to be predictive of English word recognition in the presence of noise for bilingual adults (Shi, 2015).

Table 1. Listeners' demographic, language experience background, and proficiency scores.

ID	Age (years)	Sex	Born	Home language	AoA (years)	LOR (years)	Level of education	Self-reported proficiency			Versant Test scores	
								English	Spanish	Dominance	English	Spanish
BIL003	27	F	United States	Spa-Eng	0	27	PhD	22	17	69	80	62
BIL004	36	F	United States	Spa-Eng	5	36	HS	18	22	16	55	69
BIL005	37	F	Peru	Spanish	5	1	BA/BS	20	24	-64	62	80
BIL006	21	M	Puerto Rico	Spanish	4	10	BA/BS	21	22	-47	80	80
BIL007	22	F	Venezuela	Spanish	0	0.3	BA/BS	20	24	-52	76	80
BIL008	23	M	Puerto Rico	Spanish	3	6	MA/MS	24	24	-54	80	80
BIL010	21	F	Venezuela	Spanish	14	7	BA/BS	22	24	-36	69	80
BIL011	23	F	United States	Spanish	5	23	BA/BS	24	15	75	80	65
BIL012	33	F	Mexico	Spanish	6	21	BA/BS	24	22	59	80	80
BIL013	36	F	Colombia	Spanish	28	8	BA/BS	20	24	-69	68	80
BIL014	38	M	Venezuela	Spanish	12	1	BA/BS	19	24	-63	65	80
BIL015	38	F	Venezuela	Spanish	12	18	MA/MS	24	24	-20	75	80
BIL017	34	F	Dominican Republic	Spanish	13	21	BA/BS	18	20	-33	61	80
BIL018	33	F	United States	Spanish	3	33	BA/BS	24	23	9	79	80
BIL019	23	F	United States	Spanish	5	23	MA/MS	21	19	6	68	76
BIL020	26	F	United States	Spanish	5	26	BA/BS	24	24	-13	79	79
BIL021	37	F	Mexico	Spanish	10	27	BA/BS	20	20	14	72	80
BIL022	26	M	United States	Spanish	5	26	MA/MS	24	20	54	80	80
BIL023	19	F	Dominican Republic	Spanish	4	1.5	BA/BS	23	23	-53	77	80
BIL028	35	F	Mexico	Spanish	21	16	BA/BS	21	24	-31	69	80
BIL029	19	F	Ecuador	Spanish	9	0.6	BA/BS	20	24	-87	76	79
BIL030	19	F	Dominican Republic	Spanish	6	1	BA/BS	20	24	-134	62	80
BIL038	40	F	Mexico	Spanish	25	16	BA/BS	20	24	-71	66	80
BIL043	25	F	Mexico	Spanish	15	10	HS	16	24	-99	55	
BIL044	29	M	Mexico	Spanish	21	8	HS	8	24	-139		
BIL045	26	F	Peru	Spanish	12	14	BA/BS	24	24	-28	61	78
BIL047	23	F	Colombia	Spanish	17	6	BA/BS	20	20	-39	64	69

Note. AoA = age of acquisition of English; LOR = length of residency in the United States; F = female; Spa-Eng = Spanish-English; PhD = Doctor of Philosophy; HS = high school; BA/BS = Bachelor of Arts or Bachelor of Science; M = male; MA/MS = Master of Arts or Master of Science.

The BLP consists of four modules: language history, language use, language proficiency, and language attitudes. Each listener completed each module. Within each module are questions about their history, use, proficiency, and attitude with respect to both English and Spanish. The range, mean, and standard deviation of scores obtained in each language appear in Table 2. Larger numbers indicate more language exposure in the listener's personal history, greater current language use, higher language proficiency, and a more positive attitude about language competence. On all four modules, mean scores were higher for Spanish than for English, but this was not the case for all individual participants. A summary statistic was computed for each participant and each language by taking a weighted average of the scores in the four modules; these summary statistics could theoretically range from 0 to 218. Language dominance was calculated by subtracting the total score for Spanish from the total score for English. A positive score indicates English dominance, a score near zero suggests balanced bilingualism, and a negative score indicates Spanish dominance. Dominance scores for individual listeners appear in Table 1. Four of the 26 participants obtained dominance scores greater than 20, indicating English dominance. Three out of 26 participants obtained dominance scores near zeros (≤ 20 and ≥ -20), indicating balanced bilingualism.

Twenty of the 26 participants obtained scores less than -20, indicating Spanish dominance.

English and Spanish Proficiency From the Versant Test

The Versant Test is an automated assessment of spoken language proficiency conducted over the phone. The test assesses sentence mastery, vocabulary, fluency, and pronunciation. In addition to reporting scores for each of the assessed areas, the Versant Test provides an overall score that ranges from 20 to 80. Higher scores indicate greater mastery of the language. The English Versant scores for listeners in this study ranged from 55 to 80 ($M = 71.15$, $SD = 8.44$). The Versant scores in Spanish ranged from 62 to 80 ($M = 77.64$, $SD = 5.24$). According to the Versant's qualitative assessment based on the scores, the lowest scores obtained both in English (55) and in Spanish (62) indicate that these listeners are independent users of the languages. They can understand standard spoken language, live or broadcast, on both familiar and unfamiliar topics normally encountered in personal, social, academic, or vocational life. The highest scores obtained (80) indicate that listeners speak and understand effortlessly at native speaker speeds and can contribute readily to a native-paced

Table 2. Summary statistics for each of the four bilingual language profile modules, shown separately for English and Spanish.

Module	Range of possible scores	Scores for English	Scores for Spanish
Language history	0–120	Range = 13–120 <i>M</i> = 61.6 <i>SD</i> = 33.2	Range = 65–120 <i>M</i> = 95.3 <i>SD</i> = 17.1
Language use	0–50	Range = 11–46 <i>M</i> = 27.0 <i>SD</i> = 10.2	Range = 4–48 <i>M</i> = 29.7 <i>SD</i> = 11.3
Language proficiency	0–24	Range = 8–24 <i>M</i> = 20.6 <i>SD</i> = 3.4	Range = 15–24 <i>M</i> = 22.3 <i>SD</i> = 2.5
Language attitude	0–24	Range = 0–23 <i>M</i> = 15.8 <i>SD</i> = 5.6	Range = 12–24 <i>M</i> = 21.0 <i>SD</i> = 3.0

discussion at length, maintaining the colloquial flow. Speech is completely fluent and intelligible.

Summary of Language Proficiency Measures

Summarizing, in a subjective measure (BLP) of proficiency, all listeners indicated high levels of self-reported proficiency for both languages. Similarly, high Versant Test scores for both languages in all participants objectively indicate high levels of proficiency in both languages. The BLP scores also indicated a trend for more listeners to be Spanish dominant than to be English dominant.

Stimuli and Experimental Conditions

The auditory stimuli and pictures used in this study were originally developed by Calandruccio, Gomez, Buss, and Leibold (2014) as part of a closed-set picture identification task to assess masked speech perception in monolingual and bilingual children in a laboratory setting. Targets were 30 disyllabic words selected so that the English and Spanish versions of the test share similar acoustic, phonetic, and linguistic characteristics (see the list of words and all their psycholinguistic statistics in Calandruccio et al., 2014). In general, the average word frequency was 1.36 (*SD* = 0.71) for English and 1.30 (*SD* = 0.63) for Spanish. These words were recorded in both English and Spanish by a simultaneous bilingual female talker who grew up in a Spanish/English bilingual household; the talker's mother spoke Venezuelan Spanish, whereas her father spoke American English. For the production of these tokens, no detectable dialectal accent was observed based on native Spanish speakers' subjective ratings.¹ The average duration of the words was 0.55 s (range = 0.34–0.80 s). All tokens were root-mean-square equalized using MATLAB. The competing masker for all conditions was the speech-shaped noise, which was based on a two-talker masker in English and Spanish (see Calandruccio et al., 2014, for a detailed description of the generation of this noise). The noise was

presented continuously throughout testing. All stimuli were recorded using a 44.1-kHz sampling rate. The selection and presentation of stimuli were controlled by a custom MATLAB script. Recordings were played through a 24-bit digital-to-analog converter (Avid, Fast Track Solo) and delivered diotically via headphones (Sennheiser HD-25).

Listeners were tested in each of four target word conditions: (a) English only, (b) Spanish only, (c) mixed English, and (d) mixed Spanish. In the English-only and Spanish-only conditions, 100% of the words were presented in English or in Spanish, respectively. In the mixed-English condition, 75% of the trials were English words and 25% were Spanish words. Similarly, in the mixed-Spanish condition, 75% of the trials were Spanish words and 25% were English words. The sequence of target languages in each mixed-conditions block was determined iteratively. An array, composed of variables representing 25% minority and 75% majority target language, was randomly shuffled until two criteria were met: (a) The first trial used the majority language, and (b) each instance of the minority language was preceded and followed by a trial in the majority language.

Procedure

The task was a four-alternative forced choice (4AFC). Listeners sat in front of a touchscreen monitor in a sound-attenuating booth. Four words were selected prior to each trial, one target and three foils. These selections were random with the restriction that words could not appear in sequential trials. Pictures corresponding to the four words appeared on the monitor prior to the presentation of the target word, one in each quadrant. The listener was instructed to touch the image on the monitor that represented the target word they heard.

All listeners completed a stimulus familiarization phase prior to experimental testing, consisting of a 60-trial block of either English-only or Spanish-only target words in the noise masker. The language of the target words for the familiarization block was randomly selected for each listener. The level of both the target and the masker was fixed at 65 dB SPL during the familiarization phase (0 dB

¹Native Spanish listeners who rated the production of the stimulus tokens are speakers of Mexican, Colombian, and Peruvian Spanish.

SNR). All listeners completed the familiarization block with only one or fewer incorrect responses (range = 97%–100% correct responses). During testing, the level of the masker was fixed at 65 dB SPL, and target words were presented at 51 dB SPL for all conditions. The fixed SNR of –14 dB was selected because it was expected to result in 60%–80% correct performance for most listeners in all conditions, based on the results of extensive pilot testing. Although –14 dB SNR is a difficult SNR relative to most speech recognition testing conducted in the laboratory (especially for open-set testing), it is one that is encountered in everyday environments (Flamme et al., 2012; Hodgson, Steininger, & Razavi, 2007; Rusnock & McCauley, 2012). Listeners completed two blocks of 60 trials for the English-only and Spanish-only conditions and three blocks of 60 trials for the mixed-English and mixed-Spanish conditions.

Results

Overall Performance

Listeners' percent correct word recognition scores in each language were computed for each of the four conditions. For the English-only and Spanish-only conditions, scores were based on the responses for all 120 trials (two blocks $\times$ 60 words per block). For the mixed-English and mixed-Spanish conditions, scores were computed separately for the minority language (i.e., percent correct recognition of words in the language presented 25% of the time) and the majority language (i.e., percent correct recognition of words in the language presented 75% of the time). Scores for the majority language were, therefore, based on 135 trials, and those in the minority language were based on 45 trials. The scores were converted into rationalized arcsine units prior to statistical analysis to normalize the variance at the extremes (Studebaker, 1985).

Figure 1 shows the distribution of English and Spanish word recognition scores in the speech-shaped noise for each of the target word conditions.

For the single-language conditions, the distributions of the scores represented in the box plots show that word recognition scores were higher in English ($M = 72.0\%$, $SD = 9.38$) than in Spanish ($M = 63.2\%$, $SD = 8.05$). Although scores in the mixed-language conditions showed a similar trend for English scores to exceed Spanish scores, the performance was poorer in the mixed-language condition than in the corresponding single-language condition. The pattern of results was observed for both minority English words ($M = 64.4\%$, $SD = 11.11$) and majority English words ($M = 67.7\%$, $SD = 11.71$), as well as minority Spanish words ($M = 56.2\%$, $SD = 9.42$) and majority Spanish words ($M = 58.5\%$, $SD = 8.12$).

A repeated-measures analysis of variance was performed on rationalized arcsine unit scores to test the hypothesis that masked word recognition is better for proficient English–Spanish bilingual adults when listening to a single language than when code-switching is required. This analysis included the within-subject factors of language (English,

Figure 1. Distribution of English and Spanish word recognition scores in speech-shaped noise for each experimental condition. Open boxes represent the performance for English targets, and grey boxes represent the performance for Spanish targets. The boxes represent the interquartile range (25th–75th percentile); the median score is indicated by the horizontal line dividing the boxes. Whiskers represent the 10th and 90th percentiles, and asterisks represent the maximum and minimum scores. In the mixed-language conditions, wider boxes represent the majority language (75%), while the narrower boxes represent the minority language (25%).

Spanish) and proportion of trials in the dominant language (100%, 75%, 25%). The analysis of variance yielded a significant main effect of language, $F(1, 25) = 67.22$, $p < .001$, $\eta_p^2 = .73$, indicating better performance with English than with Spanish target words. The main effect of the code-switch was also significant, $F(2, 50) = 19.65$, $p < .001$, $\eta_p^2 = .44$. The Language $\times$ Proportion of Dominant Language interaction was not statistically significant, $F(2, 50) = .15$, $p = .86$; $\eta_p^2 < .01$. Thus, while listeners performed more poorly overall in Spanish than in English, the effect of code-switching was similar for both languages. Paired comparisons (with Bonferroni adjustments, using a criterion of $\alpha = .017$) indicated significantly better performance for the single-language conditions compared with either the minority (25%) or the majority (75%) of language performance in the code-switch conditions ($p < .01$). No significant difference was observed between the minority and the majority of language performance for the code-switch conditions ($p = .14$). This result could indicate that the effect of code-switching was not solely due to unexpected input in a different language but to the requirement to change languages in the task.

Figure 2 presents data for individual listeners. Consistent with the group data, most listeners showed poorer word recognition performance in noise for the mixed-language condition compared with the corresponding single-language condition in both the majority and minority languages. The mean effect in English was 7.6 ($SD = 9.08$)

Figure 2. Differences between single-language and mixed-language scores are shown with values for individuals on top of the box plots. Open boxes represent the performance for English targets, and grey boxes represent the performance for Spanish targets. The boxes represent the interquartile range (25th–75th percentile), the median score is indicated by the horizontal line dividing the boxes, whiskers represent the 10th and 90th percentiles, and open circles represent individual scores. In the mixed-language conditions, wider boxes represent the language of the majority of the words (75%), whereas narrower boxes represent the language of the minority of words (25%).

and 4.3 ($SD = 7.41$) percentage points when English was the minority and majority languages, respectively. For Spanish, those values were 6.9 ($SD = 7.26$) and 4.7 ($SD = 8.20$) percentage points.

Discussion

This study evaluated word recognition accuracy in highly proficient Spanish/English bilingual adults while switching between their languages in a background of speech-shaped noise. Listeners completed a 4AFC task in four target conditions: (a) English only, (b) Spanish only, (c) mixed English, and (d) mixed Spanish. The results showed better masked word recognition in the English-only condition compared with the Spanish-only condition. Performance was poorer in both languages in the mixed-language conditions relative to the corresponding single-language conditions. While the influence of code-switching on masked speech recognition has not been well studied, understanding speech in noise is challenging for nonnative listeners even when code-switching is not required and they are listening to an L2 learned early in life (Florentine et al., 1984; García Lecumberri & Cooke, 2006; Mayo, Florentine, & Buus, 1997; Nábelek & Donahue, 1984; Rogers et al., 2006). These results suggest that switching between languages in the presence of noise reduces word recognition in highly proficient bilingual adults for both of their languages.

Masked Speech Recognition for English-Only and Spanish-Only Conditions

The better performance in the English-only condition than in the Spanish-only condition was unpredicted considering that all the listeners in the study had been exposed to Spanish since birth and that 19 of the listeners indicated Spanish language dominance based on the BLP. These results are also in contrast to the results of Weiss and Dempsey (2008), who tested Spanish–English bilingual adults using the Hearing in Noise Test sentences and found better performance in the L1 compared with the L2. Piccini and Garellek (2014) examined the perception of single-language and code-switched sentences in white noise. They reported that eight Mexican American heritage listeners performed better in Spanish-only sentences compared with English only, although their performance did not differ in the recognition of code-switched sentences compared with the single-language sentences.

It is not clear why listeners in this study performed better in English than in Spanish. Data reported by Hochmuth, Birger, Brand, and Jurgens (2015) are consistent with the idea that languages differ with respect to masked recognition. The authors examined speech reception thresholds in different noise maskers for listeners with normal hearing. Performance was assessed in four different languages (Spanish, Polish, Russian, and German). In a speech-shaped noise masker, the authors report higher thresholds for Spanish (-7.2 dB SNR) and German (-7.4 dB SNR) target speech compared with Russian (-10.2 dB) and Polish (-9.4 dB). In terms of the linguistic context in which participants live, it might be that relying on English for day-to-day communication impacts their word recognition in the L1. In addition, it is important to mention that there are dialectal variations in some of the Spanish words in the stimuli set (e.g., hanger: gancho, percha; woman: mujer, señora; balloon: globo, bomba, vejiga; sweater: suéter, saco, chamarra), but in English, all of the words chosen for the stimuli set do not have other options, which may decrease the overall difficulty in English.

Effect of Mixing Languages in Word Recognition in Noise

In the mixed-language conditions, the prediction was that switching from the majority language (English or Spanish when presented 75% of the time) to the minority language (English or Spanish when presented 25% of the time) would yield a cost for the minority language. The performance was indeed better when the target language was predictable across a block of trials compared with when English and Spanish were randomly intermixed within a block of trials. Mixing the two languages reduced performance by about six percentage points compared with the fixed-language baseline for both English and Spanish. However, poor performance in word recognition was also observed when the direction of change was from the minority to the majority language. These results suggest that regardless of the direction of the switch, word recognition in noise was affected by mixing the languages. Another possibility is that a difference of 25% versus 75% was

not large enough to observe effects related to language expectancy.

The observation that the effect of code-switching was equivalent across language is in line with previous findings of symmetrical code-switching costs in highly proficient bilinguals using visual stimuli to name pictures and/or numbers while switching languages (Costa et al., 2006; Costa & Santesteban, 2004; Meuter & Allport, 1999). Conversely, Olson (2017) reported asymmetric results from highly proficient bilingual listeners in the monolingual mode condition (majority of words in one language, minority of words in the other language), in which the cost was larger when switching to the L1 than to the L2. However, the stimuli used by Olson (2017) included sentences instead of words, which create linguistic demands that may impact the cost of switching languages in a different way.

In addition to the poorer performance associated with switching from the majority to the minority language in the mixed conditions, poorer word recognition when switching from the minority to the majority language was also observed. Although unexpected, this pattern of results indicates that mixing the languages alone was enough to yield a switching cost for the bilingual listeners for word recognition in noise. One interpretation of these results is that switching between languages in noisy environments taxes the listener's cognitive resources, which, in turn, hurts performance. Another alternative, however, is that intermixing languages effectively increases the number of targets that the listener must consider when selecting a response. Previous research has suggested that a small number of response alternatives can improve performance by reducing the number or quality of speech cues required to select the correct response (e.g., Miller, Heise, & Lichten, 1951). For example, if the only clearly audible phoneme is /s/, the listener is more likely to select the correct response from a small set of alternatives (e.g., sweater, candy, baby, and turkey) than a larger set (e.g., sweater/suéter, zebra/cebra, baby/bebé, and turkey/pavo). This is because the smaller set is less likely to contain multiple instances of the clearly audible phoneme. In this example, there is only one instance of /s/ in the smaller set (sweater), but multiple instances in the larger set (sweater/suéter and zebra/cebra). In order to evaluate the consequences of set size on this task, we performed a second experiment in which speech recognition performance was compared across 4AFC and seven-alternative forced choice (7AFC) conditions.

Experiment 2: Response Set Size

The second experiment evaluated the consequences of increasing the number of words associated with each picture in a forced-choice task. Theoretically, increasing the number of potential labels for each picture could increase the quality or number of acoustic cues required to perform well on the task. We hypothesized that intermixing languages effectively increases the number of targets that the listener must consider when selecting a response, which, in turn, may impact accuracy. The decision to

compare performance in the 4AFC and the 7AFC was based on the observation that some pairs of English and Spanish words used in Experiment 1 are phonetically similar (e.g., sweater/suéter). Of the set of 30 pairs of words, 12 were cognates, meaning these word pairs share phonological and semantic similarity between the languages. For example, monster and monstruo are cognates, but candy and dulce are not. In the 4AFC task of Experiment 1, listeners were presented with a mix of noncognate and cognate words. Because cognates are phonetically similar, mixing English and Spanish would not double the number of phonetic templates to consider in selecting a response. On average, only one of the four alternatives would be a cognate. In the mixed-language conditions, this would increase the number of possible target words from four to seven (e.g., monster/monstruo, candy/dulce, chicken/pollo, and table/mesa).

Method

Listeners

Twelve English monolingual adults, with normal hearing (≤ 20 dB HL bilaterally at octave frequencies from 0.25 to 8 kHz), aged between 21.3 and 31.1 years ($M = 26.2$; $SD = 3.27$) participated in the experiment.

Procedure

The target and masker stimuli were the same as used in Experiment 1. The task was the closed-set word recognition in the same speech-shaped noise masker used in Experiment 1. There were two set-size conditions. In the first set-size condition, listeners selected their answers from among four alternatives, as in Experiment 1. In the second set-size condition, listeners selected from among seven alternatives. Instead of measuring the percent correct, the signal level was adapted using the two-down, one-up stepping rule to estimate 71% correct recognition performance. The initial step size was 4 dB, reducing to 2 dB after the second track reversal. The testing continued until eight reversals were obtained. The threshold was computed as the mean of the last six reversals. Each listener provided two thresholds in each condition. A third estimate was obtained if the first two differed by more than 3 dB. Data were collected and blocked by target language, and participants completed two conditions blocked by language in random order.

Data were combined across listeners to estimate the psychometric function for each condition (see Figure 3). Data and fits to data in the 4AFC are shown on the left, and those in the 7AFC are shown on the right. Percent correct was estimated at 51 dB SPL; the signal level was evaluated in Experiment 1. That value was 82.4% for the 4AFC and 78.8% for the 7AFC. While poorer performance in the 7AFC task is qualitatively consistent with the hypothesis that a larger set size hurts the performance by increasing the phonetic detail required to do the task, it is also consistent with the lower level of chance performance associated with a larger set size. The most obvious manifestation of this effect is evident in the lower asymptote of

Figure 3. Psychometric function comparing 4AFC (left) and 7AFC (right) tasks. Percent correct word recognition was estimated at 51 dB SPL. 4AFC = four-alternative forced choice; 7AFC = seven-alternative forced choice.

the fitted functions, which is 25% for the 4AFC and 14% for the 7AFC.

To accommodate differences in chance, the performance was converted from percent correct to d' . If the larger set of response alternatives in the 7AFC hurt performance apart from effects related to lower levels of chance performance, then d' at 51 dB SPL would be smaller for the 7AFC than the 4AFC. This was not observed. For a 4AFC, a performance of 82.4% correct corresponds to a d' of 2.0. For a 7AFC task, a performance of 78.8% correct corresponds to a d' of 2.2. In other words, after correcting for effects related to guessing, the mean performance was better in the 7AFC than in the 4AFC. This result is inconsistent with the idea that the set size of response alternatives hurt performance in the mixed-language conditions in Experiment 1.

Discussion

One possible explanation for poorer performance in the mixed-language condition than in the single-language conditions of Experiment 1 is the larger number of response alternatives available when each of the four pictures provided could be identified in either English or Spanish. Experiment 2 attempted to evaluate the contribution of the number of response alternatives on performance by comparing d' for a 51 dB SPL signal for a 4AFC and a 7AFC. Listeners were all monolingual English speakers, and target words were in English. The increase in response alternatives associated with the mixed-language conditions was modeled with a 7AFC because some English/Spanish word pairs were cognates. Cognates are phonetically similar in the two languages, and the rationale was that phonetic cues supporting recognition would be similar in these cases. Results showed that there was not a difference in performance after accounting for differences in chance performance. This result provides support for the conclusion that

the results in the first experiment represent costs associated with switching languages and cannot be explained simply in terms of the increased number of picture labels available in the mixed-language conditions.

Conclusion

The main result of this study is that Spanish/English bilingual listeners' recognition of code-switched words in noise was negatively impacted relative to single-language recognition, even when the linguistic context is relatively simple and the task is not cognitively demanding. The code-switching cost observed in this study was not significantly different for Spanish and English. This finding is consistent with previous reports in code-switching studies with highly proficient bilinguals. Recall that the target words used in this study are frequent and familiar words in both English and Spanish and are part of the first-grade child lexicon (Calandruccio et al., 2014). Although demands of the task could have influenced listeners' identification of the code-switched words, this is unlikely because the 4AFC reduces stimulus uncertainty by having limited alternative response options and, in this case, the visual representation of the auditory stimuli. Future research should include listeners with a wider range of proficiency levels in each language to determine whether language proficiency influences word recognition in noise during code-switching and if the cost of code-switching can be predicted by their language dominance or the age of acquisition of the L2. There was a wide range in age of L2 acquisition for the participants in this study. However, their proficiency measures were very similar not only in the self-reported questionnaire but also in the Versant Test, which indicated that regardless of their age of acquisition, they all had attained high proficiency in L2. An analysis of how language dominance predicts word recognition in noisy environments during code-switching would offer more insight.

However, the data collected in this study do not offer enough power to conduct the analysis. Previous work demonstrated that nonnative listeners are more susceptible to masking than native speakers (Florentine et al., 1984; Mendel & Widner, 2016), and the present results indicate that mixing native and nonnative language speech may further degrade performance. Further research in support of this finding is critical to provide evidence-based counseling for bilingual patients who are struggling to communicate when listening in noise. If these data are substantiated further, it would suggest that effective counseling would include explaining to patients and their loved ones that code-switching may be detrimental to good speech recognition in noisy environments. Therefore, switching back and forth between the L1 and the L2 could make comprehension harder, not easier, for bilingual populations in noisy environments; however, other considerations, such as word familiarity and/or limited lexicon in their L2, could impact masked speech recognition and would need to be explored prior to making such claims.

Acknowledgments

This research was funded by the National Institute on Deafness and Other Communication Disorders Grant R01 DC015056. Participant recruitment was facilitated by the Clinical Measurement Core of Boys Town National Research Hospital, which is supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award P20GM109023. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Abutalebi, J., Brambati, S., Annoni, J. M., Moro, A., Cappa, S., & Perani, D. (2007). The neural cost of the auditory perception of language switches: An event-related functional magnetic resonance imaging study in bilinguals. *The Journal of Neuroscience*, 27, 13762–13769. <https://doi.org/10.1523/JNEUROSCI.3294-07.2007>
- Alvarez, R. P., Holcomb, P. J., & Grainger, J. (2003). Accessing word meaning in two languages: An event-related brain potential study of beginning bilinguals. *Brain and Language*, 87, 290–304. [https://doi.org/10.1016/S0093-934X\(03\)00108-1](https://doi.org/10.1016/S0093-934X(03)00108-1)
- Barker, G. (1947). Social functions of language in a Mexican-American community. *Acta Americana*, 5, 185–202.
- Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. Austin, TX: COERLL, University of Texas. Retrieved from <https://sites.la.utexas.edu/bilingual/>
- Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford–Kowal–Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, 13, 108–112.
- Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, 121(4), 2339–2349. <https://doi.org/10.1121/1.2642103>
- Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, 112(1), 272–284. <https://doi.org/10.1121/1.1487837>
- Calandruccio, L., Gomez, B., Buss, E., & Leibold, L. J. (2014). Development and preliminary evaluation of a pediatric Spanish/English speech perception task. *American Journal of Audiology*, 23, 158–172. https://doi.org/10.1044/2014_AJA-13-0055
- Chauncey, K., Grainger, J., & Holcomb, P. J. (2008). Codeswitching effects in bilingual word recognition: A masked priming study with event-related potentials. *Brain and Language*, 105, 161–174. <https://doi.org/10.1016/j.bandl.2007.11.006>
- Cooper, G. (2013). An exploration of intentions and perceptions of code-switching among bilingual Spanish/English speakers in the inland northwest. *Journal of Northwest Anthropology*, 47, 215–225.
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50, 491–511. <https://doi.org/10.1016/j.jml.2004.02.002>
- Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 32, 1057–1074. <https://doi.org/10.1037/0278-7393.32.5.1057>
- Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford, United Kingdom: Oxford University Press.
- Filippi, R., Karaminis, T., & Thomas, M. S. C. (2014). Language switching in bilingual production: Empirical data and computational modelling. *Bilingualism: Language and Cognition*, 17(2), 294–315. <https://doi.org/10.1017/S1366728913000485>
- Flamme, G. A., Stephenson, M. R., Deiters, K., Tatro, A., VanGessel, D., Geda, K., . . . McGregor, K. (2012). Typical noise exposure in daily life. *International Journal of Audiology*, 51(1), S3–S11. <https://doi.org/10.3109/14992027.2011.635316>
- Florentine, M., Buus, S., Scharf, B., & Canevet, G. (1984). Speech reception thresholds in noise for native and non-native listeners. *The Journal of the Acoustical Society of America*, 75, S84. <https://doi.org/10.1121/1.2021645>
- Garcia Lecumberri, M. L., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America*, 119(4), 2445–2454. <https://doi.org/10.1121/1.2180210>
- Green, W. D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81.
- Hazan, V., & Simpson, A. (2000). The effect of cue-enhancement on consonant intelligibility in noise: Speaker and listener effects. *Language and Speech*, 43, 273–294. <https://doi.org/10.1177/00238309000430030301>
- Hochmuth, S., Birger, K., Brand, T., & Jurgens, T. (2015). Influence of noise type on speech perception thresholds across four languages measured with matrix sentence test. *International Journal of Audiology*, 54, 1499–2017. <https://doi.org/10.3109/14992027.2015.1046502>
- Hodgson, M., Steininger, G., & Razavi, Z. (2007). Measurement and prediction of speech and noise levels and the Lombard effect in eating establishments. *The Journal of the Acoustical Society of America*, 121, 2023–2033.
- Macizo, P., Bajo, T., & Paolieri, D. (2012). Language switching and language competition. *Second Language Research*, 28(2), 131–149. <https://doi.org/10.1177/0267658311434893>
- 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, 940–967. [https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067))
- Mayo, L. H., Florentine, M., & Buus, S.** (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research*, 40, 686–693.
- Mendel, L., & Widner, H.** (2016). Speech perception in noise for bilingual listeners with normal hearing. *International Journal of Audiology*, 55(2), 126–134. <https://doi.org/10.3109/14992027.2015.1061710>
- Meuter, R. F., & Allport, A.** (1999). Bilingual language switching in naming: Asymmetrical switch costs of language naming. *Journal of Memory and Language*, 40, 25–40. <https://doi.org/10.1006/jmla.1998.2602>
- Miller, G. A., Heise, G. A., & Lichten, W.** (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41, 329–335.
- Moreno, E., Federmeier, K., & Kutas, M.** (2002). Switching languages, switching palabras (words): An electrophysiologic study of code switching. *Brain and Language*, 80, 188–207. <https://doi.org/10.1006/brln.2001.2588>
- Myers-Scott, C.** (2006). Natural codeswitching knocks on the laboratory door. *Bilingualism: Language and Cognition*, 9(2), 203–212. <https://doi.org/10.1017/S1366728906002549>
- Nábělek, A. K., & Donahue, A. M.** (1984). Perception of consonants in reverberation by native and non-native listeners. *The Journal of the Acoustical Society of America*, 75, 632–634.
- Olson, D.** (2017). Bilingual language switching cost in auditory comprehension. *Language, Cognition and Neuroscience*, 32(4), 494–513. <https://doi.org/10.1080/23273798.2016.1250927>
- Pearson.** (2011). *Versant English Test description and validation summary*. Retrieved from <http://www.versanttest.com/technology/VersantEnglishTestValidation.pdf>
- Piccini, P. E., & Garellek, M.** (2014). Prosodic cues to monolingual versus code-switching sentences in English and Spanish. In N. Campbell, D. Gibbon, & D. Hirst (Eds.), *Social and linguistic speech prosody: Proceedings of the 7th International Conference on Speech Prosody* (pp. 885–889). Retrieved from <http://fastnet.netsoc.ie/sp7/>
- Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B.** (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, 27, 465–485. <https://doi.org/10.1017/S014271640606036X>
- Rusnock, C., & McCauley, P.** (2012). Case study. *Journal of Occupational and Environmental Hygiene*, 9(6), D108–D113. <https://doi.org/10.1080/15459624.2012.683716>
- Shi, L. F.** (2015). How proficient is “proficient” bilingual listeners’ recognition of English words in noise. *American Journal of Audiology*, 24, 53–65. https://doi.org/10.1044/2014_AJA-14-0041
- Spivey, M., & Marian, V.** (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, 10(3), 281–284.
- Studebaker, G.** (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455–462. <https://doi.org/10.1044/jshr.2803.455>
- van Hell, J. G., & Wittenman, M. J.** (2009). The neurocognition of switching between languages: A review of electrophysiological studies. In L. Isurin, D. Winford, & K. de Bot (Eds.), *Multi-disciplinary approaches to code switching* (pp. 53–84). Philadelphia, PA: John Benjamins.
- van Wijngaarden, S. J., Bronkhorst, A. W., Houtgast, T., & Steeneken, H. J. M.** (2004). Using the speech transmission index for predicting non-native speech intelligibility. *The Journal of the Acoustical Society of America*, 115(3), 1281–1291. <https://doi.org/10.1121/1.1647145>
- Weiss, D., & Dempsey, J. J.** (2008). Performance of bilingual speakers on the English and Spanish versions of the Hearing in Noise Test (HINT). *Journal of the American Academy of Audiology*, 19, 5–17.