Research Article

Effects of Language History on Sentence Recognition in Noise or Two-Talker Speech: Monolingual, Early Bilingual, and Late Bilingual Speakers of English

Diana Regalado,^a Jessica Kong,^b Emily Buss,^a and Lauren Calandruccio^b

Purpose: Language history is an important factor in masked speech recognition. Listeners who acquire the target language later in life perform more poorly than native speakers. However, there are inconsistencies in the literature regarding performance of bilingual speakers who begin learning the target language early in life. The purpose of this experiment was to evaluate speech-in-noise and speech-in-speech recognition for highly proficient early bilingual listeners compared to monolingual and late bilingual listeners.

Method: Three groups of young adults participated: native monolingual English speakers, bilingual Mandarin–English speakers who learned English from birth (early bilinguals), and native Mandarin speakers who learned English later in life (late bilinguals). All participants had normal hearing and were full-time college students. Recognition was assessed for English sentences in speech-shaped noise and two-

talker English speech. Participants provided linguistic and demographic information, and late bilinguals completed the Versant test of spoken English abilities.

Results: All listeners performed better in speech-shaped noise than two-talker speech. Performance was similar for monolingual and early bilinguals. Late bilinguals performed more poorly overall. There was evidence for a stronger association between masked speech recognition and English dominance for late bilinguals compared to early bilinguals.

Conclusion: These results support the conclusion that bilingualism itself does not necessarily result in a disadvantage when recognizing masked speech in noise and speech in speech. For populations similar to those studied here (highly proficient early bilinguals), it would be appropriate to evaluate masked speech recognition using the same simple stimuli and normative data used for monolingual speakers of English.

It is well documented that being bilingual provides many advantages (Janus & Bialystok, 2018; Rivera Mindt et al., 2008; Treccani, Argyri, Sorace, & Sala, 2009). The obvious benefit is that bilingualism supports communication with a broad and diverse group of people. Bilingualism has also been associated with functional advantages, such as stronger executive control (Carlson & Meltzoff, 2008), improved cognitive function (Bialystok, 2015), greater auditory capacity (Motlagh Zadeh, Jalilvand

Karimi, & Silbert, 2018), lower auditory thresholds for nonspeech stimuli (Krizman, Bradlow, Lam, & Kraus, 2017), and reduced likelihood of stroke (Alladi et al., 2016), as well as greater cognitive reserves and delayed symptoms of dementia (Guzmán-Vélez & Tranel, 2015). Despite these advantages, it is important to recognize the potential disadvantages that bilingual audiology patients might experience when it comes to recognizing speech in noisy environments. Clinically, this is an important issue given that one in every five persons in the United States speaks a language other than English (Camarota & Zeigler, 2014), and clinicians need to know how to best evaluate and interpret speech recognition performance of bilingual patients.

A large body of research indicates better masked English speech recognition for native speakers of English compared to those who learned English later in life (e.g., Mayo, Florentine, & Buus, 1997; Tabri, Abou Chacra, & Pring, 2011). This is true even for people who have excellent

^aDepartment of Otolaryngology/Head and Neck Surgery, University of North Carolina at Chapel Hill

^bDepartment of Psychological Sciences, Case Western Reserve University, Cleveland, OH

Correspondence to Lauren Calandruccio:

Lauren.Calandruccio@case.edu

Editor-in-Chief: Sumitrajit Dhar

Editor: Monita Chatterjee

Received December 18, 2018

Revision received April 24, 2019

Accepted August 8, 2019

https://doi.org/10.1044/2019_AJA-18-0194

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

English speech recognition in quiet (Crandell & Smaldino, 1996; Mayo et al., 1997; Rogers, Lister, Febo, Besing, & Abrams, 2006; Tabri et al., 2011; von Hapsburg, Champlin, & Shetty, 2004), including those who have lived in the United States or Canada for long periods of time (Meador, Flege, & Mackay, 2000; Rimikis, Smiljanić, & Calandruccio, 2013; Shi, 2010). Poor masked speech recognition in bilingual listeners is often attributed to a reduced familiarity with the later-acquired language, resulting in poorer performance when access to speech cues is reduced by the presence of a masker. This view is supported by the observation that language proficiency is highly correlated with masked speech recognition in bilingual listeners (Kilman, Zekveld, Hällgren, & Rönnberg, 2014; Rimikis et al., 2013).

However, some data indicate that bilingualism itself, independent of proficiency, may be associated with a cost when it comes to masked speech recognition. This result is sometimes described as a “bilingual disadvantage” (Kraus & White-Schwoch, 2017; Krizman et al., 2017; Schmidtke, 2016). Being bilingual could reduce speech recognition by virtue of the larger overall number of lexical items in memory, leading to more lexical competition during recognition. Alternatively, reduced frequency of exposure to words in the test language, could lead to lower activation of those words in memory (Persici, Vihman, Burro, & Majorano, 2019; Schmidtke, 2016). However, there is substantial disagreement in the literature regarding the existence of a bilingual disadvantage for listeners who are highly proficient in the test language. Some studies show poorer masked speech recognition in highly proficient bilingual listeners compared to monolingual listeners (Krizman, Skoe, & Kraus, 2016; Mayo et al., 1997; Shi, 2010; Tabri et al., 2011), but other studies indicate no disadvantage for highly proficient bilinguals (Calandruccio & Zhou, 2014; Reetzke, Lam, Xie, Sheng, & Chandrasekaran, 2016; Shi, 2009).

One important factor to consider when comparing results across published studies on speech recognition in bilingual listeners is the demographic and linguistic characteristics of the listeners (von Hapsburg & Peña, 2002). In the United States, monolingual and bilingual speakers of English tend to differ along a range of dimensions that could affect masked speech recognition, including linguistic knowledge and access to educational resources (Calvo & Bialystok, 2014; Schneider, Martinez, & Owens, 2006). One of the most studied factors is the age at which the test language was acquired (Meador et al., 2000; Neave-DiToro, Rubinstein, & Neuman, 2017; Shi, 2010, 2014; Weiss & Dempsey, 2008). Late bilinguals perform more poorly than early bilinguals (e.g., Mayo et al., 1997), and age of acquisition covaries with many of the other predictors of language proficiency (Flege, 2019). The listener’s habitual language use and the listening context also affect performance (Grosjean, 1998).

Another important methodological consideration that could affect performance of bilingual listeners compared to monolingual listeners in previous studies is differences in the stimuli used to assess masked speech

recognition (Calandruccio & Smiljanić, 2012). For example, Krizman et al. (2017) reported that highly proficient bilingual listeners performed similarly to monolingual listeners on a noise-masked word recognition task, but there was a bilingual disadvantage for noise-masked sentence recognition. This result was attributed to the bilingual listeners’ more limited knowledge of the statistical co-occurrence of words in the test language. Masked speech recognition also depends critically on particular characteristics of the masker. Maskers that are composed of competing speech hurt performance via informational masking—difficulty attending to the target and segregating it from the masker signal (Brungart, Simpson, Ericson, & Scott, 2001; Carhart, Johnson, & Goodman, 1975). Speech maskers can pose a significant challenge even for monolingual listeners, and speech-in-speech recognition continues to mature into adolescence under these conditions (e.g., Corbin, Bonino, Buss, & Leibold, 2016; Wightman & Kistler, 2005). In contrast to speech maskers, noise maskers introduce a greater contribution of energetic masking, whereby acoustical overlap between the target and masker signals within the auditory periphery limits neural encoding of the target (e.g., Hawkins & Stevens, 1950). There are some data indicating that the difference between monolingual and bilingual listeners’ masked speech recognition is larger when the masker is composed of speech, compared to noise (Cooke, García Lecumberri, & Barker, 2008; García Lecumberri & Cooke, 2006; Kilman et al., 2014). This difference may also vary depending on signal-to-noise ratio (SNR; Rogers et al., 2006) and semantic context (Mayo et al., 1997). Some data indicate a difference between monolinguals and bilinguals in the rate of performance change or psychometric function slope, but this result is not consistent; some data indicate that functions are steeper for monolinguals than late bilinguals (Mayo et al., 1997), but other data indicate the opposite trend (Rogers et al., 2006).

Variability in subject populations and stimuli are important for interpreting the conflicting results in the literature regarding whether highly proficient bilinguals have greater difficulty understanding speech in unfavorable conditions compared to monolingual listeners. Most reports in the literature describe linguistic background but do not consider access to education (e.g., Rogers et al., 2006; Shi, 2010). One exception is the study by Krizman et al. (2016). The bilingual and monolingual groups in that study had similar incidences of low maternal education (an often used proxy for socioeconomic status). However, those listeners were high school freshmen, with a mean age of 14.6 years. Since speech-in-speech recognition may be immature at this age even for monolingual listeners (Wightman & Kistler, 2005), it is unclear whether results of this study represent asymptotic, mature performance, which could occur later for bilingual children.

Therefore, it remains unclear whether early bilinguals who are highly proficient in English perform similarly to their monolingual peers on English speech-in-noise and speech-in-speech tasks or whether their performance is intermediate between their monolingual and late bilingual

peers. Clinical assessment of English speech recognition performance depends on an accurate understanding of the role that language experience plays in this task. The purpose of the present experiment was to investigate sentence recognition in a steady-state noise masker and a complex two-talker speech masker for highly proficient early bilingual listeners compared to monolingual and late bilingual listeners.

To minimize factors that could confound the results, such as large differences in socioeconomic status, we recruited three groups of full-time college students from Case Western Reserve University. This recruitment strategy was used to provide a “best case” scenario for observing good performance in bilingual listeners with respect to educational resources. All bilingual participants were highly proficient in English and communicated in English on a daily basis, including in their coursework; bilinguals also spoke Mandarin. There were three main hypotheses: (a) that early bilingual adult listeners would perform similarly to their monolingual peers (see Calandruccio & Zhou, 2014, for a similar result), (b) that late bilinguals would perform more poorly overall than their monolingual and early bilingual peers (e.g., Rogers et al., 2006), and (c) that there would be a significant interaction between listener group and masker type, with late bilinguals being more adversely affected while listening to English sentences in the two-talker masker relative to the two other listener groups (e.g., Cooke et al., 2008). The primary outcome measure was speech recognition threshold (SRT), but psychometric function slope was also examined to determine how SNR might affect the pattern of results across groups.

Method

Participants

Three groups of 20 listeners were recruited from the undergraduate and graduate student population of Case Western Reserve University. Monolingual participants were born in the United States to native English-speaking parents (13 women, ages 18–27 years, $M = 21.4$ years). Early bilingual participants were born in the United States to native Mandarin-speaking parents and acquired English and Mandarin at the same time (11 women, ages 18–28 years, $M = 20.6$ years). Late bilingual participants were born in China, learned Mandarin as their first language, and acquired English as their second language (16 women, ages 19–31 years, $M = 22.1$ years).¹ Demographic information on English exposure for the late bilingual participants is provided in Table 1. Participants who were recruited from the Department of Psychological

Sciences research subject pool were compensated for their time through course credit, and others received \$15/hr.

All participants passed a hearing screening, with thresholds of ≤ 25 dB HL bilaterally at octave frequencies between 250 and 8000 Hz (American National Standards Institute & Acoustical Society of America, 2009). Bilinguals completed the English–Mandarin Bilingual Language Profile (BLP; Birdsong, Gertken, & Amengual, 2012), a self-report measure that assesses language history, proficiency, use, and attitudes for English–Mandarin bilinguals. Scores on the BLP range from –218 (indicating Mandarin dominance) to 218 (indicating English dominance); a score near 0 indicates a balanced bilingual (see Table 2). The English–Mandarin BLP is appropriate for any person who speaks both English and Mandarin, regardless of their proficiency in either language. Late bilinguals also completed the English Versant Test (Pearson), an automated speech recognition test designed to assess the spoken English language skills for speakers who acquire English as a second language. The English Versant Test provides five numeric scores: four subscores (sentence mastery, vocabulary, fluency, and pronunciation) and one overall score. Scores range from 20 to 80, with 80 indicating native-like English fluency. The English Versant Test is not appropriate for use with monolingual and early bilingual speakers with high degrees of English fluency. The Versant was therefore not administered to monolinguals or early bilinguals. Both early and late bilingual participants were also asked to self-rate their level of proficiency for reading, writing, speaking, and listening in Mandarin and in English using a scale from 1 to 5 (with 1 indicating an *ability to speak and understand phrases and sentences* and 5 indicating *native-like language abilities*).

Stimulus Development

Three 21-year-old female actors (Talkers A–C) were hired to record stimuli. All were native speakers of American English with no noticeable regional dialect. Speech for the target and the two-talker masker consisted of sentences from the Bamford–Kowal–Bench (BKB) speech corpus, a corpus of speech that was developed based on spoken sentences from deaf and hard-of-hearing children (Bench, Kowal, & Bamford, 1979). These sentences have an elementary syntax structure and vocabulary, making them appropriate for use with nonnative speakers of English (Calandruccio & Buss, 2017). The corpus includes 21 lists of 16 sentences each, for a total of 336 sentences. The BKB corpus was used as target stimuli to increase the likelihood that the vocabulary used within the sentences and the grammar of the sentences would be highly familiar to all three listener groups. Talker A was chosen as the target talker, and Talkers B and C were masker talkers. Talker A had an average speaking rate of 4.8 syllables/second and fundamental frequency (F0) of 230 Hz. Talker B had an average speaking rate of 3.7 syllables/second and F0 of 230 Hz. Talker C had an average speaking rate of 4.2 syllables/second and F0 of 235 Hz.

¹Although all of our late bilingual subjects self-identified as native speakers of Mandarin, they did not all report speaking Mandarin at home. There are hundreds of dialects of Chinese, many of which are not mutually intelligible. However, standard Chinese is a form of Mandarin. It is therefore likely that our subjects who did not report speaking Mandarin at home were nonetheless exposed to it at an early age in their local communities. All late bilinguals were formally educated in Mandarin.

Table 1. Demographic and linguistic information for late bilingual participants.

Subject ID	Age (in years)	Gender	Age of English acquisition (self-report in years)	Age moved to the United States (in years)	Highest level of education completed in Mandarin	Languages spoken in household	Overall Versant score	Length of residency in the United States (in years)
LB1	22	Male	5	22	Bachelor's degree	Mandarin	69	0
LB2	23	Female	4	21	Bachelor's degree	Mandarin	55	2
LB3	31	Male	7	25	Bachelor's degree	Mandarin	54	6
LB4	19	Female	5	16	Less than high school	Mandarin, Cantonese	58	3
LB5	21	Female	7	18	High school	Mandarin, Shanghai dialect	52	3
LB6	20	Female	6	15	Less than high school	Mandarin	72	5
LB7	23	Male	8	18	High school	Chengdu dialect	69	5
LB8	21	Female	7	17	High school	Mandarin	69	4
LB9	19	Female	4	18	High school	Mandarin	64	1
LB10	22	Female	3	22	Bachelor's degree	Mandarin	51	0
LB11	23	Female	5	16	High school	Mandarin	64	7
LB12	19	Female	8	14	Less than high school	Mandarin	67	5
LB13	22	Female	6	21	Bachelor's degree	Funing dialect	51	1
LB14	19	Female	5	14	Less than high school	Mandarin	56	5
LB15	26	Female	5	26	Some college	Mandarin	45	0
LB16	20	Female	5	18	Some college	Mandarin, Wuhan dialect	59	2
LB17	22	Female	7	22	Bachelor's degree	Mandarin, Dongbei dialect	55	0
LB18	24	Female	10	24	Bachelor's degree	Mandarin	55	0
LB19	21	Female	6	19	High school	Mandarin	66	2
LB20	24	Male	10	22	Bachelor's degree	Mandarin, Guiliu dialect, Kejia dialect	49	2

Recordings were made in a double-walled sound-isolated booth, where the speaker stood 6 in. in front of a Shure KSM-42 omnidirectional cardioid condenser microphone connected to an M-Audio M-Track 2x2 converter. Sentences were recorded using a 44.1-kHz sampling rate (16-bit). All sentences were root-mean-square (RMS) equalized using Praat. Sentences produced by each masker talker were concatenated, and the two streams were summed to generate the two-talker masker. A speech-shaped noise (SSN) masker was also generated, with the same long-term average magnitude spectrum of the two-talker speech masker. Masker stimuli were gated on/off using 50-ms raised cosine ramps, with a 500-ms fringe at the beginning and end of each target sentence.

Procedure

All testing was completed in a double-walled, sound-isolated room. Hearing thresholds were obtained using an AudioStar Pro audiometer with either supra-aural (TDH) headphones or insert (IP30 RadioEar) earphones. Differences in stimulus delivery were based on tester preference. All participants tested with insert phones also received an otoscopic evaluation to ensure clear ear canals prior to inserting the foam tips of the insert phones. Speech recognition was evaluated in the free field using a pair of loudspeakers positioned at $\pm 45^\circ$ (left/right) relative to the participant, and stimuli played diotically. The experiment was controlled using a custom MATLAB script, and stimuli were presented at 44.1 kHz.

Instructions for the sentence recognition task were to repeat back each sentence spoken by the target talker while ignoring the background noise or background speech. Participants were told that the target talker would be louder than the noise or other voices at the start of the experiment, but that it would gradually become more difficult to hear. In the event that they did not hear or understand the target, they were encouraged to guess. They were specifically instructed to repeat back all words that they thought they heard, even if the string of words did not make sense.

The target was fixed at 56 dB SPL to ensure audibility, and the masker level was adapted dependent upon the listener's response. The first trial used an SNR of 5 dB. This relatively easy SNR was chosen to ensure all participants were comfortable with and understood the task prior to more difficult trials. Trials proceeded with two interleaved one-down, one-up adaptive tracks comprising 32 sentences in total. Adjustments of SNR were made in steps of 8 dB prior to the first reversal; this was reduced to 4 dB after the second reversal and further reduced to 2 dB after the third reversal. One track used a lax criterion, decreasing the SNR if the listener got one or more key words correct. The other track used a strict criterion, decreasing the SNR if the listener got three or more key words correct. These two stepping rules ensure acquisition of data points above and below 50% correct recognition by word. Psychometric functions were fitted to the word-level data obtained in the two interleaved tracks for each listener and condition. That function was a logit, defined as $y = \frac{1}{1 + \exp(-(x-\alpha)/\beta)}$, where y

Table 2. English and Mandarin language information for early and late bilingual participants.

Listener group	Listener ID	English				Mandarin				% Daily time exposure to English	BLP dominance score	BLP proficiency scores	
		Reading	Writing	Speaking	Listening	Reading	Writing	Speaking	Listening			English	Mandarin
Early bilingual	EB1	5	5	5	5	2	2	4	4	70	71.93	54.48	38.59
Early bilingual	EB2	5	5	5	5	3	3	4	4	90	97.81	54.48	31.78
Early bilingual	EB3	5	5	5	5	4	4	4	4	90	89.82	54.48	31.78
Early bilingual	EB4	5	5	5	5	2	2	3	3	95	109.16	54.48	18.16
Early bilingual	EB5	5	5	5	5	4	4	4	4	90	40.14	45.4	38.59
Early bilingual	EB6	5	5	5	5	1	1	4	4	90	107.43	54.48	20.43
Early bilingual	EB7	5	5	5	5	3	3	4	4	80	104.44	54.48	22.7
Early bilingual	EB8	5	5	5	5	2	2	3	3	80	108.80	52.21	22.7
Early bilingual	EB9	5	5	5	5	1	1	3	3	98	106.25	54.48	13.62
Early bilingual	EB10	5	5	5	5	4	4	5	5	94	48.95	54.48	38.59
Early bilingual	EB11	5	5	5	5	3	3	4	4	80	91.63	54.48	34.05
Early bilingual	EB12	5	5	5	5	1	1	5	5	90	90.09	54.48	24.97
Early bilingual	EB13	5	5	5	5	2	2	4	4	95	105.34	54.48	31.78
Early bilingual	EB14	5	5	5	5	2	2	5	5	60	61.3	54.48	31.78
Early bilingual	EB15	5	5	5	5	1	1	4	4	95	121.87	54.48	22.7
Early bilingual	EB16	5	5	5	5	5	5	5	5	80	22.16	54.48	54.48
Early bilingual	EB17	5	5	5	5	1	1	5	5	95	59.39	54.48	29.51
Early bilingual	EB18	5	5	5	5	1	1	3	3	90	107.61	54.48	20.43
Early bilingual	EB19	5	5	5	5	3	3	3	3	97	107.98	54.48	22.7
Early bilingual	EB20	5	5	5	5	4	4	5	5	70	57.22	54.48	40.86
Late bilingual	LB1	5	5	5	5	5	5	5	5	40	-19.97	54.48	54.48
Late bilingual	LB2	3	3	3	3	5	5	5	5	50	-82.73	29.51	52.21
Late bilingual	LB3	5	5	3	3	5	5	5	5	40	-80.91	36.32	54.48
Late bilingual	LB4	4	4	4	4	5	5	5	5	65	-48.58	38.59	54.48
Late bilingual	LB5	4	4	3	3	5	5	5	5	39	-75.28	38.59	54.48
Late bilingual	LB6	5	5	5	5	5	5	5	5	90	-13.89	49.94	54.48
Late bilingual	LB7	5	5	5	5	5	5	5	5	90	-15.98	47.67	54.48
Late bilingual	LB8	4	4	5	5	5	5	5	5	70	-45.30	40.86	54.48
Late bilingual	LB9	4	4	4	4	5	5	5	5	90	-43.13	45.4	54.48
Late bilingual	LB10	4	4	4	4	5	5	5	5	85	-59.47	38.59	54.48
Late bilingual	LB11	4	4	4	4	5	5	5	5	85	-52.48	43.13	54.48
Late bilingual	LB12	4	4	5	5	5	5	5	5	50	-26.05	34.05	47.67
Late bilingual	LB13	4	4	3	3	5	5	5	5	60	-69.29	36.32	52.21
Late bilingual	LB14	4	4	5	5	5	5	5	5	80	-42.77	47.67	54.48
Late bilingual	LB15	3	3	3	3	5	5	5	5	70	-97.16	24.97	54.48
Late bilingual	LB16	4	4	4	4	5	5	5	5	40	-63.29	43.13	54.48
Late bilingual	LB17	4	4	4	4	5	5	5	5	50	-72.37	36.32	54.48
Late bilingual	LB18	4	4	3	3	5	5	5	5	50	-86.90	31.78	54.48
Late bilingual	LB19	3	3	4	4	5	5	5	5	60	-90.62	27.24	54.48
Late bilingual	LB20	4	4	4	4	5	5	5	5	40	-47.13	43.13	52.21

Note. Reading, writing, speaking, and listening scores are based on a 5-point scale. BLP dominance scores range from -218 (indicating Mandarin dominance) to 218 (indicating English dominance). BLP proficiency scores range from 0 to 54.48 (highest possible score). BLP = Bilingual Language Profile.

is the proportion correct, x is the signal level in dB SNR, α is the SRT associated with 50% correct, and β is the function slope.

All participants completed two conditions presented in a randomized order: a two-talker masker condition and an SSN masker condition. All spoken responses were scored on-the-fly during the experiment by the tester who sat in a single-walled control room adjacent to the testing room. Scoring consisted of marking key words within the sentence as correct or incorrect. Any variations of key words (e.g., tense changes, plurals) were marked as incorrect. Testing took approximately 1 hr for each participant, with breaks taken periodically as needed.

Statistical analyses were conducted using linear regression models with fixed effects of language status (monolingual, early bilingual, late bilingual), masker type (SSN, two-talker), and the interaction of these two effects. To account for correlation due to repeated masker conditions, we examined multiple approaches. One model used a random intercept for subject, and other models included no random effect but modeled the correlation in the residuals directly. The heterogeneous compound symmetric correlation matrix yielded a smaller Akaike information criterion value than a random intercept mixed model because the heterogeneous matrix allowed the variances by masker type to differ. Restricted maximum likelihood estimates are reported, and analyses were carried out in SAS 9.4 (SAS Institute).

Results

Language Profile of Bilingual Participants

Early bilingual participants reported greater proficiency and comfort using English than Mandarin. They rated their English reading, writing, speaking, and listening abilities with a score of 5. They rated their Mandarin reading and writing between 1 and 5 ($M = 2.5$) and listening and speaking between 3 and 5 ($M = 4.1$). When not at college, all participants heard Mandarin spoken at home, and all but one also heard English at home. Nineteen of the 20 participants also had experience attending Mandarin language school. All participants spoke Mandarin with immediate and extended family, and 17 of them also spoke some English with family members. All participants spoke English with friends and at school and work, and two of them also reported speaking Mandarin in these contexts. Early bilinguals reported that they were currently exposed to Mandarin on average 13% of the time. Their average BLP dominance score was 85.5 (range: 22.2–121.9), consistent with an overall dominance of English over Mandarin.

The late bilingual group reported higher proficiency and dominance in Mandarin than in English. All participants rated their Mandarin reading, writing, speaking, and listening abilities with a score of 5 (see Table 2). Their ability to read, write, speak, and listen in English was rated between 3 and 5 for all four categories ($M = 4$ for each category). On average, late bilinguals learned English at 6 years of age and became comfortable speaking English

at 16.9 years old ($SD = 4.7$ years). All participants reported speaking Mandarin with family, friends, and Mandarin-speaking classmates, coworkers, and professors. They also reported speaking English with classmates, professors, coworkers, and non-Mandarin-speaking speaking friends. Late bilinguals reported that they were currently exposed to Mandarin on average 36% of the time. The average BLP dominance score was -56.7 (range: -97.2 to -13.9), consistent with dominance of Mandarin over English. The average English Versant score for late bilingual participants was 59 (range: 45–72; see Table 1). BLP dominance scores were significantly different between the two groups of bilinguals, as evaluated with a t test assuming unequal variances, $t(77.2) = 24.07$, $p < .0001$.

Masked Sentence Recognition

Fits to the adaptive track data were generally quite good, with a median r^2 value of .92 (r^2 interquartile range: .83–.96). There were only two fits for which r^2 was less than .5, both for early bilinguals tested in the two-talker speech masker. The SRT and slope values associated with fits to the adaptive track data are shown in Figure 1. Boxes indicate the distribution of values; listener group is indicated on the x -axis, and shading reflects the masker condition, as defined in the legend. The top panel shows estimates of the SRT, and the bottom panel shows estimates of slope. Circles show data for individual listeners, with x s indicating the two data points associated with function fits resulting in $r^2 < .5$.

The SRTs were lower for the SSN masker than the two-talker masker, and SRTs were lower for the monolinguals than the late bilinguals. On average, SRTs were 4.4 dB lower for sentence recognition in the SSN masker than the two-talker speech masker. Mean thresholds were lower for monolingual than late bilingual listeners (2.7 dB), but they were not significantly different for monolingual and early bilingual groups (-0.02 dB).

To test the significance of the observed differences in thresholds across maskers and conditions, a linear mixed-model regression analysis was conducted. There was no significant interaction between listener group and masker type, $F(2, 57) = 0.29$, $p = .745$. There was a significant group difference for monolingual and late bilingual listeners, $\text{coef} = 2.90$, $t(57) = 7.37$, $p < .001$, but no difference between monolingual and early bilingual listeners, $\text{coef} = 0.06$, $t(57) = 0.21$, $p = 0.838$. The effect of masker type was also significant, $\text{coef} = 4.54$, $t(57) = 11.08$, $p < .001$. Omitting data for the two early bilinguals with poor function fits did not change this pattern of results.

Differences in slope across group and masker type were more modest. There was visual evidence that slope was shallower for the late bilinguals than other groups, particularly for the SSN masker. A second analysis was performed to evaluate the significance of these observations. Values of slope underwent a natural log transformation prior to analysis because the slope values were left skewed. There was no significant interaction between listener group

Figure 1. Distribution of speech recognition thresholds (SRTs; top panel) and psychometric function slope (bottom panel) for each listener group, as indicated on the x-axis. Boxes span the 25th to 75th percentiles, and vertical lines span the 10th and 90th percentiles. Data for individual listeners are indicated with circles. Shading indicates masker type, which was either speech-shaped noise (open boxes and filled symbols) or two-talker speech (filled boxes and open symbols). The two data points associated with a poor fit ($r^2 < .5$) are indicated with xs.

and masker type, $F(2, 57) = 5.54$, $p = .120$. There was a significant group difference for monolingual and late bilingual listeners, $\text{coef} = 0.44$, $t(57) = 4.16$, $p < .001$, and no difference between monolingual and early bilingual listeners, $\text{coef} = 0.03$, $t(57) = 0.25$, $p = .806$. There was no significant effect of masker type, $\text{coef} = 0.01$, $t(57) = 0.05$, $p = .961$.

One possible explanation for the finding of shallower function slopes in the late bilinguals compared to the other two groups is that some of the vocabulary in the target sentences could have been sufficiently unfamiliar as to affect performance. This would result in greater variability in recognition threshold across words for the late bilinguals than for the other two groups, for whom all words were sufficiently familiar to support recognition at a low SNR. To evaluate this possibility, an analysis based on word frequency was performed. A pair of SRTs was estimated for each listener and condition, one based on responses to the lower frequency words within the corpus and one based on the higher frequency words. Word frequency was based on CELEX log-frequency per million values

(the Collins Birmingham University International Language Database [COBUILD] corpus; Davis, 2005), with lower frequency words being those with a score below the median score (1.86). Approximately 10% of the data fits had an $r^2 < .5$ and were omitted from further analysis. Results indicate lower thresholds for high-frequency than low-frequency words in late bilinguals tested in the SSN masker (1 dB), $t(17) = 3.25$, $p = .002$, one-tailed. There was a trend for this difference in the late bilinguals tested in the two-talker masker, but the effect was small and did not reach significance (0.3 dB), $t(16) = 0.98$, $p = .174$, one-tailed. There was no evidence of a threshold difference between low- and high-frequency words in either masker condition for the monolingual and early bilingual listeners ($p \geq .214$, one-tailed). The lack of an effect of word frequency in monolingual and early bilingual listeners is not surprising given the overall simplicity of the BKB sentence corpus.

The association between BLP dominance scores and SRTs for bilingual participants is shown in Figure 2. Symbol shape reflects listener group, and fill indicates masker condition, as defined in the legend. Lines indicate linear trends within group and condition. The BLP dominance score was negative for all late bilinguals, reflecting the fact that these listeners rated themselves more proficient in Mandarin than English. In contrast, dominance scores for all early bilingual participants were positive, indicating greater self-rated proficiency in English than Mandarin. Overall, higher BLP scores were associated with lower SRTs in both masker conditions, as predicted if greater English dominance supports masked English speech recognition at a lower SNR.

A linear model with SRT as the outcome measure evaluated the fixed effects of BLP score, listener group (early bilingual, late bilingual), and masker type (SSN, two-talker), as well as all two-way interactions. The interactions of Listener Group $\times$ Masker Type, $\text{coef} = -0.08$, $t(37) = -0.06$, $p = .954$, and BLP $\times$ Masker Type, $\text{coef} =$

Figure 2. The speech recognition threshold (SRT) as a function of Bilingual Language Profile (BLP) for bilingual participants. Symbol shape indicates listener group, which was either late bilinguals (diamonds) or early bilinguals (circles), and shading indicates the masker type, which was either two-talker speech (open) or speech-shaped noise (SSN; filled). Lines indicate linear trends for each group and masker type.

0.002, $t(37) = 0.23$, $p = .816$, were not significant. However, there was a nonsignificant trend for the interaction between BLP and listener group, $\text{coef} = 0.02$, $t(36) = 1.78$, $p = .084$. There was a significant main effect of BLP score, $\text{coef} = -0.03$, $t(36) = -3.05$, $p = .004$, and masker type, $\text{coef} = 4.28$, $t(37) = 6.90$, $p < .001$. However, no significant main effect of listener group was observed, $\text{coef} = -0.65$, $t(36) = -0.66$, $p = .513$. Nevertheless, there appears to be a weaker association between BLP and SRT for the early bilinguals than the late bilinguals.

For the late bilinguals, thresholds were predicted by the age at which the person moved to the United States. Note that we do not use the phrase “U.S. immigration,” as not all of the participants were seeking permanent U.S. residency. When predicting threshold as the outcome measure, masker type was highly significant (with two-talker speech resulting in poorer thresholds than SSN), so masker type remained in this model. When “age moved to the United States” was added to the model, “age moved” was significant, coefficient = 0.27, $t(37) = 3.20$, $p = .003$. Better thresholds were observed for people who moved to the United States at a younger age. It should be noted that, for our late bilinguals, “age moved to the United States” was significantly correlated with “duration of U.S. residency” ($r = -.59$, $p = .006$). Therefore, for this data set, “duration of U.S. residency” also predicted SRT (however, see Rimikis et al., 2013, for data supporting a larger effect of “age moved to the United States” than “duration of U.S. residency” on masked sentence recognition; see also Flege, 2019, for discussion of this topic).

Discussion

Overview of Basic Issues and Results

This study evaluated speech-in-noise and speech-in-speech recognition for monolingual, early bilingual, and late bilingual young adults with normal hearing. The late bilingual group performed more poorly than the monolingual group, consistent with a large number of previous studies (e.g., Mayo et al., 1997; Shi, 2010). In contrast to late bilinguals, the early bilingual group performed like the monolingual group. This replicates the results of some previous studies (Calandruccio & Zhou, 2014; Reetzke et al., 2016; Shi, 2009) and is inconsistent with a bilingual disadvantage that is sometimes reported (Krizman et al., 2016).

Several aspects of the results from this study suggest that the difference between performance of early and late bilinguals is due to differences in their English language skills. For the two groups of bilingual listeners, there was a negative correlation between the BLP English dominance score and the SRTs. There was some evidence of a trend for this association to be stronger in the data of late than early bilinguals, but this effect did not reach significance. Additional evidence of a stronger association between English language abilities and speech recognition in late than early bilinguals comes from the analysis of SRT for

low- and high-frequency words. Late bilinguals had higher SRTs in noise for low- than high-frequency words, but no such difference was observed for the other two groups. These results suggest that masked speech recognition may require some minimum familiarity with the target words, but that additional familiarity may not impact performance. We also observed that, for our late bilinguals, SRTs were positively correlated with the age at which they moved to the United States and were first immersed in an English-speaking country (i.e., lower thresholds predicted by a younger age of U.S. entry).

Bilingual Disadvantage Versus Frequency Effect

Several different explanations have been offered for poorer masked speech in bilingual speakers of the test language. The process of word recognition is often described using the neighborhood activation model (Luce & Pisoni, 1998). In this model, the acoustic-phonetic patterns characterizing words in the listener's lexicon are activated by auditory input. Familiar words are more easily activated than less familiar words, and word recognition is determined based on the node in the model with the highest activation. Understood in this way, a bilingual disadvantage might occur due to the larger number of words that the listener knows, including words in both languages (Colomé, 2001; Marian & Spivey, 2003). This effect would persist regardless of proficiency in the target language.

Another mechanism that could produce poorer masked speech recognition in bilingual listeners is based on their more limited exposure to words in the target language (Schmidtke, 2016). Whereas monolingual listeners are exposed to linguistic input in a variety of contexts (e.g., with reverberation) and with a variety of productions (e.g., accents), bilingual listeners' exposure is inherently more limited due to the fact that they are sharing time between two languages. This could affect the accuracy of the key acoustic-phonetic features of words in the listener's lexicon, as well as the baseline activation.

Results of this study support the idea that there is not an inherent disadvantage associated with bilingualism. The observation that early bilinguals performed comparably to monolinguals runs counter to the idea that competing activation associated with bilingualism limits performance. The association between English dominance and SRTs is more consistent with the idea that language skills and/or familiarity of the target speech is the critical factor affecting performance. The lower thresholds for high- than low-frequency words observed for late bilinguals provides additional support for the idea that familiarity is the critical factor responsible for degraded masked speech recognition in some bilingual listeners. In this data set, no significant interaction was found between listener groups and the two different types of maskers used (steady-noise masker vs. two-talker masker). Although it has been suggested that late bilinguals are more adversely affected by complex speech-based maskers than steady-noise maskers, our data did not support this conclusion.

Less familiarity with the language, combined with reduced spectral and temporal information of the speech signal in the context of a background masker, could be the dominant factor determining masked speech recognition for late bilingual listeners. One way to test this idea would be to compare performance across stimuli that vary in linguistic complexity. Linguistically complex stimuli should be associated with higher SRTs than linguistically simple and familiar stimuli, and if less familiarity with the language is a dominant factor in predicting performance, this difference should be larger for bilingual listeners than monolingual listeners. The rationale for using linguistically simple BKB stimuli in the current experiment was twofold—to evaluate performance with clinically available materials and to ensure that the lexical content was familiar to all listeners.

Listener Variables Other Than Age of Acquisition

Although it is not clear why some studies have found differences between monolingual and early bilingual participants (Krizman et al., 2017; Neave-Di Toro et al., 2017) and others have not (Calandruccio & Zhou, 2014; Reetzke et al., 2016; Shi, 2009), there are some listener variables that could explain this discrepancy. These factors include, but are not limited to, listener age, access to education, and daily language use. For example, Krizman et al. (2017) studied English-speaking monolingual and Spanish–English bilingual adolescents (14–15 years), with > 50% of the listeners in each group reporting low maternal education, which is an often-used proxy for socioeconomic status. In contrast, listeners in this study were older (18–31 years) and had access to high-level educational resources. The bilingual groups tested in this study were no longer living with their families of origin and reported only using Mandarin 13% (early bilinguals) and 36% (late bilinguals) of the time. Although not reported by Krizman et al. (2017), it is likely that the bilingual adolescents in their study were living with their families and used Spanish more regularly. It is possible that the discrepancy in results across studies was due to some combination of listener age, access to educational resources, and daily language use.

Language Specific Factors

This study assessed the difference in performance in a speech recognition task between English monolinguals and Mandarin–English bilinguals. Although it seems likely that these results would generalize to other bilingual populations, language-specific factors cannot be ruled out. While some studies have evaluated bilinguals with diverse language backgrounds (with respect to experience and exposure), others have tested bilinguals with more uniform language histories. For example, studies of more homogenous bilingual populations showing a difference have evaluated English–Spanish bilinguals (Krizman et al., 2017; Rogers et al., 2006), and studies failing to show a difference have evaluated English–Greek bilinguals

(Calandruccio & Zhou, 2014) and English–Mandarin bilinguals (this study).

Due to their common Latin roots, English and Spanish share a large number of cognates—words with shared meaning that are phonemically similar—as well as similarities in syntax and other features. In contrast, there are fewer cognates and structural similarities between English and Mandarin or Greek. Although the predominance of English/Spanish cognates facilitates vocabulary acquisition (Perez, Peña, & Bedore, 2010) and verbal fluency (Blumenfeld, Bobb, & Marian, 2016), it is unclear whether word-level similarities could increase confusion between English and Spanish in the context of a masked sentence recognition task. The finding of a bilingual disadvantage for English–Spanish bilinguals, but not for English–Mandarin or English–Greek bilinguals, is broadly consistent with an explanation based on the prevalence of cognates.

Spanish, Greek, and Mandarin also differ with respect to their acoustic–phonetic similarity with English, and it is possible these features could play a role in masked speech recognition. For example, extensive overlap between the phonemes in a bilingual listener’s two languages could increase competition in the activation of the two lexicons. The percentage of phonemes shared with English is 78% (Spanish), 48% (Greek), and 41% (Mandarin; American Speech-Language-Hearing Association, n.d.-a, n.d.-b, n.d.-c, n.d.-d; Mennen & Okalidou, 2007). The observation that Spanish has greater phonemic overlap with English than Greek or Mandarin is broadly consistent with the idea that acoustic–phonetic similarity between a bilingual’s two languages reduces masked speech recognition. However, more research is needed to assess this possibility.

Practical Implications

The results of this study found poorer masked speech recognition for late bilinguals than early bilinguals or monolinguals, but no evidence of a bilingual disadvantage for highly proficient early bilinguals tested with linguistically simple, clinically common stimuli. Results also indicate that bilingual students in university settings may struggle to understand masked speech even if they are highly proficient in the target language but have learned English later in life. This result is consistent with that of Neave-Di Toro et al. (2017). That study evaluated masked speech recognition multitalker babble, with and without reverberation, in university students who were either native English speakers or late learners of English. Late bilinguals performed more poorly than native speakers, and they were more susceptible to reverberation (see also Shi, 2010). Those results and results from this study support the importance of maintaining a quiet and nonreverberant environment to facilitate bilingual students’ speech recognition in classrooms across education levels, including university.

The present results are also pertinent to speech perception testing in a clinical setting. Masked speech recognition is an important component of an audiological

evaluation, candidacy for various interventions (e.g., cochlear implantation), demonstrating need for services (e.g., an FM system at school), and efficacy of treatment (e.g., hearing aid outcomes). For patients who are not proficient in English, language-appropriate materials are needed for valid assessment, as well as an audiologist with sufficient proficiency in the target language to score patient responses appropriately (Polley, 2010). This is a significant limitation for many clinics in the United States, given the severe shortage of bilingual audiologists (American Speech-Language-Hearing Association, 2017). Results of this study suggest that using English language materials and test norms is appropriate for early bilingual patients who are highly proficient in the test language and spent their life in an environment in which they had to continuously use English daily.

In contrast to early bilinguals, even highly proficient late bilinguals are at a disadvantage when recognizing speech in background noise. This poses a serious challenge for speech perception testing, where poor performance could reflect effects of language history, hearing loss, or some combination. This leaves the question: How do we develop a full auditory–linguistic profile for our bilingual patients who learned English later in life? In many cases, patients may use their native language on a very limited basis at the time of clinical intervention but continue to underperform compared to monolingual English speakers on masked speech recognition. In the current data, late bilinguals reported only using Mandarin approximately 36% of the time and only had an average age of 22 years. By the time many of these listeners need audiological intervention (e.g., due to presbycusis), they may use Mandarin for an even smaller percentage of the day if they continue to reside in the United States. In this scenario, it might be advisable to test them in English to get an accurate estimate of their functional auditory difficulties. However, test materials (including nonlinguistic stimuli; e.g., see Archer-Boyd, Southwell, Deeks, Turner, & Carlyon, 2018; Aronoff & Landsberger, 2013) and speech perception models need to be developed that can parse out difficulties due to hearing loss versus those caused by limited familiarity of the test language.

Limitations

A potential limitation of this study—and one reason we might not have seen a difference between our monolingual and early bilingual groups—is that both groups lived their entire lives immersed in a U.S. population that is predominantly English speaking. Though this would be true for many early bilinguals living in the United States, these results may not be representative of the broader population of early bilinguals who use both of their languages to a similar extent (e.g., in dual-language countries). Furthermore, if the task had been more difficult by, for example, adding reverberation or using linguistically more complex speech material, it is possible that differences between the groups might have emerged. Lastly, this study

evaluated a “best case” scenario, where all of our listeners had access to high levels of education. Perhaps some of the inconsistencies for masked speech recognition reported in the literature between monolinguals and early bilinguals have to do with differences in education across populations studied.

Conclusion

There are many advantages associated with bilingualism, from greater employment opportunities to enhanced frequency discrimination (Skoe, Burakiewicz, Figueiredo, & Hardin, 2017). As of 2013, there were over 60 million people over the age of 5 years in the United States who spoke a language other than English at home, and this number continues to grow (U.S. Census Bureau, Population Division, 2008). It is essential that the literature addressing bilingualism in audiology captures the needs of this particular population. The data obtained in this study demonstrate important differences in performance between monolingual and late bilingual listeners, even for a group of college students who are highly educated and successfully communicating in the target language on a daily basis. The correlation between English speech recognition scores and English dominance, as well as the effect of word frequency on performance of late bilinguals, suggests that performance reflects language ability and familiarity rather than active interference between the bilingual listener’s two languages. In contrast, early bilinguals performed similarly to their monolingual peers. The bilingual experience is very different from one individual to the next based on their surroundings, and this diversity of experience is reflected in their speech recognition performance. These results argue against the existence of a consistent bilingual disadvantage.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grants R03 DC 015074 (Principal Investigator: Lauren Calandruccio) and R01 DC 015056 (Co-Principal Investigator: Emily Buss). Jacob Oleson provided advice about the statistical analysis.

References

- Alladi, S., Bak, T. H., Mekala, S., Rajan, A., Chaudhuri, J. R., Mioshi, E., . . . Kaul, S. (2016). Impact of bilingualism on cognitive outcome after stroke. *Stroke*, 47, 258–261.
- American National Standards Institute, & Acoustical Society of America. (2009). *Methods for manual pure-tone threshold audiometry*. S3.21-2004.
- American Speech-Language-Hearing Association. (n.d.-a). *Arabic phonemic inventory*. Retrieved from <http://www.asha.org/uploaded/files/practice/multicultural/arabicphonemicinventory.pdf>
- American Speech-Language Hearing Association. (n.d.-b). *English phonemic inventory*. Retrieved from <http://www.asha.org/uploadedFiles/practice/multicultural/EnglishPhonemicInventory.pdf>
- American Speech-Language-Hearing Association. (n.d.-c). *Mandarin phonemic inventory*. Retrieved from <http://www.asha.org/>

- uploadedFiles/practice/multicultural/MandarinPhonemicInventory.pdf
- American Speech-Language-Hearing Association.** (n.d.-d). *Spanish phonemic inventory*. Retrieved from <http://www.asha.org/uploadedFiles/Spanish-Phonemic-Inventory.pdf>
- American Speech-Language-Hearing Association.** (2017). *Demographic profile of ASHA members providing bilingual services*. Retrieved from <http://www.asha.org/practice/multicultural/Resources-Recruitment-Bilingual-Service-Providers/>
- Archer-Boyd, A. W., Southwell, R. V., Deeks, J. M., Turner, R. E., & Carlyon, R. P.** (2018). Development and validation of a spectro-temporal processing test for cochlear-implant listeners. *The Journal of the Acoustical Society of America*, 144, 2983.
- Aronoff, J. M., & Landsberger, D. M.** (2013). The development of a modified spectral ripple test. *The Journal of the Acoustical Society of America*, 134(2), EL217–EL222.
- Bench, J., Kowal, Á., & Bamford, J.** (1979). The BKB (Bamford–Kowal–Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, 13, 108–112.
- Bialystok, E.** (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives*, 9, 117–121.
- Birdsong, D., Gertken, L. M., & Amengual, M.** (2012). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. COERLL/University of Texas at Austin. Retrieved from <http://sites.la.utexas.edu/bilingual>
- Blumenfeld, H. K., Bobb, S. C., & Marian, V.** (2016). The role of language proficiency, cognate status and word frequency in the assessment of Spanish–English bilinguals’ verbal fluency. *International Journal of Speech-Language Pathology*, 18, 190–201.
- Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R.** (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, 110, 2527–2538.
- Camarota, S. A., & Zeigler, K.** (2014, October, 3). One in five U.S. residents speaks foreign language at home, record 61.8 million. Center for Immigration Studies. Retrieved from <http://cis.org/One-Five-US-Residents-Speaks-Foreign-Language-Home-Record-618-million>
- Calandruccio, L., & Buss, E.** (2017). Spectral integration of English speech for non-native English speakers. *The Journal of the Acoustical Society of America*, 142(3), 1646–1654.
- Calandruccio, L., & Smiljanić, R.** (2017). New sentence recognition materials developed using a basic non-native English lexicon. *Journal of Speech, Language, and Hearing Research*, 55(5), 1342–1355. [https://doi.org/10.1044/1092-4388\(2012/11-0260\)](https://doi.org/10.1044/1092-4388(2012/11-0260))
- Calandruccio, L., & Zhou, H.** (2014). Increase in speech recognition due to linguistic mismatch between target and masker speech: Monolingual and simultaneous bilingual performance. *Journal of Speech, Language, and Hearing Research*, 57(3), 1089–1097.
- Calvo, A., & Bialystok, E.** (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, 130, 278–288.
- Carhart, R., Johnson, C., & Goodman, J.** (1975). Perceptual masking of spondee by combinations of talkers. *The Journal of the Acoustical Society of America*, 58, S35.
- Carlson, S. M., & Meltzoff, A. N.** (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, 11, 282–298.
- Colomé, Á.** (2001). Lexical activation in bilinguals’ speech production: Language-specific or language-independent? *Journal of Memory and Language*, 45, 721–736.
- Cooke, M., García Lecumberri, M. L., & Barker, J.** (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America*, 123, 414–427.
- Corbin, N., Bonino, A. Y., Buss, E., & Leibold, L. J.** (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear and Hearing*, 37, 55–63.
- Crandell, C. C., & Smaldino, J. J.** (1996). Speech perception in noise by children for whom English is a second language. *American Journal of Audiology*, 5, 47–51.
- Davis, C. J.** (2005). N-watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, 37(1), 65–70.
- Flege, J.** (2019). A non-critical period for second-language learning. In *A sound approach to language matters: In honor of Ocke-Schwen Bohn*, Aarhus University. Open access e-book at Aarhus University Library.
- García 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, 2445–2454.
- Grosjean, F.** (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1(2), 131–149.
- Guzmán-Vélez, E., & Tranel, D.** (2015). Does bilingualism contribute to cognitive reserve? Cognitive and neural perspectives. *Neuropsychology*, 29, 139–150.
- Hawkins, J. E., & Stevens, S. S.** (1950). The masking of pure tones and of speech by white noise. *The Journal of the Acoustical Society of America*, 22, 6–13.
- Janus, M., & Bialystok, E.** (2018). Working memory with emotional distraction in monolingual and bilingual children. *Frontiers in Psychology*, 9, 1582.
- Kilman, L., Zekveld, A., Hällgren, M., & Rönneberg, J.** (2014). The influence of non-native language proficiency on speech perception performance. *Frontiers in Psychology*, 5, 651.
- Kraus, N., & White-Schwoch, T.** (2017). The bilingualism paradox. *The Hearing Journal*, 70(1), 40–42.
- Krizman, J., Bradlow, A. R., Lam, S. S.-Y., & Kraus, N.** (2017). How bilinguals listen in noise: Linguistic and non-linguistic factors. *Bilingualism: Language and Cognition*, 20, 834–843.
- Krizman, J., Skoe, E., & Kraus, N.** (2016). Bilingual enhancements have no socioeconomic boundaries. *Developmental Science*, 19, 881–891.
- Luce, P. A., & Pisoni, D. B.** (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1–36.
- Marian, V., & Spivey, M.** (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, 6, 97–115.
- 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.
- Meador, D., Flege, J. E., & Mackay, I. A.** (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition*, 3, 55–67.
- Mennen, I., & Okalidou, A.** (2007). Greek speech acquisition. In S. McLeod (Ed.), *The international guide to speech acquisition*, (pp. 398–401). Clifton Park, NY: Thomson Delmar Learning.
- Motlagh Zadeh, L., Jalilvand Karimi, L., & Silbert, N. H.** (2018). Bilingualism leads to greater auditory capacity. *International Journal of Audiology*, 57, 831–837.
- Neave-Di Toro, D., Rubinstein, A., & Neuman, A. C.** (2017). Speech recognition in nonnative versus native English-speaking

- college students in a virtual classroom. *Journal of the American Academy of Audiology*, 28(5), 404–414.
- Perez, A. M., Peña, E. D., & Bedore, L. M.** (2010). Cognates facilitate word recognition in young Spanish–English bilinguals’ test performance. *Early Child Services (San Diego)*, 4, 55–67.
- Persici, V., Vihman, M., Burro, R., & Majorano, M.** (2019). Lexical access and competition in bilingual children: The role of proficiency and the lexical similarity of the two languages. *Journal of Experimental Child Psychology*, 179, 103–125.
- Polley, K. B.** (2010). *Accuracy of English speakers administering word recognition score tests in Mandarin* (Dissertation). Brigham Young University, Provo, UT.
- Reetzke, R., Lam, B. P., Xie, Z., Sheng, L., & Chandrasekaran, B.** (2016). Effect of simultaneous bilingualism on speech intelligibility across different masker types, modalities, and signal-to-noise ratios in school-age children. *PLOS ONE*, 11, e0168048.
- Rimikis, S., Smiljanić, R., & Calandruccio, L.** (2013). Nonnative English speaker performance on the Basic English Lexicon (BEL) sentences. *Journal of Speech, Language, and Hearing Research*, 56, 792–804.
- Rivera Mindt, M., Arentoft, A., Kubo Germano, K., D’Aquila, E., Scheiner, D., Pizzirusso, M., . . . Gollan, T. H.** (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review*, 18, 255–268.
- 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.
- Schmidtke, J.** (2016). The bilingual disadvantage in speech understanding in noise is likely a frequency effect related to reduced language exposure. *Frontiers in Psychology*, 7, 678.
- Schneider, B., Martinez, S., & Owens, A.** (2006). Barriers to educational opportunities for Hispanics in the United States. In M. Tienda & F. Mitchell (Eds.), *Hispanics and the future of America* (pp. 179–227). Washington, DC: National Academies Press.
- Shi, L.** (2009). Normal-hearing English-as-a-second-language listeners’ recognition of English words in competing signals. *International Journal of Audiology*, 48, 260–270.
- Shi, L.** (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of Speech, Language, and Hearing Research*, 53, 821–835.
- Shi, L.** (2014). Validating models of clinical word recognition tests for Spanish/English bilinguals. *Journal of Speech, Language, and Hearing Research*, 57, 1896–1907.
- Skoe, E., Burakiewicz, E., Figueiredo, M., & Hardin, M.** (2017). Basic neural processing of sound in adults is influenced by bilingual experience. *Neuroscience*, 349, 278–290.
- Tabri, D., Abou Chacra, K. M., & Pring, T.** (2011). Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language & Communication Disorders*, 46, 411–422.
- Treccani, B., Argyri, E., Sorace, A., & Sala, S. D.** (2009). Spatial negative priming in bilingualism. *Psychonomic Bulletin & Review*, 16, 320–327.
- U.S. Census Bureau, Population Division.** (2008). *Projections of the population by sex, race, and Hispanic origin for the United States: 2010 to 2050*. Retrieved from <https://www.census.gov/data/tables/2008/demo/popproj/2008-summary-tables.html>
- von Hapsburg, D., Champlin, C. A., & Shetty, S. R.** (2004). Reception thresholds for sentences in bilingual (Spanish/English) and monolingual (English) listeners. *Journal of the American Academy of Audiology*, 15, 88–98.
- von Hapsburg, D., & Peña, E. D.** (2002). Understanding bilingualism and its impact on speech audiometry. *Journal of Speech, Language, and Hearing Research*, 45(1), 202–213.
- 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(1), 5–17.
- Wightman, F. L., & Kistler, D. J.** (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *The Journal of the Acoustical Society of America*, 118, 3164–3176.