Journal of the
American Academy of Audiology

Bilingualism and Speech Understanding in Noise: Auditory and Linguistic Factors

Journal:	*Journal of the American Academy of Audiology*
Manuscript ID	17-082.R2
Manuscript Type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Skoe, Erika; University of Connecticut, Speech, Language, and Hearing Sciences Karayanidi, Kateryna; University of Connecticut, Speech, Language, and Hearing Sciences
Keywords:	Speech perception, bilingual, speech in noise

SCHOLARONE™
Manuscripts

For Review

FINAL PREPRINT PRIOR TO PUBLICATION

Skoe, E., & Karayanidi, K. (2018). Bilingualism and speech understanding in noise: Auditory and linguistic factors. Journal of the American Academy of Audiology.

doi:10.3766/jaaa.17082.

Bilingualism and Speech Understanding in Noise:

Auditory and Linguistic Factors

Erika Skoe[1,2,3†,] and Kateryna Karayanidi [1,3]

Affiliations:

[1]Department of Speech, Language, and Hearing Sciences; [2]Department of Psychological Sciences, Cognitive Sciences Program, Connecticut Institute for Brain and Cognitive Sciences; [3]University of Connecticut, Storrs, CT 06269 USA

†Corresponding author:

Erika Skoe, Ph.D.

University of Connecticut

850 Bolton Rd, U-1085

Storrs, CT 06129 USA

860-486-3685

Email: erika.skoe@uconn.edu

Submission date: 11/8/2017

For Peer Review

Abstract

Background: Bilingual speakers often have difficulty understanding speech in noisy and acoustically-degraded conditions.

Purpose: The first aim was to examine the potential source(s) of the difficulties that English-proficient bilingual listeners experience when hearing English speech in noise. The second aim was to assess how bilingual listeners perform on a battery of central auditory processing tests.

Research Design: A mixed design was used in this study.

Study Sample: Normal-hearing college students (n=24) participated in this study. The bilingual participants (n=12) self-reported that they learned a second language before age 9 and the monolingual participants reported that they only knew American English. All participants considered themselves to be native speakers of American English.

Data Collection and Analysis: Participants were administered the Revised Speech in Noise (R-SPIN) test to assess whether bilingual listeners' speech understanding in noise reflects auditory factors, linguistic factors, or a combination of the two. To minimize the influence of short-term memory and motor movements, only the final word of a sentence is repeated for this test. Sentence-final words were presented in two linguistic contexts: in the high predictability condition, the final word can be deduced from the context created by the preceding words, and in the low predictability condition it cannot. The R-SPIN test was administered at two signal-to-noise ratios (SNR) (0 dB, 3 dB). In addition, participants were given a reading comprehension test to measure their ability to use context when linguistic stimuli are delivered to the visual, not auditory, modality. The central auditory test battery consisted of three tests: Competing

For Peer Review

Sentences, Dichotic Digits, NU-6 Time-Compressed Speech with Reverberation. All test materials were given in American English.

Results: The bilingual and monolingual groups performed similarly in the low context condition of the R-SPIN test. However, in comparison to the age-matched monolingual group, the bilingual group did not derive the same level of benefit from contextual cues, as seen by a smaller improvement in performance between the low and high predictability R-SPIN conditions. The bilingual and monolingual groups showed a similar decrement in performance when the SNR dropped. In addition, bilingual individuals underperformed on the Competing Sentences test, when instructed to attend to the left ear. However, the bilingual and monolingual groups performed equally well on the reading comprehension test, as well as on the Time-Compressed Speech with Reverberation Test and Dichotic Digits.

Conclusions: We show that individuals who are exposed to two languages from an early age, and self-report as having a high level of proficiency in English, perform like their monolingual counterparts in acoustically-degraded conditions where context is not facilitative, but they underperform in conditions where sentence-level linguistic context is facilitative to understanding. We conclude that deficits observed in noise are likely not due to a perceptual deficit or a lack of linguistic competence, but instead reflect a linguistic system that performs inefficiently in noise. In addition, we do not find evidence of an auditory processing weakness or advantage in our bilingual cohort; however, the use of speech materials to assess auditory processing is a confound.

Key Words: bilingual, speech in noise

For Peer Review

Abbreviations:

AoA:	age of acquisition
ANL:	Acceptable Noise Level
CAP:	Central Auditory Processing
CAPD:	Central Auditory Processing Disorder
CST:	Competing Sentences Test
DDT:	Dichotic Digits Test
DPOAE:	distortion product otoacoustic emissions
HINT:	Hearing in Noise Test
L1:	first language
L2:	second language
MLV:	monitored live voice
NEL:	non-English language
PTA	pure tone average
R-SPIN:	Revised Speech Perception in Noise
SIN:	speech in noise
SL:	sensation level
SNR:	signal-to-noise ratio
SRT:	speech recognition threshold
SUN:	speech understanding in noise

For Peer Review

Introduction

There is a preponderance of evidence that bilingual speakers have more difficulty understanding speech in noisy and acoustically degraded conditions than their monolingual counterparts (Mayo, Florentine, and Buus 1997; Rogers et al. 2006; Bradlow and Alexander 2007; Shi 2010; Tabri, Abou Chacra, and Pring 2011; Shi 2012; Hervais-Adelman, Pefkou, and Golestani 2014; Krizman et al. 2016). This increased perceptual difficulty is akin to having a mild hearing loss, even for listeners with clinically normal audiometric thresholds. Lucks Mendel and Widner (2016) suggest that the bilingual disadvantage for speech in noise is the consequence of "auditory processing degradation", although other work suggests a bilingual advantage for low-level auditory processing (Krizman et al. 2016). Bilingualism, thus, provides an interesting test case for examining the relative roles of auditory versus linguistic contributions to speech understanding in noise. The current study explores the potential source(s) of the perceptual difficulties that bilingual individuals experience when listening to speech in noise, and we specifically focus on bilingual listeners who consider themselves to be proficient in the test language.

Speech-in-noise (SIN) testing is a routine part of audiological practice. SIN tests are an attractive clinical tool because they access the most common complaint that brings a patient to the audiologist in the first place, namely difficulty understanding speech in noisy backgrounds. The tests in most wide-scale use include the Hearing in Noise Test (HINT) (Nilsson, Soli, and Sullivan (1994)), the Words in Noise Test (Wilson, Abrams, and Pillion 2003), and the QuickSIN test (Killion et al. 2004) (For a review see, Lagacé, Jutras, and Gagné (2010)). Most of these tests can be administered in a matter of minutes and require the listener to repeat a single word or the entire sentence. The requirement to have the listener repeat what they heard as a way

For Peer Review

of indexing perceptual acuity is a complication to test interpretation because performance depends not only on the listener's auditory percept, but also on working memory, the ability to form and execute speech motor plans, top-down linguistic knowledge, among other variables. Most of these SIN tests, however, lack the specificity to isolate whether the listener's weakness is due to auditory factors, linguistic factors, cognitive, memory factors, motor factors (etc.) or some combination thereof. Understanding which factor, or set of factors contribute to decreased performance in noise is essential for providing appropriate clinical counseling as well compensatory and/or remediation strategies.

Lagacé, Jutras, and Gagné (2010) proposed using the Revised Speech in Noise (R-SPIN) test (Bilger et al. 1984) to evaluate whether decreased speech perception in noise is the result of weak auditory processing, weak language processing, or a combination of the two. It has been argued that both types of impairments can manifest in poor speech understanding in noise (SUN), although the underlying mechanisms are presumed to be different. To minimize the influence of short-term memory and motor movements, only the final word of a sentence is repeated in the R-SPIN test. In addition, the target words are presented as part of a design that uses two levels of linguistic predictability. In the high predictability condition, the final word can be deduced from the context created by the preceding words in the sentence, and in the low predictability condition it cannot. For example, "The lion gave an angry roar" (high predictability final word) vs. "He is thinking about the roar" (low predictability final word). The high and low predictability conditions can then each be presented with different levels of masking to manipulate the signal-to-noise ratio (SNR), creating low SNR and high SNR conditions. Comparing performance on the low and high SNR conditions can give insight into the auditory processes contributing to SUN. According to Lagacé, Jutras, and Gagné (2010), this

For Peer Review

factorial design, which manipulates both predictability and SNR, provides an method for dissociating auditory processing from linguistic processing. The authors posit that a listener with a "pure" auditory processing issue (i.e., an auditory processing problem without concomitant language problem) will perform more poorly than a typical listener as the SNR becomes less favorable; however, they will receive the same, if not potentially greater, benefit when the word can be deduced from the linguistic context. Partial evidence for this pattern of findings can be found in Lagacé and colleague's small scale study of children with central auditory processing disorder (CAPD) (Lagacé et al. 2011). Lagacé et al. (2011), argue that if reduced perception of speech in noise is the result of language-specific processes, the benefit from context will be small (if not at all) compared to typical listeners, but the listener will not be inordinately affected by changes in SNR. If an individual underperforms as the SNR decreases and they also do not benefit from context in the typical manner, this, Lagacé, Jutras, and Gagné (2010) argue, should be taken as evidence that both auditory and linguistic processes are contributing to poor perception of speech in noise. As a logical extension of the argument made by Lagacé, Jutras, and Gagné (2010), if a listener experiences less of a performance decrement than typical listeners when the SNR is decreased, this would be an indication of an auditory processing advantage.

We focused on bilingual adults who self-reported as being highly-proficient speakers of English. The bilingual college students in our sample learned English in combination with another language from a young age (before age 9), they were proficient at the native level in both their languages (self-report), were college students at an American university, and spoke English without a noticeable accent. Thus, the bilingual participants in our sample are likely to have been mistaken for monolingual English speakers during routine, daily communication by other native, monolingual speakers of English. For listeners who are still learning the test language, or

For Peer Review

who consider themselves to be non-native speakers of the test language, poor performance on SIN tests can be explained by an under-developed knowledge of the target language that results from reduced exposure to that language compared to monolingual speakers of that language. For some bilingual listeners, this reduced language exposure is the necessary consequence of having learned the target language at a later age (Shi 2010) and in most examples cited in the literature, underperformance on SIN tasks can be attributed to measurable differences in language proficiency, years of exposure, or the degree of balance between the two languages (Shi 2010, 2012). The current study asks whether such disadvantages persist even when the listener has achieved a high level of language proficiency in the target language. Schmidtke (2016) recently reported that bilingual speakers have poorer word recognition ability on a modified version of the R-SPIN, even after controlling for differences in verbal ability between the bilingual and monolingual subsamples. This lead to the hypothesis that underperformance on SIN tasks is an inevitable consequence of splitting resources across multiple languages (Schmidtke 2016).

The primary aim of the current study was to explore the auditory and/or linguistic basis of the bilingual disadvantage for speech understanding in noise (SUN) using the R-SPIN test. The second aim of our study was to assess the auditory processing skills of our bilingual listeners using a battery of tests that are commonly used in a clinical setting to diagnose CAPD. The CAPD test battery included Time-Compressed Speech with Reverberation, Dichotic Digits, and Competing Sentences. These tests, like the R-SPIN, utilize linguistic stimulation, but unlike the R-SPIN, they do not utilize background babble as a distractor. Krizman et al. (2016) recently hypothesized that bilingual listeners develop stronger auditory processing skills as a way to compensate for the challenges they face for SUN. This hypothesis emerged from a recent string of studies showing superior performance to monolingual listeners for processing auditory

For Peer Review

signals, including signals that are masked by noise. Krizman et al. (2016), reported that adolescents who learned two language from an early age had lower thresholds for simultaneous and backward masking tasks compared to monolingual counterparts. Montagni and Peru (2011) provide additional evidence that early exposure to a second language confers an advantage to auditory processing tasks across both linguistic and musical stimuli. Moreover, early exposure to two languages (for both children and adults) has also been associated with more robust (pre-attentive) neural processing of speech sounds in both quiet and background babble conditions (Krizman et al. 2012; Krizman et al. 2014; Krizman et al. 2015; Skoe et al. 2017).

In addition to this recent evidence of an auditory processing advantage in bilingual individuals, there is compelling evidence that bilingual individuals have stronger executive functions, in particular, stronger inhibition of task-irrelevant stimuli that are both auditory and visual in nature (Soveri et al. 2011; Krizman et al. 2012; Bak, Vega-Mendoza, and Sorace 2014; Bialystok 2015). For bilinguals, inhibitory control has been theorized to emerge as a byproduct of needing to suppress one language when the other is the target language (Green 1998), although a more modern account is that increased inhibitory control is the result of needing to monitor which language to produce in different communication settings (Costa et al. 2009). There is a small body of literature suggesting that increased inhibitory control contributes to heightened dichotic processing in bilingual listeners (Soveri et al. 2011; Gresele et al. 2013). This literature predicts that bilingual listeners should outperform monolingual listeners on tests of dichotic listening, such as Competing Sentences and Dichotic Digits. When listening to dichotic speech stimulation, there is a bias towards listening to the right ear, even when instructed to attend to the left ear. This right-ear bias, which is well described in the scientific literature, is presumed to be the outcome of the right ear having a more direct pathway to speech-

For Peer Review

specialized regions in the left temporal lobe than the left ear. As a consequence of this circuitry, attending to the left ear is theorized to require more executive processing than attending to the right ear under dichotic stimulation (reviewed in Hugdahl et al. (2009). By this argument, heightened executive function is expected to boost bilingual listeners' ability to attend to the left ear. However, the whole notion that bilinguals have advantages in executive function has recently been called into question (Paap, Johnson, and Sawi 2014, 2015). In addition, because CAP tests are intended to assess impaired (not extraordinary) auditory processing, and because they typically utilize speech materials, they may lack the sensitivity to observe bilingual advantages for auditory processes and/or executive function.

Methods

Participants

The study included 12 monolingual speakers (9 females) and 12 bilingual speakers (9 females), all students at the University of Connecticut with a negative history of hearing impairment and negative history of speech or language pathology. All procedures were approved by the University of Connecticut Institutional Review Board. Participants gave their written informed consent to participate and they were either paid or compensated through course extra credit for their participation (their choice). Consent, and all testing materials and instructions were delivered in American English. Testing was conducted in a 1.5-2 hour session, with breaks given between tests. All testing was administered by the second author, an English-Russian bilingual, who learned English at age 11. The potential confounds of having a bilingual test administrator are addressed in the Clinical Implications section of the Discussion.

For Peer Review

Participants completed a survey of their bilingual background and language exposure. The survey was a modified version of the survey developed by Garcia-Sierra et al. (2012). On the survey, participants rated their ability to use English and all other languages that they knew, using a Likert scale from 1-10, with 10 being labeled as "expert". To validate the consistency of their ratings, at a later point in the survey, the participants were asked to indicate whether or not they considered themselves to have native-like proficiency for English and each non-English language (NEL). With respect to English, all participants rated themselves as 9 or 10, and described their proficiency as "native-like". The survey also included questions about confidence reading in their NEL. In addition, participants were given a musical training questionnaire adapted from the one created by Kraus and colleagues (Slater and Kraus 2016), because of the literature showing an association between musical training and SUN advantages (Parbery-Clark et al. 2009; Bak, Vega-Mendoza, and Sorace 2014; Slater et al. 2015).

On the bilingual background survey, participants were instructed to indicate the degree to which they were exposed to English vs. their NEL at different points in their life, broken down in increments of three years (i.e. 0-3, 3-6, 6-9, etc.), using a rating of 0, 25, 50, 75, or 100%, with 100% indicating exposure to NEL only and 0% indicating exposure to English only. The monolinguals indicated that they were not exposed to a language other than English during their day-to-day communication at any point in their life. For all but four bilingual participants, exposure to English and the NEL began during the first three years of life. For the remaining four, English was learned as the second language after age 3 but before age 9. In all cases, the non-English language (NEL) was spoken by one or both of their parents. For two of the bilingual participants, one parent was a native speaker of English and the other was a native speaker of the NEL. At the time of testing, average exposure to the NEL was 29.1% (SD =

For Peer Review

14.43%). In terms of language use, all participants were English dominant at the time of testing. The non-English languages included Bangla, Japanese, Mongolian, Polish, Portuguese (x2), Serbian, Spanish (x3), Tamil, and Telugu (Table 1).

The bilingual and monolingual groups were matched with respect to age, self-rated English proficiency, bilateral pure tone averages (0.5, 1, 2 kHz), as well as maternal education, a commonly used index of socio-economic status ($p>0.05$ for all comparisons) (Table 2). In addition, the groups were matched with respect to the number of years of musical training (~3.4 years on average for the monolingual group and ~4.5 years on average for the bilingual group) (Table 2). In the case of the bilingual participants, all but one reported having received voice or instrumental training in the past but none were currently active in music activities. For the monolingual group, two were still active in music activities, and five reported never having received any musical training in the past.

[Table 1]

[Table 2]

For Peer Review

Test Battery

All testing was completed in a double-walled sound booth (IAC Acoustics) in the Auditory Brain Research Laboratory at the University of Connecticut. Prior to any SIN testing, participants were first verified to have normal otoscopy, normal bilateral air conduction thresholds < 25 dB HL for octaves 125 to 8000 Hz (GSI 61 audiometer), and normal outer hair cell function as confirmed by a distortion product otoacoustic emissions (DPOAEs) screening protocol performed using a handheld screener (Madsen Alpha OAE+ Screener, GN Otometrics). Speech recognition thresholds (SRTs) were obtained using the modified Hughson-Westlake method for the right and left ears (separately) via monitored live voice, after first familiarizing the participants with the spondee words: ice cream, baseball, toothbrush, airplane, outside, mushroom and sunshine. Binaural SRTs were obtained in the same manner, except spondees were presented to both ears at the same time via monitored live voice. The test order was right SRT, left SRT, and then binaural SRT. All subsequent test materials were delivered relative to the SRT; for binaural tests, like the R-SPIN, binaural SRTs were used. The tests were administered in the following order: Competing sentences (right ear first), Dichotic Digits, NU-6 Time-Compressed Speech with Reverberation (right ear first), R-SPIN, and then the Passage Comprehension test (Woodcock-Johnson Mastery Tests of Achievement IIII, WRMT-III). The Competing Sentences, Dichotic Digits, and Time-Compressed Speech tests are distributed by Auditech, Inc. and test administration followed recommended guidelines. Test materials were delivered from a desktop computer routed through a two-channel GSI 61 audiometer to ER-2 insert earphones.

Competing Sentences Test (CST) (Willeford 1978)

For Peer Review

This test of binaural separation includes 20 pairs of simple sentences spoken by a man, with the sentences being six to seven words in length. For each pair, one sentence is presented to the right ear and the other is presented simultaneously to the left ear. The two sentences in the pair have a similar theme. For illustrative purposes, here are two example pairs: (1) "I was late today" and "This watch keeps good time" and (2) "We had to repair the car" and "We usually take a taxi". The participant was instructed to listen and repeat back the sentence presented to one ear while ignoring the sentence presented to the other. For the first 10 sentences, the participant was instructed to repeat back the entire sentence presented to the right ear and for the final set of 10 sentences they were instructed to repeat back the sentence presented to the left ear. The target sentence was presented at 35 dB SL (re: SRT) and the competing sentence was presented at 50 dB SL (re: SRT). Each sentence is worth 10 points (2.5 per word), and there are 10 sentences, yielding a total possible maximum score of 100. For the purposes of assessing central auditory processing disorder, a score <90% for the right ear and < 90% for the left ear is considered abnormal for adults (11+ years).

Dichotic Digits Test (DDT) (Musiek 1983)

In this test of binaural integration, the participant repeats what s/he heard in both ears. For each of the 20 trials, two digits are presented to each ear (80 total digits, 40 per ear) at 50 dB SL (re: SRT). The digits include monosyllabic numbers between 1 and 10 (i.e., all numbers except 7). The participant is instructed to listen to both ears and repeat the numbers without concern about the order. Participants are encouraged to guess if they were not sure of what they heard. Prior to administering the test material, the participant is given three practice trials. To score the test, the number of correctly repeated digits is totaled for each ear separately (40 points per ear) and converted to a percentage. For the purposes of assessing central auditory processing disorder, a

For Peer Review

score <90% for the right ear or <90% for the left is considered abnormal for adults. Like Competing Sentences, a right ear advantage is expected. Compared to the other tests that were administered, this has a relatively light linguistic load.

NU-6 Time-Compressed Speech with Reverberation (Wilson et al. 1994)

In this monaural test, the speech materials, spoken by a female, are time-compressed (45%) with 0.3 seconds of reverberation. Participants are told that they will hear a woman's voice and that she will sound as if she is in a gymnasium. They are instructed to verbally repeat the word that they hear the woman say. Each sentence starts with the carrier phase "Say the word _____", with the final word being drawn from the NU-6 list of words. In this test, the final word cannot be derived from the preceding word. For each sentence, the listener must repeat back the final word that they heard, and guess if they are uncertain. Sentences were presented at 50 dB SL (re: SRT), with 50 target words per ear, starting with the right ear. The test is scored based on the number of correctly repeated words (50 points/ear). For the purposes of assessing central auditory processing disorder, a score <35% for the right ear or <35% for the left is considered abnormal. NU-6 List 5 was administered to the right ear and NU-6 List 6 was administered to the left ear.

Revised Speech Perception in Noise Test (R-SPIN) (Bilger et al. 1984).

For the R-SPIN test, participants are told that they will hear a man say a sentence in an environment that sounds as if he is at a party. They are instructed to repeat back only the last word of the sentence that the man says and to guess if they are uncertain. All sentences are syntactically correct, and, in all cases, the final word is a monosyllabic noun; however, in half of the sentences the final word is predictable from context. The final word predictability is pseudo-

For Peer Review

randomly varied from sentence to sentence. The test is administered diotically with both the target sentence and 12-talker babble delivered to the left and right ears at the same time. In this study, the target sentence was delivered at 50 dB SL (re: binaural SRT), beginning first with the babble set to be 47 dB SL (+3 dB SNR). Two lists of 50-sentences were administered, with List 1 being presented at + 3 dB SNR and List 2 being presented at 0 dB. This yielded four conditions: +3 dB high predictability, +3 dB low predictability, 0 dB high predictability, and 0 dB low predictability. R-SPIN List 1 and 2 contain the same set of 50 target words, but the predictability of each target word is different across the two lists. For example, "spoon" occurs in a high predictability context in List 1 ("Stir your coffee with a spoon") and in a low predictability context in List 2 ("Bob could have known about the spoon"). According to the test developers, the two lists contain the same types of syllables, vowels, and consonants and when administered at the same SNR, they are equivalent in terms of difficulty and reliability. Using the Corpus of Contemporary American English (https://corpus.byu.edu/coca/), we also confirmed that the lexical frequency of the high and low predictably target words was matched within a list.

The two SNR levels used in this study were selected based on pilot testing. Before starting the 0 dB condition, the participant was first instructed that the task would be the same but that it might be more difficult to hear the man's voice. The percent correct for each condition was then calculated, with 25 being the highest possible raw score for each of the four conditions.

Passage Comprehension (Woodcock-Johnson III Tests of Achievement)

As a complement to the R-SPIN test, we administered an English reading comprehension that was taken from the Woodcock-Johnson III Tests of Achievement, in which a missing word must

For Peer Review

be filled in from context. Passage Comprehension evaluates written language comprehension at the sentence level by assessing the ability to make use of vocabulary, syntactic, and semantic knowledge to infer missing elements. For the purposes of this study, we selected this test to evaluate the ability to make use of top-down linguistic knowledge in a non-auditory condition (Woodcock, Mather, and McGrew 2001). For ages 14-30, the Passage Comprehension Test has high inter-test correlation (r >= 0.6) with the Oral Vocabulary, Oral Comprehension, Letter-Word Identification, and Spelling subtests of the Woodcock-Johnson III Tests of Achievement.

The Passage Comprehension Test was administered in a quiet room with the participant sitting across the table from the test administrator. During this test, the participant silently reads a sentence containing a missing word and then verbalizes the word that they think would best complete the sentence based on the context created by the other words in the sentence. This test is performed without a time limit. The test item was counted correct if it was included in the set of possible answers provided by the test manufacturers, or if it was a synonym of an answer provided by the manufacturer. The test includes a total of 38 items, beginning with simpler vocabulary and scaling to more advanced vocabulary. The first item administered was sentence 19, which is considered to be at a Grade 10 level. When tabulating the final score, the participant received credit for the first 19 sentences. Standard scores and percentiles were then calculated according to the test manufacturer's guidelines.

Statistical analyses

Percent scores were converted to rationalized arcsine transform units (RAU) for statistical analysis (Studebaker 1985). RAU linearizes percent scores, making the values better suited for analysis via linear tests (e.g., ANOVAs, t-tests).

For Peer Review

Results

Dichotic Digits, Competing Sentence, Time-Compressed Speech

Participants performed at or near ceiling on the Dichotic Digits and Competing Sentences tests (Figure 1). For these tests, a mixed-model repeated-measures Analysis of Variance (ANOVA) was performed using group (bilingual, monolingual) as the between-subjects factor, and ear (left vs. right) as the within-subjects factors. All variables met the assumptions of Sphericity.

For the Dichotic Digits test, there was the expected main effect of ear ($F(1,22) = 15.60$, $p = 0.001$, Partial Eta Squared (ηp^2) = 0.41), with lower accuracy for the left ear compared to the right ear. Two of the participants (1 bilingual, 1 monolingual) fell below the 90% cutoff for the left ear, but achieved perfect or near perfect scores in the right ear. For the Dichotic Digits test, neither the main effect of group ($F(1,22) = 0.05$, $p = 0.83$, $\eta p^2 = 0.002$), nor the ear-by-group interaction was significant ($F(1,22) = 0.05$, $p = 0.82$, $\eta p^2 = 0.002$).

For the Competing Sentences test, the main effect of ear was trending ($F(1,22) = 3.28$, $p = 0.09$, $\eta p^2 = 0.13$), with performance being lower for the left ear relative to the right ear. In this case, three participants (all bilingual) scored below the 90% cutoff for the left ear. For this test, a group effect emerged ($F(1,22) = 5.30$, $p = 0.03$, $\eta p^2 = 0.19$), but the interaction between ear and group was only trending ($F(1, 22) = 2.27$, $p = 0.12$, $\eta p^2 = 0.11$). Planned post-hoc analysis revealed that the bilingual group underperformed the monolinguals on the left ear ($t(22) = 2.37$, $p = 0.03$, $d = 1.0$), but that the groups were matched on the right ear condition, with both groups scoring ~98% ($t(22) = 0.53$, $p = 0.59$, $d = 0.22$). For the bilingual group, the average score was 98.5% for the right ear (SD = 1.67) compared to 93.92% for the left ear (SD = 6.14) ($t(11) = 2.02$, $p = 0.07$, $d = 1.38$).

For Peer Review

For the Time-Compressed Speech With Reverberation Test, the effect of ear was trending ($F(1,22) = 3.65$, $p = 0.07$, $\eta p^2 = 0.12$); however, neither the main effect of group ($F(1,22) = 0.27$, $p = 0.61$, $\eta p^2 = 0.002$), nor the ear-by-group interaction was significant ($F(1,22) = 0.04$, $p = 0.84$, $\eta p^2 = 0.002$). On this test, all participants were in the clinically normal range.

[Figure 1]

R-SPIN

For the R-SPIN test, a mixed-model repeated measures ANOVA was performed using group (bilingual, monolingual) as the between subjects factor, and linguistic predictability (high vs. low) as well as SNR (0, 3 dB) as within-subjects variables (Figure 2).

We start by reporting the within-subjects comparisons followed by the group comparisons: As expected, main effects of SNR and Predictability were observed (SNR: $F(1,22) = 39.82$, $p<0.005$, $\eta p^2 = 0.62$; Predictability: $F(1,22) = 511.89$, $p<0.005$, $\eta p^2 = 0.96$), with less accurate final word recognition observed in the 0 dB SNR condition compared to the 3 dB SNR condition (mean (standard deviation) = 70.5% (SD = 10.06) vs. 79.17% (SD = 6.43)) and also less accurate word recognition in the low predictability compared to the high predictability conditions (58.92% (SD = 9.62) vs. 90.75% (SD = 6.61)). The facilitative influence of linguistic context, however, was different across the two SNR conditions (SNR x Predictability Interaction, $F(1,22) = 13.54$, $p = 0.001$, $\eta p^2 = 0.38$), with greater benefits of context observed for the 0 dB SNR condition than the 3 dB condition.

With respect to group comparisons, the overall main effect of group was trending towards significance ($F(1,22) = 3.05$, $p = 0.09$, $\eta p^2 = 0.12$). Moreover, the SNR by group interaction was not significant ($F(1,22) = 0.21$, $p = 0.65$, $\eta p^2 = 0.009$), with both groups showing a performance

For Peer Review

decrement of ~8% when the SNR dropped from 3 to 0 dB. This can be seen visually in Figure 2; the distance between the square markers (3 dB condition) and the circle markers (0 dB) is matched for the two groups. Thus, on the R-SPIN test, our bilingual group was not inordinately affected by background noise compared to the monolinguals on the R-SPIN test. However, the bilingual group did differ from the monolingual group in terms of how much they benefitted from the linguistic predictability of the final word. In Figure 2, this manifests as difference in the slope of the lines connecting the low and high predictability conditions, with the slope being less steep for the bilingual group compared to the monolingual group. Collapsing across the two SNR conditions, the bilingual group had an average performance boost of 28.17% for the high predictability sentences over the low predictability sentences, compared to a 35.5% increase for the monolingual group. This is a small, yet, significant effect (Predictability x Group interaction ($F(1,22) = 14.27$, $p = 0.001$, $\eta p^2 = 0.39$). Planned post-hoc comparisons revealed that the groups had equivalent performance in the low predictability condition ($t(22) = -0.21$, $p = 0.84$, $d = 0.08$) but differed on the high predictability condition ($t(22) = 2.72$, $p = 0.01$, $d = 1.2$). Finally, the three-way interaction between Group, SNR, and Predictability was not significant ($F(1,1,1,22) = 0.01$, $p = 0.91$, $\eta p^2 = 0.001$), suggesting that the differential effect of predictability for the two groups did not differ as a function of SNR.

[Figure 2]

Passage Comprehension

The two groups performed similarly on the Passage Comprehension test ($t(22) = 0.10$, $p = 0.92$, $d = 0.04$). For the monolingual group, the average standard score was 113.50 (SD = 7.78), with a range from 103-126 (58^{th} to 96^{th} percentile). For the bilingual group, the average standard score was 113.83 (SD = 8.34), with a range from 96-126 (39^{th} to the 96^{th} percentile).

For Peer Review

Discussion

Speech understanding in noise is a complex process that reflects many different factors. Successful performance relies not only on the integrity of peripheral hearing and central auditory processes, but also on the ability to map the neural representation of the acoustic signal to a phonetic unit, match this phonetic information to lexical items, and use top-down linguistic knowledge including lexical, syntactic, semantic, and pragmatic information to interpret missing or obscured bottom-up information. Lagacé, Jutras, and Gagné (2010) propose that the R-SPIN test has the advantage over other SIN tests by being able to dissociate whether the SIN weakness has its roots in auditory or language-based functions. Using the R-SPIN, together with a battery of three commonly used tests of Central Auditory Processing Disorder, we do not find any evidence that bilingual listeners who self-rate as being proficient in the test language differ from monolingual listeners on their global auditory processing skills when utilizing testing materials that involve speech stimuli. Instead, our constellation of findings suggest that differences between monolingual and bilingual individuals on the R-SPIN test reflect less efficient top-down processing of speech. However, we are careful to point out that any apparent weakness or disadvantage observed in our bilingual group, should not be construed as an impairment, given the overall high level performance seen across all tests.

In the following sections, we examine the degree to which the bilingual disadvantage in SUN reflects linguistic and not global auditory processes, the possibility that SUN weaknesses are an inevitable byproduct of speaking two languages, how our findings fit within the debate on the advantages and disadvantages of speaking two languages, and the clinical implications of this

For Peer Review

line of work. Throughout these sections, we will highlight the novelty and limitations of our study, and propose new avenues for investigating SUN in bilingual speakers, including the need for multimodal testing and other forms of SIN testing to further evaluate how bilingual listeners perform in acoustically-degraded conditions.

The bilingual SUN disadvantage reflects linguistic not general auditory processing abilities

Our cohort of bilingual speakers was found to underperform on the R-SPIN test but only under specific conditions. When manipulating the level of the background noise relative to the target sentence, we found that bilingual and monolingual listeners received a similar level of performance benefit when the SNR was more favorable, contradicting the claim that bilinguals listeners are experiencing "auditory processing degradation" (Lucks Mendel and Widner 2016). Instead, the outcomes of the R-SPIN test conditions suggest that bilingual listeners, even those who self-rate as being highly-proficient in the test language, are weaker at utilizing compensatory cues to aid their speech understanding in noise, leading them to underperform on R-SPIN but only in conditions where the final (English) word can be restored from context. For some of the R-SPIN sentences, the contextual cues are so strong that the final word can be deduced without any clues about the auditory signal. However, when asked to fill in a missing word based on context in a written sentence, with no time limits and no auditory clues, the two groups performed similarly, with both achieving high, near ceiling scores on the Passage Comprehension WRMT-III test. The bilingual group also performed on par with their monolingual counterparts in acoustically-degraded conditions when the target English word was not predictable from context, as seen in the low predictability conditions of the R-SPIN test and the Time-Compressed Speech with Reverberation test, in which listeners are prompted to repeat a target word without any aiding linguistic context ("Say the word ____"). From this collective

For Peer Review

evidence, we conclude that the bilingual college students in our sample were able to access and utilize top-down linguistic knowledge but that they may not have been able to capitalize on it to the same degree as monolingual listeners when sensory input was degraded (as in the R-SPIN test).

Consistent with the source of the bilingual disadvantage for SUN being top-down not bottom-up in nature, previous work has shown that the performance gap between bilingual and monolingual listeners does not increase as the amount of energetic masking increases (Rogers et al. 2006) or as the amount of time-compression increases (Shi and Farooq 2012), supporting the finding that our bilingual group is not inordinately affected by increasing levels of background noise on such tests. Moreover, when word- and sentence-level factors are stripped away from the SIN task, and the focus is shifted to identifying individual English phonemes, monolinguals and non-native listeners have also been found to perform similarly (Cutler et al. 2004). In addition, there is evidence to suggest that noise tolerance levels, as measured by the Acceptable Noise Level (ANL) test, are matched between bilingual and monolingual individuals, even when speech understanding is reduced (von Hapsburg and Bahng 2006); however, other findings suggest that ANL scores are influenced by the listener's language background (Shi, Azcona, and Buten 2015).

In our healthy young adult population, using the R-SPIN test, we isolated the bilingual disadvantage to top-down linguistic factors, although in our case the disadvantage (relative to monolingual peers) is small. This weakness in leveraging top-down information is consistent with what has been found for listeners who are less proficient in the target language. For example, von Hapsburg and Bahng (2006) found that individuals who self-report as being moderately proficient in the test language (English) derived less benefit from context on the R-

For Peer Review

SPIN test than monolingual individuals, but those who self-report as having low proficiency in the target language show no benefit of linguistic predictability. Likewise, Mayo, Florentine, and Buus (1997) reported that bilingual listeners who acquire the target language late (after age 12), reach native-like proficiency on sentence recognition in quiet but they do not derive the same benefit from cross-word context when the speech signals are presented in noise, again suggesting less efficient top-down linguistic processing. A similar pattern of findings was reported for bilinguals speakers who were asked to recall English passages in noise: compared to monolingual speakers, bilinguals speakers with high self-rated English proficiency, but more wide ranging ages of acquisition than the current study, did not derive the same benefit from the linguistic cues afforded by inter-connected, linguistically-related English sentences (Shi 2012). These findings in diverse bilingual populations echo what has been shown in non-native listeners who likewise have less efficient use of top-down cues in their non-native compared to their native language in background noise but also quiet conditions (Hervais-Adelman, Pefkou, and Golestani 2014). Similar to the von Hapsburg and Bahng (2006) findings for bilingual speakers with (self-rated) low proficiency in the target language, Bradlow and Alexander (2007) reported that non-native speakers could not take advantage of sentence-level context unless the speech signal was produced in clear speech, a type of speaking style often adopted by talkers in adverse communication environments in which the speaking rate is slowed and individual speech sounds are more discernable. Thus, whereas non-native speakers may not benefit from top-down linguistic cues to facilitate word recognition in noise until the speech signal becomes perceptually favorable, bilingual speakers with more native-like proficiency in the test language appear to be able to benefit from such cues, even in adverse listening conditions, but the benefit is not as great as a monolingual might achieve. Consistent with our findings, Schmidtke (2016)

For Peer Review

reported that bilingual listeners continued to underperform on R-SPIN sentences with high predictability relative to the monolingual listeners, even when using subsamples of bilingual and monolingual participants who were matched in language proficiency (as assessed by multiple standardized tests), suggesting that differences in language proficiency cannot fully account for these group differences. For a similar account see, Shi (2011).

Taken together, this combined evidence suggests that perceptual weaknesses for speech understanding in noise that are observed in bilingual individuals are not necessarily due to a lack of knowledge in the target language, or a lack of linguistic knowledge more generally, but that they are instead more likely due to a linguistic system that underperforms when the bottom-up acoustic input is less reliable. However, a central limitation of our study, is that we relied on self-report to estimate language proficiency. Moreover, although the high scores on the test of Passage Comprehension provide confirmatory evidence that the bilingual speakers in our sample have good mastery of English, this single test cannot provide a complete picture of language proficiency. Thus, although bilingual and monolingual groups were matched with respect to self-rated proficiency and performance on the Passage Comprehension tests, it is premature to conclude that they are necessarily matched on all aspects of English language use and knowledge. While it is common practice to rely on self-report, Shi (2011) calls this practice into question, especially for late language learners who are more likely to over-estimate their abilities. Another factor that is not adequately addressed by our study, or by the literature more generally, is how the quality of the exposure to the test language affects SIN performance and self-ratings of language proficiency, although the effect of language quantity and quality on language development are well recognized (Hart and Risley 1995; Ramírez-Esparza, García-

For Peer Review

Sierra, and Kuhl 2014). Our study findings, and comparisons to the broader literature, should, therefore, be interpreted with these limitations in mind.

Are SUN weaknesses inevitable for bilingual listeners?

It has recently been proposed that bilingual disadvantage for SUN is an inevitable consequence of being bilingual (Schmidtke 2016). Under this theoretical framework, weaknesses on SUN tasks are not due to a lack of knowledge about the target language but they are instead considered to be byproduct of knowing two languages and having to split one's time, as well as lexicon and phonetic inventory, across multiple languages. Thus, even when a bilingual speaker is a native speaker of the target language, she may still be at a communicative disadvantage, compared to monolinguals, when listening to speech in noise. Evidence supporting this "inevitability" viewpoint comes from current, well-accepted models of speech processing.

Current models of speech processing posit that upon hearing a (target) word, other similar sounding words and semantically related words, are simultaneously activated in the mental lexicon (McClelland and Elman 1986; Luce and Pisoni 1998; Magnuson et al. 2007). These simultaneously activated words compete internally for recognition with the target word, and the listener must select the word that is deemed most plausible. When the bottom-up sensory signal is obscured by noise or otherwise degraded, the signal becomes less reliable, creating less certainty about what was said, and, this, it is theorized, leads to a greater number of candidate words being activated, which in turn increases the processing load. In such cases, the listener must rely more on non-auditory processes to consider the plausibility of each candidate word and discard those deemed least probable based on lexical knowledge, such as word frequency and other top-down linguistic information. A variety of evidence suggests that bilingual individuals face an increased processing load as the result of both languages being activated in parallel

For Peer Review

during speech processing (Weber and Cutler 2004; Kerkhofs et al. 2006; Paulmann et al. 2006; Marian, Blumenfeld, and Boukrina 2008). This dual-activation results in both within- and across-language competitors being activated, producing a greater number of lexical competitors for bilingual compared to monolingual speakers. So, for example, when a monolingual English speaker hears the word, "kite", phonological neighbors such as "bite" and "right" will be activated. However, for a bilingual listener, the set of activated (competing) words may also include non-English words with similar phonology. As an illustration, a German-English bilingual speaker may also activate words like "kein" (none) or "weit" (far) (pronounced "kine" and "vite", respectively), when hearing "kite", creating more internal competition for bilingual listeners, who then must depend more on top-down knowledge to select the appropriate target word.

Further compounding these theorized lexical selection inefficiencies for bilingual individuals are lexical frequency effects. For all listeners, whether they are proficient in one or multiple languages, faster recognition and recall times are seen for frequently encountered words compared to less frequent words (Taft 1979). However, in the case of bilingual individuals, the disadvantage for low frequency words is exacerbated (Gollan et al. 2008). Consistent with this finding, Schmidtke (2016) found that low frequency words and words in less predictable linguistic conditions were recognized less accurately by bilingual listeners compared to monolingual listeners on a modified version of R-SPIN. Golan and colleagues (2008) propose that this disadvantage for low frequency words is the result of weaker connections between a lexical item and its phonological form, which arises because bilingual listeners activate each word in their lexicon less frequently than monolingual listeners simply because they know more words, on average, than a monolingual individuals. For example, a bilingual individual and

For Peer Review

monolingual individual may encounter the English word "kite" the same number of times but because the bilingual individuals has a larger number of phonological representations within their lexicon by virtue of knowing two languages, the word "kite" will be processed/activated as if it is less frequent compared to the monolingual individual, leading to slower and less efficient lexical recall for "kite". In further support of weaker lexical recall in bilinguals, bilingual speakers have also been shown to have lower verbal recall, slower reactions times on verbal recall measures, and poorer memory span for verbal information in both languages (Mägiste 1979; Bialystok et al. 2009)

This literature, in combination with our findings, support the notion that bilinguals individuals, even those who are native speakers of the test language, are at a comparative disadvantage, compared to monolinguals, when performing SUN tasks, as a consequence of less efficient lexical retrieval. Our study illustrates that the disadvantage can, in some cases, be quite small.

Bilingual advantages and disadvantages

In interpreting our findings, we are mindful that our participant sample was limited in size, that we used self-ratings of language proficiency, and that although our dataset captured an array of different language families, it was by no means representative of the diversity of languages worldwide. Nevertheless, our small study provides an important data point in the larger discussion on the potential disadvantages of being bilingual by helping to delineate the conditions under which the bilingual disadvantage for SUN may or may not emerge for bilinguals speakers who self-report as having native-like abilities in the test language. In addition, our study adds to the conversation on the potential benefits of being bilingual by providing evidence that the auditory processing advantages that have emerged for non-linguistic

For Peer Review

stimuli (Bak, Vega-Mendoza, and Sorace 2014; Krizman et al. 2016) and for auditory-evoked potentials to passively-attended speech stimuli (Vihla, Kiviniemi, and Salmelin 2002; Krizman et al. 2012; Krizman et al. 2015; Skoe et al. 2017) do not lead to any apparent behavioral benefits on the R-SPIN test nor on three tests routinely used to clinically assess Central Auditory Processing Disorder. However, as seen in the current study, most of the participants (regardless of group) performed at or near ceiling on Dichotic Digits and Competing Sentences tests, suggesting that these linguistic-based tests of central auditory processing lack the sensitivity to evaluate individual or group-level differences in central auditory processing for high performing (non-impaired) listeners. In addition, although Lagacé et al. (2011) theorize that the R-SPIN can be used to delineate auditory factors from linguistic factors, this claim has not undergone extensive scrutiny in the literature. Because it utilizes speech materials, we cannot rule out the possibility that the auditory processing component of the test (i.e., the manipulation of SNR in the R-SPIN) may reflect linguistic processing, at least to some degree. Phenomena such as the Ganong Effect further illustrate the difficulty of separating linguistic and perceptual processes when the stimuli are speech or speech-like (Ganong 1980). To better distill what specifically is being measured by the auditory dimension of the R-SPIN test, performance on the R-SPIN test should, in future investigations, be compared to performance non-linguistic tests to determine whether an auditory processing advantage observed on non-linguistic tests is associated with better performance for low SNR R-SPIN conditions. Until then, we also leave open the possibility that the auditory processing advantages observed in previous work in bilinguals may counteract disadvantages for SUN by increasing the fidelity of the bottom-up signal, as suggested by Krizman et al. (2016). Thus, enhanced basic auditory processing may help to level

For Peer Review

the playing field for processing speech in noise, although the data we present here do not provide evidence that either supports or refutes that possibility.

In addition, we did not observe evidence of a dichotic listening advantage in our bilingual participants. This stands in contrast to previous evidence of heightened dichotic processing in bilinguals (Soveri et al. 2011; Gresele et al. 2013). Gresele et al. 2013, like the current study, used the Dichotic Digits test, but that study covered a much wider age range (18-59) and did not control for differences in educational level. Theses differences could account for the discrepant findings. Soveri et al. (2011), in contrast, administered a phonemic version of a dichotic listening test in which two syllables (constant-vowel) were played dichotically at the same intensity (dB level not reported) to Finnish-Swedish bilingual listeners and Swedish monolingual listeners between the ages of 30 and 74. In the "non-forced" condition, where the listener was not given explicit instructions as to which ear to attend to, and was instead told to report back which syllable they heard best/first, both groups displayed a right-ear advantage and the groups did not differ in the degree of this advantage. This finding is consistent with what we observed for the Dichotic Digits test, in which the listener is asked to report back the numbers that they heard without selectively attending to one ear. In addition to the non-forced condition, Soveri et al. (2011) included two other listening conditions, where the listeners were instructed to attend to either the right or the left ear and report back what they heard. When attending to the left ear in dichotic listening situations, the listener must inhibit this right ear bias, and as a consequence, attending to the left ear is theorized to require more executive processing than attending to the right ear under dichotic stimulation (reviewed in Hugdahl et al. (2009)). In the Soveri et al. (2011) study, bilingual individuals had more accurate recall than the monolinguals for the left-

For Peer Review

attend and the right-attend conditions, which the authors took as evidence that bilingual individuals have stronger executive function.

As mentioned above, the Dichotic Digits Test administered in our study does not include a selective attention condition (at least not as part of standardized procedures), which limits comparison with the Soveri et al. (2011) study. However, we did administer the Competing Sentences Test, which does provide a more comparable analog to the Soveri et al. (2011) study. The Competing Sentences Test uses full sentences not syllables, and unlike the dichotic syllables test used by Soveri et al. (2011), the target sentence is presented at a lower intensity than the distractor. However, a right-ear advantage is still expected for the Competing Sentences Test, even under conditions where the signal to the right ear is 15 dB below the signal to the left ear (reviewed in Hugdahl et al. (2009)). Using the Competing Sentences Test, in our cohort of college students (who were younger than the Soveri et al. (2011) sample), we found that performance was at or near ceiling for both the left-attend and right-attend conditions, with the scores ranging from a low of 92.5% to a high of 100%. Yet, even in the face of these high scores, the bilingual group still showed a small but statistically significant drop in performance compared to the monolingual group. This was most evident when the task was to focus on the left ear compared to the right ear, consistent with the right-ear advantage for this task. The monolingual group, by contrast, showed no ear bias by performing at or near ceiling on both conditions. One interpretation of this finding is that the monolingual group in our dataset has more refined executive function than the bilingual group. Thus, while previous studies suggest that bilingual individuals may be able to draw on more refined executive skills to outperform monolingual individuals on selective auditory attention tasks that involve non-linguistic stimuli (Bak, Vega-Mendoza, and Sorace 2014) or simple linguistic stimuli such as numbers (Krizman

For Peer Review

et al. 2012) and syllables (Soveri et al. 2011), our findings suggest that the purported bilingual advantage in harnessing executive function does not necessarily transfer to more linguistically-challenging stimuli or that it may not manifest at all. A different, yet not contradictory interpretation is that listening tasks that require more cognitive control expose an underlying (subtle, in our case) difference in linguistic processing.

Although our small study, like recent more large-scale studies of bilinguals (Paap, Johnson, and Sawi 2014, 2015), does not provide direct or even indirect evidence of enhanced executive function in our young adult bilingual cohort, this does not necessarily discount the possibility that bilingualism may advantage certain aspects of executive function and/or auditory processing at points in life (Bialystok, Martin, and Viswanathan 2005). Future studies that include comprehensive assessments of language, executive, and auditory function are needed to more fully explore the linguistic and non-linguistic conditions under which selective listening advantages emerge for bilingual speakers at different points in life.

Is the bilingual disadvantage in noise modality specific?

We now turn to the question of whether the bilingual disadvantage is modality specific. In the case of dyslexia, difficulties understanding speech in noise has been theorized to be the outcome of a sensory-wide difficulty with excluding noise for both auditory (speech and non-speech), as well as visual conditions (Sperling et al. 2005). For musically trained populations, advantages have been seen for adverse (i.e., degraded or distorted) conditions across both visual and auditory modalities (Anaya, Pisoni, and Kronenberger 2016). In the case of bilingual individuals, current data suggest that the disadvantages that face bilingual speakers in noise are specific to speech and that they do not generalize to non-speech signals (Krizman et al. 2016); however, more work needs to be done to examine whether the bilingual weakness in noise is

For Peer Review

specific to language within the auditory modality or whether it might be more sensory pervasive and extend to written forms of language.

In the current investigation, we incorporated a visual test of language processing (Passage Comprehension) that assesses the ability to use top-down linguistic information. While both the auditory-based R-SPIN test and visual-based Passage Comprehension assess top-down language skills, the two tests are not true analogs (Bellis, Billiet, and Ross 2011). In the case of R-SPIN, the sensory input was degraded, but for the Passage Comprehension test the sensory (i.e., visual input) was not. Although we did not test the R-SPIN tests in a "quiet" condition without multi-talker babble, we can infer from the results of the Time-Compressed Speech Test with Reverberation and previous work in bilingual and trilingual speakers (Mayo, Florentine, and Buus 1997; Rogers et al. 2006; Tabri, Abou Chacra, and Pring 2011; Shi 2012) that the groups would likely perform similarly in an R-SPIN condition without background noise, at least on the low predictability condition. Another way in which the R-SPIN and Passage Comprehension tests differ is in their performance loads. The Passage Comprehension Test is administered without a time limit. Although the R-SPIN test is not a timed test, per se, the listener is expected to keep to the pace that the test materials are delivered in the digital recordings. The comparatively slower pace of the Passage Comprehension Test may have allowed the bilingual individuals to reach monolingual-like levels. Future research should consider incorporating a visual analog of the R-SPIN test in which the text is difficult to make out (e.g., blurred, faintly colored text, visual masker) to illuminate whether this bilingual weakness with top-down information under degraded condition is specific to auditory input (Zekveld and Kramer 2014; Anaya, Pisoni, and Kronenberger 2016). An investigation using noise in both auditory and visual would not only shed light on mechanisms of why bilingual speakers are likely to underperform

For Peer Review

on SUN tasks, but also help to guide clinical recommendations for accommodating such weakness.

Clinical implications

Specialists in all areas of healthcare, including audiologists and speech-language pathologists, are now treating a larger percentage of bilingual patients. Census reports from 2010 estimated that roughly 20% of the U.S. population is bilingual, with a nearly 40% rise in bilingualism between 1980 and 2010; however, clinical services, and the number of bilingual audiologists, have not necessarily kept pace with the growing bilingual population. The challenge with caring for bilingual populations is that most clinical norming criteria are based on monolingual datasets, and utilize English-only materials, and therefore do not take into account that bilingual individuals might have different performance baselines (von Hapsburg and Peña 2002). Our finding emphasize that (1) bilinguals listeners, even those with normal hearing, no noticeable accent, and who consider themselves to have high proficiency in English, may underperform on English SUN tests, under certain conditions and (2) the choice of test materials is critical. Shi (2011) recommended that for English-dominant, early bilinguals that proficiency-ratings of 8 or better (out of 10) are required for using monolingual normative values for the NU-6 test, although our findings suggest that this recommendation does not generalize to all SIN and CAP tests. The idea of developing bilingual-specific norms, and language-specific materials, for audiological use, as well as training more bilingual audiologists, is intuitively appealing; however, it is an inherently complex, potentially fraught process, given the diversity of bilingual backgrounds (different languages, different proficiency levels, etc.) that may be encountered in a clinical setting. Another area that needs further exploration is the potential impact of having a bilingual audiologist administer and score SIN tests. In the case of the current study, the test

For Peer Review

administrator was bilingual, and given that test scoring for many audiologist tests is contingent on the test administrator's perception of what the patient says, this could be viewed as a potential confound.

An alternative to developing language-specific norms is to use a test like the R-SPIN that can (theoretically) dissociate auditory and linguistic factors, and/or to use SIN tests that have a low linguistic load. A recent study suggests that children who learn two languages simultaneously can achieve the same level of performance on SIN tests as SES-matched children, when the age of English acquisition is matched and speech materials use simple vocabulary that minimizes linguistic load/bias (Reetzke et al. 2016). Reetzke et al. (2016) replicated this finding across multiple SNR conditions, under different types of maskers, and in auditory-only and audio-visual conditions, providing strong converging evidence that bilingual and monolingual individuals can achieve similar levels of performance in noise when linguistic materials are adequately controlled. For the purposes of assessing central auditory function in bilingual (as well monolingual) speakers, there is also value in incorporating non-linguistic measures of central auditory function (Moore et al. 2010; Ludwig et al. 2014), utilizing speech in noise tests that do not rely on a direct report of speech understanding, such as electrophysiological testing (Krizman et al. 2012), creating tests that allow for a greater spread of performance among non-impaired listeners, and/or administering subjective tests of noise tolerance and/or listening effort in noise (von Hapsburg and Bahng 2006; Shi, Azcona, and Buten 2015).

However, a first step in developing more bilingual-focused care, is to establish more widespread clinical recognition and scientific exploration of the specific advantages and disadvantages that bilinguals may display on tests that are routinely administered in the

For Peer Review

evaluation of central and peripheral auditory function. This will be key for developing more tailored strategies for counseling and remediation in disordered bilingual populations, as well as developing hearing conservation programs that address the specific difficulties faced by bilingual individuals.

Acknowledgements

This study was completed as partial fulfillment of the second author's Au.D. Degree. We'd like to extend our thanks to Parker Tichko for his comments on an earlier version of this manuscript, and Kelly Linehan for her assistance with data coding and copyediting.

Author Contributions

K.K and E.S. designed the study, K.K collected the data, K.K and E.S. analyzed the data, and E.S. wrote the paper with input from K.K.

References

Anaya EM, Pisoni DB, Kronenberger WG. (2016) Long-term musical experience and auditory and visual perceptual abilities under adverse conditions. *J Acoust Soc Am* 140: 2074-2081.

Bak TH, Vega-Mendoza M, Sorace A. (2014) Never too late? An advantage on tests of auditory attention extends to late bilinguals. *Front Psychol* 5.

Bellis TJ, Billiet C, Ross J. (2011) The utility of visual analogs of central auditory tests in the differential diagnosis of (central) auditory processing disorder and attention deficit hyperactivity disorder. *J Am Acad Audiol* 22: 501-14.

Bialystok E, Martin MM, Viswanathan M. (2005) Bilingualism across the lifespan: The rise and fall of inhibitory control. *Int J Biling* 9: 103-119.

Bialystok E, Craik FI, Green DW, Gollan TH. (2009) Bilingual minds. *Psychol Sci Public Interest* 10: 89-129.

Bialystok E. (2015) Bilingualism and the development of executive function: The role of attention. *Child development perspectives* 9: 117-121.

Bilger RC, Nuetzel JM, Rabinowitz WM, Rzeczkowski C. (1984) Standardization of a test of speech perception in noise. *J Speech Hear Res* 27: 32-48.

Bradlow AR, Alexander JA. (2007) Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *J Acoust Soc Am* 121: 2339-2349.

Costa A, Hernández M, Costa-Faidella J, Sebastián-Gallés N. (2009) On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition* 113: 135-149.

Cutler A, Weber A, Smits R, Cooper N. (2004) Patterns of english phoneme confusions by native and non-native listeners. *J Acoust Soc Am* 116: 3668-3678.

For Peer Review

Ganong WF. (1980) Phonetic categorization in auditory word perception. *J Exp Psychol: Human Percept Perform* 6: 110.

Garcia-Sierra A, Ramirez-Esparza N, Silva-Pereyra J, Siard J, Champlin CA. (2012) Assessing the double phonemic representation in bilingual speakers of spanish and english: An electrophysiological study. *Brain Lang* 121: 194-205.

Gollan TH, Montoya RI, Cera C, Sandoval TC. (2008) More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *J Mem Lang* 58: 787-814.

Green DW. (1998) Mental control of the bilingual lexico-semantic system. *Bilingualism: Lang Cog* 1: 67-81.

Gresele AD, Garcia MV, Torres EM, Santos SN, Costa MJ. (2013) Bilingualism and auditory processing abilities: Performance of adults in dichotic listening tests. *Codas* 25: 506-12.

Hart B, Risley TR. (1995) *Meaningful differences in the everyday experience of young american children*. Paul H Brookes Publishing.

Hervais-Adelman A, Pefkou M, Golestani N. (2014) Bilingual speech-in-noise: Neural bases of semantic context use in the native language. *Brain Lang* 132: 1-6.

Hugdahl K, Westerhausen R, Alho K, Medvedev S, Laine M, Hamalainen H. (2009) Attention and cognitive control: Unfolding the dichotic listening story. *Scand J Psychol* 50: 11-22.

Kerkhofs R, Dijkstra T, Chwilla DJ, de Bruijn ER. (2006) Testing a model for bilingual semantic priming with interlingual homographs: Rt and n400 effects. *Brain Res* 1068: 170-83.

Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. (2004) Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 116: 2395-2405.

Krizman J, Marian V, Shook A, Skoe E, Kraus N. (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proc Natl Acad Sci USA* 109: 7877-81.

———. (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences USA* 109: 7877-81.

Krizman J, Skoe E, Marian V, Kraus N. (2014) Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain Lang* 128: 34-40.

Krizman J, Slater J, Skoe E, Marian V, Kraus N. (2015) Neural processing of speech in children is influenced by extent of bilingual experience. *Neurosci Let* 585: 48-53.

Krizman J, Bradlow AR, Lam SS-Y, Kraus N. (2016) How bilinguals listen in noise: Linguistic and non-linguistic factors. *Bilingualism: Lang Cog*: 1-10.

Lagacé J, Jutras B, Gagné J-P. (2010) Auditory processing disorder and speech perception problems in noise: Finding the underlying origin. *Am J Audiol* 19: 17-25.

Lagacé J, Jutras B, Giguère C, Gagné J-P. (2011) Speech perception in noise: Exploring the effect of linguistic context in children with and without auditory processing disorder. *Int J Audiol* 50: 385-395.

Luce PA, Pisoni DB. (1998) Recognizing spoken words: The neighborhood activation model. *Ear Hear* 19: 1-36.

Lucks Mendel L, Widner H. (2016) Speech perception in noise for bilingual listeners with normal hearing. *Int J Audiol* 55: 126-34.

Ludwig AA, Fuchs M, Kruse E, Uhlig B, Kotz SA, Rubsamen R. (2014) Auditory processing disorders with and without central auditory discrimination deficits. *J Assoc Res Otolaryngol* 15: 441-64.

Mägiste E. (1979) The competing language systems of the multilingual: A developmental study of decoding and encoding processes. *J Verbal Learning Verbal Behav* 18: 79-89.

Magnuson JS, Dixon JA, Tanenhaus MK, Aslin RN. (2007) The dynamics of lexical competition during spoken word recognition. *Cog Sci* 31: 133-156.

Marian V, Blumenfeld HK, Boukrina OV. (2008) Sensitivity to phonological similarity within and across languages. *J Psychol Res* 37: 141-170.

Mayo LH, Florentine M, Buus S. (1997) Age of second-language acquisition and perception of speech in noise. *J Speech Hear Res* 40: 686-93.

McClelland JL, Elman JL. (1986) The trace model of speech perception. *Cog Psychol* 18: 1-86.

Montagni I, Peru A. (2011) Effects of language and music learning on pitch and duration perception: An experimental investigation. *Int J Psychol Behavior Sci* 1: 33-40.

Moore DR, Ferguson MA, Edmondson-Jones AM, Ratib S, Riley A. (2010) Nature of auditory processing disorder in children. *Pediatrics* 126: e382-90.

Musiek FE. (1983) Assessment of central auditory dysfunction: The dichotic digit test revisited. *Ear Hear* 4: 79-83.

Nilsson M, Soli SD, Sullivan JA. (1994) Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 95: 1085-99.

Paap KR, Johnson HA, Sawi O. (2014) Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *J Cogn Psychol* 26: 615-639.

———. (2015) Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69: 265-278.

Parbery-Clark A, Skoe E, Lam C, Kraus N. (2009) Musician enhancement for speech-in-noise. *Ear Hear* 30: 653-61.

Paulmann S, Elston-Guttler KE, Gunter TC, Kotz SA. (2006) Is bilingual lexical access influenced by language context? *Neuroreport* 17: 727-31.

Ramírez-Esparza N, García-Sierra A, Kuhl PK. (2014) Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Dev Sci* 17: 880-891.

Reetzke R, Lam BP-W, 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.

Rogers CL, Lister JJ, Febo DM, Besing JM, Abrams HB. (2006) Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholing* 27: 465-485.

Schmidtke J. (2016) The bilingual disadvantage in speech understanding in noise is likely a frequency effect related to reduced language exposure. *Front Psychol* 7.

Shi L-F. (2010) Perception of acoustically degraded sentences in bilingual listeners who differ in age of english acquisition. *J Speech Hear Res* 53: 821-835.

———. (2011) How "proficient" is proficient? Subjective proficiency as a predictor of bilingual listeners' recognition of english words. *Am J Audiol* 20: 19-32.

———. (2012) Contribution of linguistic variables to bilingual listeners' perception of degraded english sentences. *J Speech Hear Res* 55: 219-234.

Shi L-F, Farooq N. (2012) Bilingual listeners' perception of temporally manipulated english passages. *J Speech Hear Res* 55: 125-138.

Shi L-F, Azcona G, Buten Ll. (2015) Acceptance noise level: Effects of the speech signal, babble, and listener language. *J Speech Hear Res* 58: 497-508.

Skoe E, Burakiewicz E, Figueiredo M, Hardin M. (2017) Basic neural processing of sound in adults is influenced by bilingual experience. *Neurosci* 349: 278-290.

Slater J, Skoe E, Strait DL, O'Connell S, Thompson E, Kraus N. (2015) Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behav Brain Res* 291: 244-52.

Slater J, Kraus N. (2016) The role of rhythm in perceiving speech in noise: A comparison of percussionists, vocalists and non-musicians. *Cogn Process* 17: 79-87.

Soveri A, Laine M, HÄMÄLÄInen H, Hugdahl K. (2011) Bilingual advantage in attentional control: Evidence from the forced-attention dichotic listening paradigm. *Bilingualism: Language and Cognition* 14: 371-378.

———. (2011) Bilingual advantage in attentional control: Evidence from the forced-attention dichotic listening paradigm. *Bilingualism: Lang Cog* 14: 371-378.

For Peer Review

Sperling AJ, Lu ZL, Manis FR, Seidenberg MS. (2005) Deficits in perceptual noise exclusion in developmental dyslexia. *Nat Neurosci* 8: 862-3.

Studebaker GA. (1985) A "rationalized" arcsine transform. *J Speech Hear Res* 28: 455-462.

Tabri D, Abou Chacra KM, Pring T. (2011) Speech perception in noise by monolingual, bilingual and trilingual listeners. *Int J Lang Commun Disord* 46: 411-22.

Taft M. (1979) Recognition of affixed words and the word frequency effect. *Mem Cogn* 7: 263-272.

Vihla M, Kiviniemi K, Salmelin R. (2002) Auditory cortical activation in finnish and swedish speaking finns: A magnetoencephalographic study. *Neurosci Let* 322: 141-4.

von Hapsburg D, Peña ED. (2002) Understanding bilingualism and its impact on speech audiometry. *J Speech Hear Res* 45: 202-213.

von Hapsburg D, Bahng J. (2006) Acceptance of background noise levels in bilingual (korean-english) listeners. *J Am Acad Audiol* 17: 649-658.

Weber A, Cutler A. (2004) Lexical competition in non-native spoken-word recognition. *J Mem Lang* 50: 1-25.

Willeford JA. (1978) Sentence tests of central auditory dysfunction. In: Katz J, Medwetsky L, Burkard R, eds. *Handbook of clinical audiology*, Wolters Kluwer/Lippincott Williams & Wilkins, Philadelphia, PA

Wilson RH, Preece JP, Salamon DL, Sperry JL, Bornstein SP. (1994) Effects of time compression and time compression plus reverberation on the intelligibility of northwestern university auditory test no. 6. *J Am Acad Audiol* 5: 269-77.

Wilson RH, Abrams HB, Pillion AL. (2003) A word-recognition task in multitalker babble using a descending presentation mode from 24 db to 0 db signal to babble. *J Rehabil Res Dev* 40: 321-7.

Woodcock R, Mather N, McGrew K. (2001) *Woodcock-johnson iii tests of cognitive abilities examiner's manual.* Riverside, Itasca.

Zekveld AA, Kramer SE. (2014) Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiol* 51: 277-284.

For Peer Review

Figure Captions

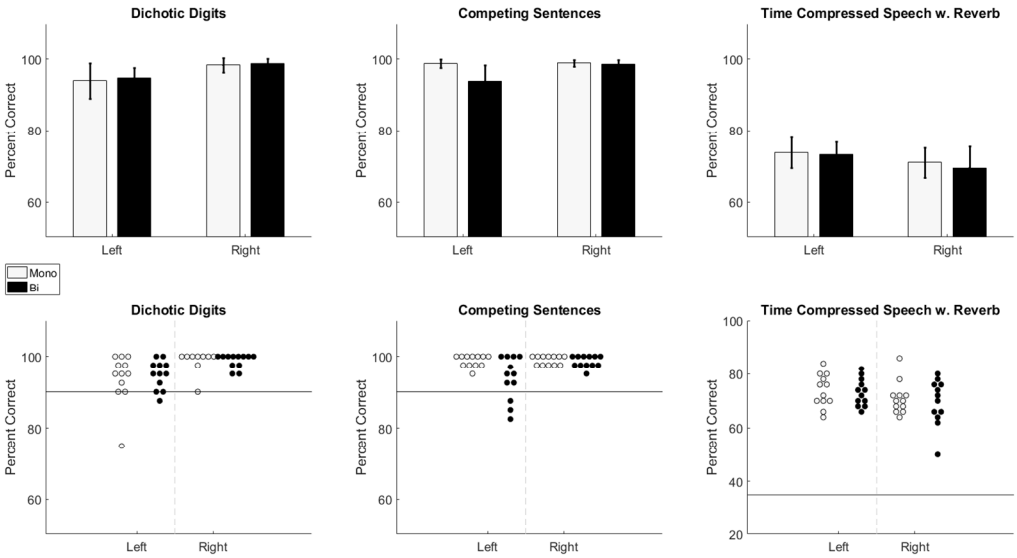


Figure 1. Comparisons between the Monolingual (gray) and Bilingual (black) groups on the Dichotic Digits, Competing Sentences, and Time-Compressed Speech with Reverberation Tests. In the top row, group means are plotted for each test, with error bars representing one standard error of the mean. In the bottom row, one-dimensional scatter plots show the distribution of scores across groups, ear, and tests. The horizontal line represents the cutoff score used for evaluating Central Auditory Processing Disorder (CAPD); scores below the line are considered abnormal. Note the number of perfect scores (100%) for the right ear for both Dichotic Digits and Competing Sentences. In the case of Dichotic Digits, also note that a number of the data points are overlapping for the monolingual group for the right ear.

Peer Review

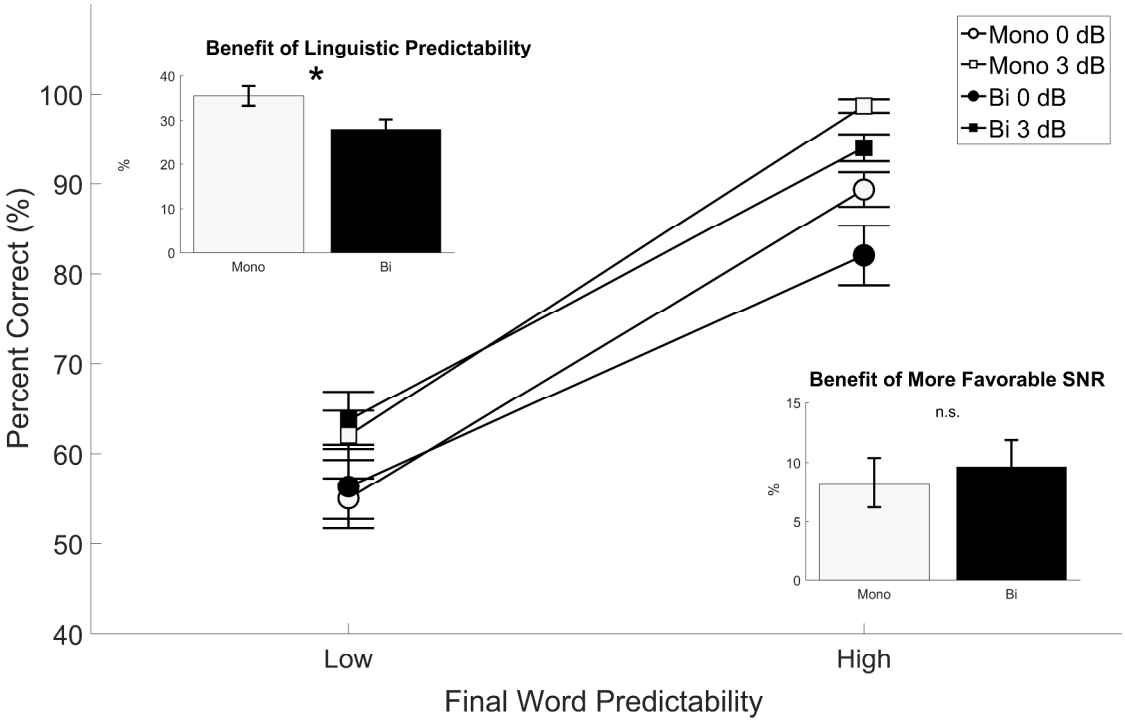


Figure 2. Comparisons between the Monolingual (gray) and Bilingual (black) groups on the R-SPIN Test. Center plot: Results of the four conditions for the two groups, with 0 dB SNR conditions plotted with squares and the 3 dB condition plotted with circles. Across both the Low (left) and High (right) predictability conditions, performance improved as the SNR increased from 0 to 3 dB; however, both group benefited to the same degree (bottom right inset panel). There was also a sharp improvement in performance when the final word could be deduced from context (high predictability condition) compared to when it could not (low predictability condition). In this case the extent of the improvement was greater for the monolingual group than the bilingual group (top left inset panel).

For Peer Review

Age window of First Exposure to English (years)	Age	Sex	Non-English Language (NEL)	English Proficiency (/10)	NEL Proficiency (/10)	Literate in NEL	Live outside U.S
0-3	26	F	Spanish	10	8	Yes	
0-3	21	F	Portuguese	10	9	Yes	
0-3	19	F	Telugu	9	8	No	
0-3	22	F	Spanish	10	10	Yes	2 years
0-3	20	F	Spanish	10	7	No	
0-3	20	F	Mongolian	10	9	No	2 years
0-3	23	F	Bangla	10	8	No	
0-3	18	M	Tamil	10	8	No	3 years
3-6	22	M	Japanese	10	9	Yes	4 years
3-6	20	F	Polish	10	7	No	
3-6	20	F	Serbian	10	7	Yes	2 years
6-9	24	M	Portuguese	9	10	Yes	14 years

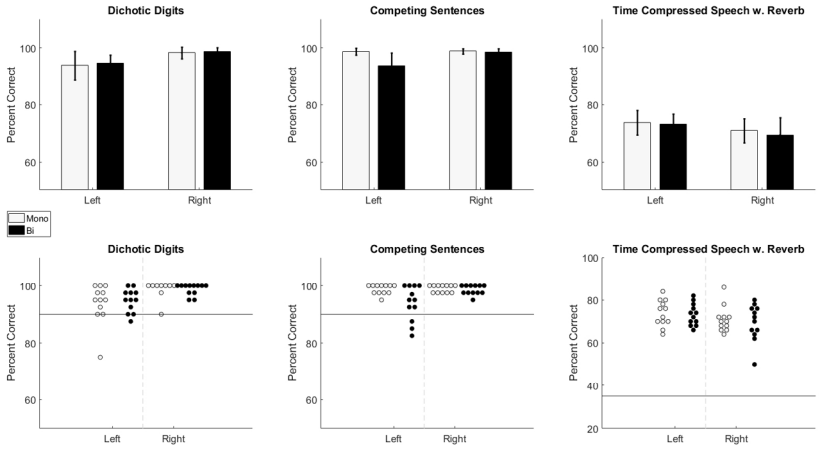
Table 1. Bilingual group demographics

For Peer Review

		N	Mean	Std. Deviation
Age (years)	Mono	12	20.30	1.22
	Bi	12	21.31	2.22
L1 Self-Rated Proficiency (/10)	Mono	12	9.92	0.29
	Bi	12	9.25	1.22
L2 Self-Perceived Proficiency (/10)	Mono	12	3.17	1.64
	Bi	12	8.92	1.00
English Self-Rated Proficiency (/10)	Mono	12	9.92	0.29
	Bi	12	9.75	0.45
Musical Training (years)	Mono	12	3.42	4.33
	Bi	12	4.50	3.82
Maternal Education (years) *	Mono	12	15.33	2.95
	Bi	11	14.73	3.26
PTA: RIGHT (dB HL)	Mono	12	12.22	5.92
	Bi	12	10.56	3.65
PTA: LEFT (dB HL)	Mono	12	10.00	4.55
	Bi	12	8.75	3.42

Table 2. Group means and standard deviations for age, self-rated current proficiency of L1, L2 and English, musical training, maternal education, and pure tone audiometric thresholds (pure tone averages (PTA)). (*One of the bilingual participants did not answer this question as the result of a photocopying error.)

For Peer Review



Comparisons between the Monolingual (gray) and Bilingual (black) groups on the Dichotic Digits, Competing Sentences, and Time-Compressed Speech with Reverberation Tests. In the top row, group means are plotted for each test, with error bars representing one standard error of the mean. In the bottom row, one-dimensional scatter plots show the distribution of scores across groups, ear, and tests. The horizontal line represents the cutoff score used for evaluating Central Auditory Processing Disorder (CAPD); scores below the line are considered abnormal. Note the number of perfect scores (100%) for the right ear for both Dichotic Digits and Competing Sentences. In the case of Dichotic Digits, also note that a number of the data points are overlapping for the monolingual group for the right ear.

508x255mm (96 x 96 DPI)

For Review