RESEARCH ARTICLE

Effect of Simultaneous Bilingualism on Speech Intelligibility across Different Masker Types, Modalities, and Signal-to-Noise Ratios in School-Age Children

Rachel Reetzke[1], Boji Pak-Wing Lam[1], Zilong Xie[1], Li Sheng[1¤a],
Bharath Chandrasekaran[1,2,3]*

1 Department of Communication Sciences and Disorders, The University of Texas at Austin, Austin, Texas, United States of America, 2 Institute for Neuroscience, The University of Texas at Austin, Austin, Texas, United States of America, 3 Cognitive Neuroscience, The University of Texas at Austin, Austin, Texas, United States of America

¤a Current address: Department of Communication Sciences and Disorders, University of Delaware, Newark, Delaware, United States of America

* bchandra@utexas.edu

Abstract

Recognizing speech in adverse listening conditions is a significant cognitive, perceptual, and linguistic challenge, especially for children. Prior studies have yielded mixed results on the impact of bilingualism on speech perception in noise. Methodological variations across studies make it difficult to converge on a conclusion regarding the effect of bilingualism on speech-in-noise performance. Moreover, there is a dearth of speech-in-noise evidence for bilingual children who learn two languages simultaneously. The aim of the present study was to examine the extent to which various adverse listening conditions modulate differences in speech-in-noise performance between monolingual and simultaneous bilingual children. To that end, sentence recognition was assessed in twenty-four school-aged children (12 monolinguals; 12 simultaneous bilinguals, age of English acquisition ≤ 3 yrs.). We implemented a comprehensive speech-in-noise battery to examine recognition of English sentences across different modalities (audio-only, audiovisual), masker types (steady-state pink noise, two-talker babble), and a range of signal-to-noise ratios (SNRs; 0 to -16 dB). Results revealed no difference in performance between monolingual and simultaneous bilingual children across each combination of modality, masker, and SNR. Our findings suggest that when English age of acquisition and socioeconomic status is similar between groups, monolingual and bilingual children exhibit comparable speech-in-noise performance across a range of conditions analogous to everyday listening environments.

Introduction

Speech communication rarely occurs in favorable listening conditions. Take, for example, the classroom. Various noise types, such as the loud heating and cooling system or students

chatting in an adjacent hallway, compete with target speech signals and critically tax perceptual, linguistic, and cognitive processes. 20% of school-age children in the United States speak a language other than English at home, and this percentage is predicted to increase in the coming years [1]. Emerging work has shown that bilinguals demonstrate task-specific perceptual and cognitive advantages relative to monolinguals [2, 3], and disadvantages in some linguistic processes such as lexical retrieval and vocabulary knowledge (for a review see, [4]). Given the disparity between bilinguals and monolinguals on these fundamental processes, do the groups significantly differ in recognizing speech in challenging listening conditions? This question becomes further complex when considering bilingual children who are in the process of acquiring two languages at the same time [5]. Most of the existing literature on speech-in-noise abilities in bilingual children has focused on children for whom English is their second language [6, 7]. While these studies are important, there is a significant gap in speech-in-noise evidence for bilingual children who learn two languages *simultaneously*. The goal of the present study was to investigate the extent to which simultaneous bilingualism in school-age children impacts English sentence recognition in noise using a comprehensive battery that tests sentence recognition under various adverse listening conditions that differentially affect perceptual and cognitive processes.

Speech perception in noise loads on perceptual, cognitive, and linguistic processes

Successful communication in challenging listening conditions requires a complex interdependence among perceptual, cognitive, and linguistic processes. In order to select a desired signal from competing sound sources, the auditory system must extract key acoustic features from the incoming stimulus stream. Interference with perceptual processes occurs when degradation caused by background noise renders portions of the target signal imperceptible [8]. In turn, this may lead to challenges in mapping key acoustic features to phonetic representations, and consequently to lexical representations (for a review see, [9]). A large body of work also implicates a role for cognitive processes in the extraction of speech from background noise. For example, the ease of language understanding (ELU) model indicates that for speech comprehension to occur sublexical information is Rapidly, Automatically, and Multimodally Bound into a PHOnological representation (RAMBPHO), a type of temporary storage system [10]. Poor speech signal quality due to noise results in ambiguous information in RAMBPHO which cannot be easily matched to stored speech representations. In such cases, working memory is needed to explicitly reprocess the incoming speech signal to resolve the mismatch. Working memory, a subcomponent of executive function [11], is engaged during speech recognition in adverse listening conditions to temporarily store salient acoustic cues, such as the fundamental frequency of the target speaker [9]. Cognitive demands tend to be greatest when the competing sound source has meaningful content, e.g. informational maskers [9, 12–14]. Like energetic masking (e.g. steady-state noise), informational masking involves selective attention and signal separation in the auditory periphery. However, recognizing speech in informational maskers not only requires listeners to cope with signal degradation due to energetic masking, but also to expend greater cognitive resources to ignore the meaningful distractors and attend to the target signal [13, 15]. Finally, speech recognition in challenging listening conditions is also dependent upon linguistic factors, such as vocabulary knowledge and lexical access. For example, a recent study demonstrated a positive correlation between phonemic restoration benefit and receptive vocabulary size and verbal intelligence, as measured by the Peabody Picture Vocabulary Test (PPVT-III-NL[16])[17]. The ELU additionally posits that individual differences in the quality of speaker-internal phonological representations in long-

term memory also contributes to individual differences in the perception of speech in adverse listening conditions [10]. Taken together, these findings suggest that linguistic knowledge significantly contributes to the restoration of degraded speech.

Bilingualism differentially impacts perceptual, cognitive, and linguistic processes

The bilingual experience differentially shapes perceptual, cognitive, and linguistic mechanisms across the lifespan resulting in outcomes ranging from better to poorer performance, when compared to monolingual age-matched peers [4]. Constant communication in two languages can fine-tune subcortical auditory responses to incoming speech signals [18]. Bilinguals have also shown advantages in executive function (specifically, selective attention and inhibitory control) across the lifespan [19–23], with evidence suggesting that these processes emerge earlier in bilingual children, compared to age-matched monolingual peers [21]. These studies have shown that bilingual children are less distracted by irrelevant stimulus features compared to monolinguals, and in turn demonstrate faster or more accurate identification of target stimulus features [20, 24]. Finally, studies investigating verbal fluency and lexical retrieval often report poorer performance in bilinguals relative to monolingual peers [25–28]. The exact reason for the corroborated bilingual deficit in verbal fluency and lexical access is unclear. Some studies have suggested that verbal fluency differences are ameliorated between monolinguals and bilinguals when language proficiency is matched between groups [29]. Other studies that have specifically investigated lexical access abilities in bilinguals and monolinguals suggest that poorer performance in bilinguals may arise from greater linguistic processing demands due to competition between two lexicons [30]. This assumption stems from a large body of evidence that indicates that bilinguals co-activate both languages during language comprehension [30, 31].

As reviewed, bilingualism differentially impacts perceptual, cognitive, and linguistic processes. Therefore, it can be argued that a significant advantage or disadvantage in one or more of these underlying perceptual, cognitive (e.g. better executive control), or linguistic processes (e.g. poorer vocabulary knowledge), may contribute to differences in speech perception in noise performance for bilingual listeners compared to monolingual peers. Likewise, performance differences between groups may arise across varied speech-in-noise experimental designs, since speech perception in noise loads on perceptual, cognitive, and linguistic processes in distinct ways through different listening conditions (e.g. various noise levels and noise types).

Bilingualism and audio-only speech perception in noise: Mixed evidence

Prior work has revealed mixed evidence for the impact of bilingualism on audio-only speech-in-noise performance. Some studies have shown that early and simultaneous bilingual listeners exhibit poorer performance on speech-in-noise tasks relative to age-matched monolingual peers [32–34]. For example, one study investigated the impact of early bilingualism on speech perception in noise. While there was no difference between Spanish-English bilingual (age of English acquisition: $\leq$ six yrs.) and monolingual participants' monosyllabic word perception in quiet, bilingual participants exhibited lower performance than monolingual participants in noise conditions. In another study, the Speech Perception in Noise (SPIN) test [35] was used to investigate the impact of speech noise and reverberation on the ability to recognize high- and low-predictable words in five groups of adult listeners who differed in their age of English acquisition [33]. The results revealed that while simultaneous and early bilinguals (age of English acquisition $\leq$ 5–7 yrs.) performed comparably to monolingual controls in mildly

degraded listening conditions, monolinguals outperformed the two groups in more challenging listening conditions. In sum, these findings suggest that differences between monolingual and bilingual performance on speech-in-noise tasks may only be observable in conditions with high task demand (i.e. more challenging listening conditions). The results also indicate that age of English acquisition is a significant predictor of bilingual listeners' identification of degraded target words in various combinations of noise, reverberation [33, 34] and context [33].

In contrast, recent evidence has demonstrated that simultaneous bilinguals exhibit *comparable* performance on speech-in-noise tasks, relative to monolingual age-matched peers [36, 37]. For example, one study found similar performance between English monolingual and simultaneous Spanish-English children (age of English acquisition ≤ 5 yrs.) on a forced-choice, picture-pointing paradigm that assessed speech reception thresholds across both English and Spanish two-talker masking conditions, as well as spectrally matched noise [37]. In another study with young adult listeners similar performance was found between English monolingual and simultaneous Greek–English bilingual listeners (age of English acquisition ≤ 2 yrs.) for recognition of target words across both English and Greek speech masker types [36]. In contrast to studies that revealed poorer speech-in-noise performance for bilinguals [33, 34], these studies [36, 37] suggest that simultaneous bilingualism does not negatively impact speech perception in noise. Several methodological variations can be noted in the type of target stimuli selected (e.g. monosyllabic words, SPIN sentences, BEL sentences), the maskers implemented (e.g. reverberation, spectral noise, various multi-talker babble tracks), as well as the SNR at which the stimuli were presented (e.g. +4 dB to -5 dB SNR).

These conflicting results and methodological variations pose a challenge to clinicians who need to determine the extent to which English speech-in-noise assessment tools can reliably test bilingual listeners with early English language acquisition [38]. Further, with the exception of a single study [37], there is a dearth of evidence for simultaneous bilingual children's performance on English speech recognition in noise tasks.

Bilingualism and audiovisual speech perception in noise: lack of evidence

The existing studies on early and simultaneous bilingual speech-in-noise performance have only considered the contribution of auditory information to speech perception in adverse listening conditions. This is surprising since it is well-established that viewing a speaker's articulatory movements improves speech-in-noise for both monolingual and non-native listeners [39–41]. Currently, the role of visual cues in bilingual speech perception has been examined in infants simultaneously learning two languages [42–44] and early bilingual adult listeners perceiving speech in a non-native language [41, 45–50]. In sum, these studies reveal that although bilingual infants show an audiovisual *advantage* relative to age-matched monolingual counterparts, bilingual adults do not. For example, bilingual infants exhibit better discriminability of silent talking faces [42, 44], and also exploit audiovisual speech cues more than monolingual infants [43]. These observations have led some to propose that bilingual infants may capitalize on audiovisual cues more than monolinguals to disambiguate their native languages as they simultaneously acquire two languages [51]. In contrast, studies that have investigated audiovisual phoneme identification in bilingual adults have shown that although monolingual and bilingual listeners both benefit from visual cues, monolingual adults outperform bilingual age-matched peers [50]. The evidence here suggests that reliance on audiovisual cues may change as a function of age in bilinguals. The use of audiovisual speech cues should be investigated in *school-age children* to gain a better understanding of this developmental change in reliance on visual cues in bilinguals since evidence is current lacking for this age group.

Bilingualism and audiovisual speech perception in noise: Study design and motivation

As reviewed, prior studies have yielded mixed results regarding bilingual speech perception in noise performance across different experimental designs. While this body of evidence is important, these studies have not examined audiovisual speech recognition in noise for bilingual listeners. Expanding on previous research, we investigated the effect of simultaneous bilingualism through a comprehensive speech-in-noise battery that assessed sentence recognition across: different modalities [audio-only (AO), audiovisual (AV)], since monolingual and bilingual listeners have been found to rely on visual cues differently across the lifespan; and two masker types that distinctively engage cognitive processes (energetic masker: steady-state pink noise; informational masker: two-talker babble). Perceiving speech in competing information maskers is more critically dependent upon executive processes, such as selective attention and inhibitory control [9], cognitive processes for which bilingual children have demonstrated advantages [20, 23, 52]. Finally, our comprehensive speech-in-noise battery assessed sentence recognition across a more expanded range of signal-to-noise ratios (SNRs; 0 to -16 dB) in school-age children. The implementation of more challenging SNRs has been found to tax the auditory system in such a way to elicit individual differences in the breakdown of speech sound encoding that may not otherwise occur in more favorable listening conditions [53, 54]. To control for linguistic bias between groups, we recruited a group of simultaneous bilingual children with similar linguistic backgrounds as similultaneous bilingual participants reported in [37] for: English language onset (age of English acquisition ≤3 years), as well as proficiency and usage of both languages (see Table 1). We also ensured that our two groups were matched on socioeconomic status (SES), since SES and bilingualism are found to contribute independently to cognitive and linguistic development [23, 55]. Prior studies have not matched monolingual and bilingual groups on SES [6, 7, 33, 34, 37]. Finally, we utilized the BEL sentences, which were developed to have simpler English lexicon and syntax compared to

Table 1. Age of English acquisition (AoEA), daily language usage, and reported language proficiency of the twelve bilingual participants.

Participant	Age at testing (years)	Sex	AoEA (months)	% of hours of use (per day)		Speaking Proficiency		Comprehension Proficiency	
				Other Language	English	Other Language	English	Other Language	English
1	6	Female	18	25	75	4	5	5	5
2	7	Female	12	30	70	5	5	4	5
3	6	Male	20	30	70	5	5	5	5
4	7	Male	36	35	65	5	5	4	4
5	7	Male	36	20	80	5	5	5	5
6	8	Female	0	25	75	4	5	4	5
7	10	Female	0	25	75	4	5	4	5
8	6	Male	0	30	70	5	5	5	5
9	7	Female	0	25	75	3	5	4	5
10	8	Female	0	25	75	4	5	4	5
11	9	Female	36	40	60	4	4	4	4
12	7	Female	36	40	60	4	4	4	4

Note. Parents reported the estimated percentage of hours per day that both languages were used within the past 24 months. Speaking Proficiency: parents reported the extent to which their child could be understood in English and the other language on the following scale: 1 = *does not speak any;* 2 = *a couple of words or phrases*; 3 = *can have a simple conversation;* 4 = *frequently;* 5 = *all of the time*. Comprehension Proficiency: parents reported how often their child could understand English and the other language on the following scale: 1 = *does not understand anything*; 2 = *understands a couple of words or phrases*; 3 = *understands basic directions*; 4 = *understands the majority of what they are told*; 5 = *understands everything they are told*.

other standardized speech-in-noise sentences [56]. With this experimental design, we aimed to assess whether performance differences between monolingual and simultaneous bilingual school-age children, if present, are restricted to certain listening conditions when socioeconomic status and age of English acquisition is similar between groups.

Materials and Methods

All materials and procedures were approved by the Institutional Review Board at the University of Texas at Austin. All children, as well as their parents, provided written informed consent before their participation in the study.

Participants

Twenty-four elementary school-age children (age range: 6–10 years) from the Austin community participated in the experiment. Inclusionary criteria consisted of: normal hearing defined as hearing thresholds < 20 dB, normal hearing level (nHL) for octaves from 250 to 8000 Hz, no history or current diagnosis of a speech, language, or neurodevelopmental disorder (confirmed via parent report), and normal intelligence ($M = 100$, $SD = 15$), as measured by the Kaufman Brief Intelligence Test-Second Edition, KBIT-2, matrices subtest [57]; monolinguals: $M = 104.83$, $SD = 17.02$; bilinguals: $M = 118.42$, $SD = 19.51$; [$F(1,22) = 3.30$, $P = 0.08$]. The KBIT-2 matrices subtest assesses non-verbal intelligence through the evaluation of an individual's ability to perceive relationships and complete visual analogies [57]. Monolinguals (n = 12; 6 female) and simultaneous bilinguals (n = 12; 8 female) were matched on age (monolinguals: $M = 7.33$ yrs, $SD = 1.23$ yrs; bilinguals: $M = 7.33$ yrs, $SD = 1.23$ yrs; [$F(1,22) = 0$, $P = 1$]) and family socioeconomic status (Family-SES; monolinguals: $M = 55.75$, $SD = 6.18$; bilinguals: $M = 54.77$, $SD = 14.36$; [$F(1,22) = 0.05$, $P = 0.83$]). The Family-SES score was calculated based on the Four Factor Index of Social Status [58] This metric takes into account parents' occupation, education, sex and marital/cohabitation status and returns a six-tier classification of social strata based on scores ranging from 8 to 66. All participants' family social strata fell into two categories: (a) medium business, minor professional, technical (social stratum range = 40–54), or (b) major business and professional (social stratum range = 55–66). SES is an important factor to control, as SES and bilingualism have both been found to contribute independently to cognitive and linguistic development [23].

Each parent completed a language history questionnaire [59]. Results derived from this language history questionnaire have been found to correlate well with linguistic proficiency [37]. Through the questionnaire parents rated their child's speaking and comprehension proficiency on a scale ranging from 1 to 5 (see Table 1). The questionnaire specified that speaking proficiency referred to how easily the child could be understood in each language, while comprehension proficiency indicated how easily the child could understand each language. The 12 bilingual speakers consisted of 7 Chinese-English, 4 Arabic-English, and 1 English-Spanish participant. All parents of bilingual participants reported that their child was born in the United States, except one participant. The parent of this participant reported that the child was born in Japan, and soon after, the family immigrated to the United States. Five participants were reported to start acquiring both languages from birth, three before 20 months, and four before 36 months of age. All parents reported that their child's use of the other language exceeded 20% throughout the day (see Table 1 for detailed demographics of bilingual participants).

Based on previously reported bilingualism classification, we labeled these children as simultaneous bilinguals. The parents of the 12 monolingual speakers stated that their child began acquiring English from birth and was not exposed to a second language throughout the child's lifespan.

Target Stimuli

Disadvantages for bilinguals, relative to monolinguals, have been found for vocabulary knowledge (for reviews, [4, 22]). Therefore, it is unclear if poor speech-in-noise performance in bilinguals arises because of poor linguistic proficiency in the target language. To control for the influence of linguistic knowledge on speech perception in noise, we selected 80 target sentence stimuli from the Basic English Lexicon [56]. The BEL corpus was developed for native and non-native English speech-recognition testing with the specific aim to select lexical items that would be familiar to non-native listeners who may have limited knowledge of English lexicon and syntax [56]. This complete corpus consists of 20 lists of 25 sentences. The 80 sentences that were selected for the current experiment comprised of simple vocabulary that would be present in an elementary school-age child's lexicon (e.g. "mouse", "cheese", "rabbit", "toy", "tree"). Each sentence contained four keywords for intelligibility scoring. For example, "*The black$_1$ cat$_2$ climbed$_3$ the tree$_4$*." A target word was counted incorrect if the child added or deleted morphemes. For example, if the child reported: "cats" for "cat" or "climb" for "climbed," the target words would be scored as incorrect. One male native speaker of American English was recorded producing the full set of 80 sentences. The audio was recorded at a sampling rate of 48000 Hz. The video and audio for each sentence were segmented and separated, and the audio tracks were equalized for RMS amplitude using Praat [60]. Similar stimuli have been utilized across a range of studies from our lab (e.g. [41, 61, 62]).

Maskers

The protocol for creating the maskers, as well as mixing the target speech with maskers, was additionally consistent with methodology reported in these past publications. There were two masker types created for this experiment: two-talker babble and steady-state pink noise. For the two-talker babble tracks, two male native speakers of American English were recorded in a sound-attenuated booth at Northwestern University as part of the Wildcat Corpus project [63]. Each speaker produced a set of 30 simple, meaningful English sentences (from [64]). All sentences were segmented from the recorded files and equalized for RMS amplitude in Praat [60]. The sentences from each talker were concatenated in random order to create 30-sentence strings without silence between the sentences. Two of these strings were mixed using the mix paste function in Audacity (Version 1.2.5; www.audacity.sourceforge.net) to generate two-talker babble. A second masker track consisting of 10 seconds of steady-state pink noise was created using the Noise Generator option in Audacity 1.2.5. Both masker tracks were equated for RMS amplitude to 50dB, 54dB, 58dB, 62dB and 66dB to create ten discrete, long masker tracks. Each masker track was segmented using Praat to create 80 unique noise clips, for a total of 5 noise clips per target sentence per masker type.

Mixing targets and maskers

Each audio clip was mixed with the five corresponding noise clips from each masker track using Adobe Audition to create ten final stimuli of the same target sentence with the following signal to noise ratios: 0 dB, -4 dB, -8 dB, -12 dB, and -16 dB. The mixed audio clips served as the stimuli for the audio-only condition. For the audiovisual condition, audio clips were reattached to the corresponding videos using Final Cut Pro. A freeze frame of the speaker was displayed during the 500 ms noise leader and 500 ms noise trailer. The final mixed target sentence with masker was approximately 2000 ms. In total, for each noise type, there were 400 final audio files (80 sentences × 5 SNRs) and 400 corresponding audiovisual files (80 sentences × 5 SNRs).

Procedures

Speech perception in noise. Participants completed the experiment in a sound attenuated booth. The sentence stimuli were binaurally presented to participants through Sennheiser HD-280 Pro headphones. A trained research assistant was present to type the verbal responses of each child and to ensure that the child attended to the computer screen through the entire duration of the experiment. Four sentences with 16 keywords for scoring were presented in each combination of masker (stead-state pink noise, two-talker babble), SNR (0 to -16 dB), and presentation modality (AO, AV). The presentation order was randomized. The child was instructed first to listen to the sentence produced by the speaker and then repeat the exact sentence that they heard out loud. The research assistant further informed the child that the speaker would begin talking after the noise started. Finally, the child was instructed that even if they only heard a few words, to say those words out loud, and if they were unsure of what they heard, to make their best guess. If they did not understand any words, they were asked to say 'X.' The speech perception in noise task lasted approximately 15–20 minutes, and the administration of the Kaufman Brief Intelligence Test-Second Edition (KBIT-2) [57] also lasted approximately 15–20 minutes. In sum, the total experimental session lasted approximately 35–40 minutes.

Data Analysis. The speech intelligibility data was analyzed with logistic mixed effects modeling implemented in lme4 package using a binomial logit link in R [65]. This type of analysis models mixed effects logistic regression, where the estimates of the model output correspond to the log odd or probability of producing a correct response. We have utilized mixed effects modeling in past speech-in-noise publications from our laboratory [41, 61, 66]. In the current analyses, participants' target word identification in sentences was coded as "correct" or "incorrect" on each trial. This trial-by-trial accuracy was treated as the dichotomous dependent variable. Differences in child speech-in-noise performance have been reported for speech masked by competing talkers compared to speech masked by steady-state noise [12, 67–69]. To confirm this result, a simple linear regression analysis was conducted to examine differences in the proportion of correct keywords identified between the two masker types in the current study. The analysis indicated a significant effect of masker type, $F(1,46) = 12.84$, $P = <0.001$, $R^2 = 0.22$, $b = 0.04$, suggesting that the proportion of correct keyword identification was higher in the steady-state, pink noise masker condition compared to the two-talker maker condition. Therefore, we conducted two separate mixed effects analyses to investigate language group differences in the two-talker and steady-state pink noise masker conditions. The fixed effects of interest were SNR, modality [AO (reference level) versus AV], and listener group [monolingual (reference level) versus bilingual], and their interactions. SNR was mean-centered and treated as a continuous variable. Modality and listener group were treated as categorical variables. In both models, subjects and sentences were specified as random factors to account for individual differences and potential linguistic variation among sentences, respectively. Two alternative random effects structures were considered to determine the optimal model: (1) by-sentence intercept and by-subject sentence slope, and (2) by-subject and by-sentence intercepts [41]. The first model failed to converge, therefore the second model was used.
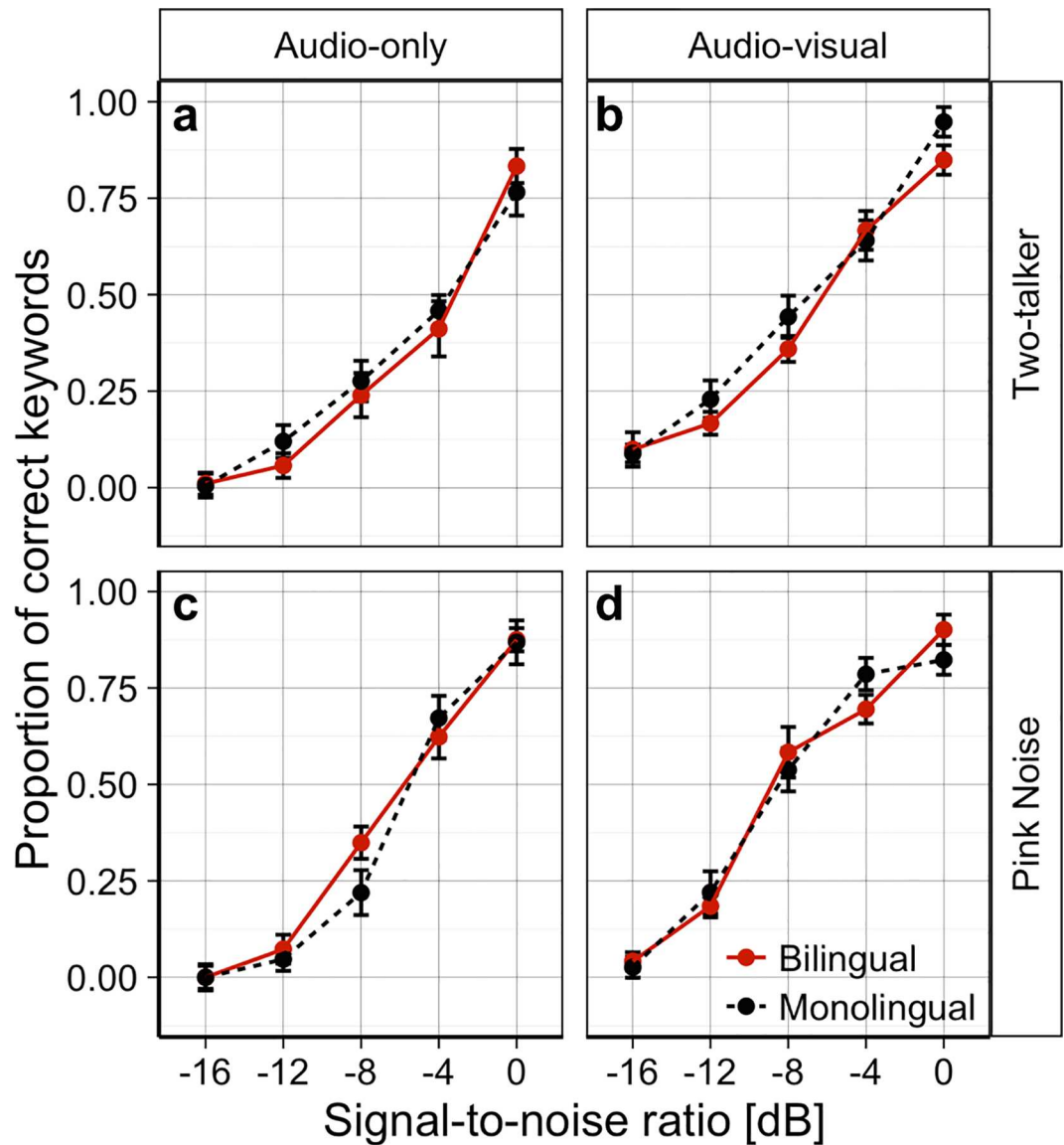
Results

[Fig 1] shows the mean proportion of correctly identified target words as a function of SNR in monolingual and bilingual children across the four masker-modality conditions.

Steady-state pink noise masker

In the mixed-effects model, the intercept was significant, $b = -3.65$, $SE = 0.36$, $Z = -10.02$, $P < 0.001$. The simple effect of language group was not significant, $b = 0.60$, $SE = 0.42$,

Fig 1. Mean proportion of correct keywords as a function of signal-to-noise ratio across all conditions. The top panels show the mean proportion of correct keywords identified by bilingual (red) and monolingual (black dashed) school-age children in the two-talker masker audio-only (a) and audio-visual (b) conditions. The bottom panels show the mean proportion of correct keywords identified by the two groups in the steady-state pink noise masker audio-only (c) and audio-visual (d) conditions. Error bars denote one standard error from the mean.

doi:10.1371/journal.pone.0168048.g001

$Z = 1.42$, $P = 0.156$, indicating that proportion of correctly identified target words in noise were not significantly different between the two language groups. The simple effect of SNR was significant $b = 0.53$, $SE = 0.03$, $Z = 17.02$, $P < 0.0001$, where improving SNR increased the probability of correct target word identification. The simple effect of modality was significant, $b = 1.79$, $SE = 0.26$, $Z = 6.87$, $P < 0.0001$, suggesting that the probability of correct target word identification was significantly higher in AV compared to the AO condition. The SNR by modality interaction was significant, $b = -0.12$, $SE = 0.04$, $Z = -3.28$, $P = 0.001$, indicating that slope of increase for intelligibility benefit from the improvement of SNR was less for the AV condition compared to the AO condition. There were no other significant two- or three-way interactions.

Two-talker masker

In the mixed-effects model, the intercept was significant, b = -3.37, SE = 0.44, Z = -7.65, P < 0.0001. The simple effect of language group was not significant, b = -0.46, SE = 0.56, Z = -0.82, P = 0.414, demonstrating that proportion of correctly identified target words in noise was not significantly different between the two language groups. The simple effect of SNR was significant, b = 0.42, SE = 0.03, Z = 16.10, P < 0.0001, where improving SNR increased the probability of correct target word identification. The simple effect of modality was significant, b = 1.34, SE = 0.24, Z = 5.57, P < 0.0001, suggesting that target words in sentences were more correctly identified in AV relative to AO conditions. There were no significant two- or three-way interactions.

Discussion

Prior studies have found mixed evidence for the impact of bilingualism on speech-in-noise performance. One body of literature has demonstrated that early and simultaneous bilingual listeners demonstrate poorer performance on speech-in-noise tasks relative to age-matched monolingual peers [33, 34]. In contrast, recent studies have found no difference in speech-in-noise performance between monolingual and simultaneous bilingual listeners [36, 37]. Here we examined the extent to which speech-in-noise performance differences, if present, were modulated by certain listening conditions in monolingual and simultaneous bilingual children. To that end, we assessed the impact of simultaneous bilingualism on speech perception through a comprehensive speech-in-noise battery, across different modalities, masker types, and SNRs. Our results revealed no difference in performance between monolingual and simultaneous bilingual children (age of English acquisition ≤ 3) in each combination of presentation modality (AO, AV), masker (steady-state pink noise, two-talker babble), and SNR (0 to -16 dB SNR).

The main effects of the tested factors observed in the current study are in line with previous studies examining speech intelligibility in adverse listening conditions in samples of school-age children. The two-talker masker condition was more difficult for all children, compared to the steady-state pink noise condition [12, 67–69]; all listeners exhibited better keyword identification in sentences masked by easier SNRs [70–73], and AV speech was more intelligible than AO speech [39, 40, 50, 74].

Apart from typical developmental factors and noise, research has demonstrated that the language experience of children further affects their ability to recognize speech-in-noise [7]. Successful recognition of the speech signal from speaker to listener has been described in the literature through signal-dependent and signal-independent factors [9]. Signal-dependent factors are related to the quality of the incoming speech signal (e.g. the level and type of background noise). In contrast, signal-independent factors refer to characteristics internal to the listener (e.g. the linguistic proficiency of the listener or familiarity with the content being conveyed by the speaker) [9, 75]. In the current study, our findings indicate that when signal-independent factors are similar between monolinguals and bilinguals, such as comparable early exposure to the target language and socioeconomic status, both groups exhibit the same degree of speech-in-noise difficulties across a range of signal-dependent factors (i.e. different noise types and noise levels). SES is another important signal-independent factor that was matched between groups in the current study. The fact that SES was similar between groups may be another reason the current study did not yield poorer speech-in-noise performance for bilingual participants; a finding that was previously shown in speech-in-noise studies that did not control for SES [6, 7, 33, 34]. This assumption is supported by recent evidence that performance on linguistic and cognitive tasks is uniquely influenced by SES and bilingualism [23].

The current results are in contrast to prior studies that have demonstrated poorer audio-only speech-in-noise performance in early and simultaneous bilingual performance relative to

monolingual peers [32–34]. The discrepancy in findings across studies may be due to differences in target stimuli used. Studies that have shown lower speech-in-noise performance in early and simultaneous bilinguals have utilized different English monosyllabic word and sentence stimuli [32–34] from studies that did not demonstrate differences in speech-in-noise performance between groups [36, 37]. For example, some studies [32, 33] implemented English sentences from the Speech-Perception-in-Noise (SPIN) test [35, 76], while others [37] employed target sentences from the Basic English Lexicon (BEL) corpus. Like [37], the present study implemented target BEL sentences across different masking conditions and did not find differences between monolingual and simultaneous bilingual school-age children. The findings from the current study suggest that the simpler lexicon and grammar in BEL sentences may eliminate linguistic bias in the testing of bilingual children simultaneously learning two languages, compared to other available speech-in-noise test materials. However, one limitation of the present study is that there was no direct comparison between the perception of BEL sentences and the perception of sentences from another standardized speech-in-noise measure. Future studies should empirically compare speech-in-noise abilities in simultaneous bilingual children perceiving BEL sentences in a comprehensive speech-in-noise battery, relative to the perception of other clinically implemented English speech-in-noise sentences.

The cross-study differences in early and simultaneous bilingual speech-in-noise performance may also be due to differences in English language learning environments across bilingual participant samples. For example, the majority of the simultaneous bilingual children in the current study, like the majority of participants in studies that also did not show bilingual disadvantages for speech-in-noise [36, 37], were born and raised in the United States. The majority of other bilingual speech-in-noise studies did not report the country of birth of their participants [33, 34]. Therefore, it is unclear whether the bilingual participants who took part in these studies were born and raised in the United States, or whether they emigrated from another country. If the subjects emigrated from another country, even at an early age, this would have an impact on the English language development of the participants. This is because it is well-established that the amount of language input strongly affects the rate of language growth for both monolingual and bilingual children [55, 77, 78].

Our results across different masker types in the AO conditions are in line with recent work that has demonstrated no difference in speech recognition performance for simultaneous bilinguals compared to age-matched monolingual peers across both speech (i.e. two-talker babble) and steady-state noise [36, 37]. Early bilinguals' lower performance on speech perception in steady-state noise has been explained as the outcome of increased linguistic processing demand [34]. These linguistic demands have been suggested to arise from competition between two co-activated lexicons, and in turn, lead to difficulties in sound-to-meaning mapping in challenging listening environments. Such increased linguistic demands would also predict poorer speech perception in speech noise bilingual performance, relative to monolingual peers. On the other hand, bilingual advantages in cognitive processes, such as selective attention and inhibitory control, should lead to better performance for bilingual speech-in-noise–especially, in perceiving speech in competing talkers. Speech perception in speech noise critically depends on ignoring irrelevant information while selecting target information, cognitive abilities that have been found to be enhanced in bilingual children [20, 24]. This hypothesis is further supported by the ease of language understanding model (ELU) [10], which posits that individual differences in the quality of speaker-internal working memory capacity contributes to individual differences in the recovery of degraded acoustic information in noise. [10].

The current findings did not reveal either of the previously observed differences between language groups, even at more difficult SNRs. Similar bilingual and monolingual performance demonstrated in the two-talker babble conditions in the current study may be evidence for an

optimum point between increased linguistic demands, and enhanced executive function in simultaneous bilinguals. This finding may also suggest that the bilingual advantage in executive function, which has been mostly demonstrated through non-linguistic tasks [20], does not generalize to the specific task demands of perceiving speech in noise. This idea is in line with a recent review which concluded that managing two languages does not result in general executive function advantages. The review instead suggests that a bilingual advantage in executive function may not exist at all or may be confined to very specific task-dependent circumstances [79]. Although the current experimental design presented sentences in both steady-state noise and two-talker babble across a range of more challenging SNRs compared to those implemented in past studies, it is likely that this task was still not sensitive enough to elicit either of the previously observed differences between language groups. For example, prior studies that have demonstrated poorer performance in early and simultaneous bilinguals due to other environmental acoustic factors (e.g. reverberation). Therefore, additional research is needed to better understand the extent to which other aspects of speech in noise processing modulate differences between monolingual and simultaneous bilingual listeners.

To the best of our knowledge, the current study is the first to evaluate the contribution of visual cues during speech perception in noise in school-age simultaneous bilingual children. We assessed the extent to which simultaneous bilingual children would use visual cues differently during speech-in-noise tasks, compared to age-matched monolingual peers. Our results demonstrated that simultaneous bilingual and monolingual children benefited equally from visual cues during speech perception in noise. While bilingual infants have been found to exploit audiovisual speech cues [43], bilingual adults have exhibited less reliance on audiovisual phoneme discrimination tasks, compared to monolingual peers [50]. In considering the current evidence with these prior findings, it may be suggested that as bilinguals develop from infancy to early childhood and become proficient in both languages, the need to capitalize on audiovisual cues to discriminate their two languages becomes less necessary [80].

Similar to communicating across two languages, experience playing a musical instrument has also been found to lead to enhanced perceptual and cognitive abilities [81–83]. For example, enhanced neural responses to pitch changes during speech processing has been demonstrated for musicians relative to nonmusicians [84, 85]. These observed enhancements in musicians have been found to be contingent upon the extent of musical expertise and training [84, 86], just as observed bilingual benefits are dependent upon the degree of proficiency and usage of both languages. Like bilingualism, there is mixed evidence for the transfer of enhanced perceptual and cognitive skills, resulting from music training, to enhanced speech-in-noise [87–91]. Speech-in-noise benefit for musicians has been suggested to be task-dependent, with musicianship advantages found to emerge more in informational masking conditions relative to energetic masking conditions [87, 92], but not always [88]. The fact that different types and levels of auditory experiences result in different levels of speech-in-noise outcomes highlights the need for more carefully defined participant groups, and implementation of comprehensive speech-in-noise test batteries, with a range of masker types, modalities, and SNRs. These methodological considerations will allow for a better understanding of the interactions between varying degrees and types of auditory expertise (signal-independent factors) and speech perception in noise.

While the current study provides new insights into simultaneous bilingual performance on speech-in-noise, a few limitations remain. First, while a large body of evidence has shown that bilinguals demonstrate advantages in executive function (specifically, selective attention and inhibitory control) [19–23], we did not implement any cognitive measure to confirm these findings and relate them to speech-in-noise performance. Future work should investigate the links between bilingualism in school-age children and measures of executive function to

provide better insight into how these cognitive enhancements, if present, relate to individual differences in speech perception in noise. Second, while we carefully controlled for differences in age, socioeconomic status, and non-verbal intelligence between the monolingual and bilingual groups, we did not investigate language-specific bilingual influences on speech-in-noise performance. Future studies should explore the extent to which differences in English speech perception in noise occur as a function of first language background (e.g. tonal vs. non-tonal language). A third limitation is the generalizability of these findings to the classroom setting. In the current study, we did not observe differences between monolingual and simultaneous bilingual children's speech perception in noise across a range of listening conditions, using high-frequency target stimuli from the BEL corpus. The current study lays the foundation for future work to pursue more classroom oriented investigations, such as the extent to which monolinguals and bilinguals differ on novel word learning in classroom noise.

Conclusion

In conclusion, our results indicate that when the age of English acquisition and socioeconomic status are similar between groups, monolingual and simultaneous bilingual children exhibit the same degree of speech-in-noise difficulties across a range of adverse listening conditions. For these simultaneous bilingual listeners, bilingualism does not negatively affect English speech recognition across a combination of masker (steady-state pink noise, two-talker babble), SNR (0 to -16 dB), and presentation modality (AO, AV). These findings suggest that despite increasing linguistic demands that may arise from two competing lexicons, simultaneous bilingual listeners have the ability to perform equally to monolinguals on the identification of degraded target words in masked sentences.

Acknowledgments

Author Contributions

Conceptualization: RR LS BC.

Data curation: RR ZX.

Formal analysis: RR BPWL ZX.

Funding acquisition: RR LS BC.

Investigation: RR.

Methodology: RR LS BC.

Project administration: RR LS BC.

Resources: RR LS BC.

Software: RR ZX BC.

Supervision: RR LS BC.

Validation: RR BPWL ZX.

Visualization: RR.

Writing – original draft: RR BC.

Writing – review & editing: RR LS BC BPWL ZX.

References

1. Shin HB, Ortman JM, editors. Language projections: 2010 to 2020. Federal Forecasters Conference, April; 2011.

2. Krizman J, Skoe E, Marian V, Kraus N. Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. Brain and language. 2014; 128(1):34–40. doi: 10.1016/j.bandl.2013.11.006 PMID: 24413593

3. Krizman J, Slater J, Skoe E, Marian V, Kraus N. Neural processing of speech in children is influenced by extent of bilingual experience. Neuroscience letters. 2015; 585:48–53. doi: 10.1016/j.neulet.2014.11.011 PMID: 25445377

4. Bialystok E. Bilingualism: The good, the bad, and the indifferent. Bilingualism: Language and cognition. 2009; 12(01):3–11.

5. Nicoladis E, Genesee F. Language development in preschool bilingual children.[L'apprentissage du langage chez les enfants bilingues d'age prescolaire]. Journal of Speech Language-Pathology and Audiology. 1997; 21(4):258–70.

6. Crandell CC, Smaldino JJ. Speech perception in noise by children for whom English is a second language. American Journal of Audiology. 1996; 5(3):47–51.

7. Nelson P, Kohnert K, Sabur S, Shaw D. Classroom Noise and Children Learning Through a Second LanguageDouble Jeopardy? Language, Speech, and Hearing Services in Schools. 2005; 36(3):219–29. PMID: 16175885

8. Shinn-Cunningham BG. Object-based auditory and visual attention. Trends in cognitive sciences. 2008; 12(5):182–6. doi: 10.1016/j.tics.2008.02.003 PMID: 18396091

9. Mattys SL, Davis MH, Bradlow AR, Scott SK. Speech recognition in adverse conditions: A review. Language and Cognitive Processes. 2012; 27(7–8):953–78.

10. Rönnberg J, Lunner T, Zekveld A, Sörqvist P, Danielsson H, Lyxell B, et al. The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. 2013.

11. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive psychology. 2000; 41(1):49–100. doi: 10.1006/cogp.1999.0734 PMID: 10945922

12. Baker M, Buss E, Jacks A, Taylor C, Leibold LJ. Children's perception of speech produced in a two-talker background. Journal of Speech, Language, and Hearing Research. 2014; 57(1):327–37. doi: 10.1044/1092-4388(2013/12-0287) PMID: 24687476

13. Tun PA, O'Kane G, Wingfield A. Distraction by competing speech in young and older adult listeners. Psychology and aging. 2002; 17(3):453. PMID: 12243387

14. Cooke M, Lecumberri MG, Barker J. The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. The Journal of the Acoustical Society of America. 2008; 123(1):414–27. doi: 10.1121/1.2804952 PMID: 18177170

15. Conway AR, Cowan N, Bunting MF. The cocktail party phenomenon revisited: The importance of working memory capacity. Psychonomic bulletin & review. 2001; 8(2):331–5.

16. Bell NL, Lassiter KS, Matthews TD, Hutchinson MB. Comparison of the peabody picture vocabulary test—Third edition and Wechsler adult intelligence scale—Third edition with university students. Journal of clinical psychology. 2001; 57(3):417–22. PMID: 11241372

17. Benard MR, Mensink JS, Başkent D. Individual differences in top-down restoration of interrupted speech: Links to linguistic and cognitive abilities. The Journal of the Acoustical Society of America. 2014; 135(2):EL88–EL94. doi: 10.1121/1.4862879 PMID: 25234920

18. Krizman J, Marian V, Shook A, Skoe E, Kraus N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proceedings of the National Academy of Sciences. 2012; 109(20):7877–81.

19. Bialystok E, Poarch G, Luo L, Craik FI. Effects of bilingualism and aging on executive function and working memory. Psychology and aging. 2014; 29(3):696. doi: 10.1037/a0037254 PMID: 25244487

20. Poulin-Dubois D, Blaye A, Coutya J, Bialystok E. The effects of bilingualism on toddlers' executive functioning. Journal of experimental child psychology. 2011; 108(3):567–79. doi: 10.1016/j.jecp.2010.10.009 PMID: 21122877

21. Bialystok E. Bilingualism in development: Language, literacy, and cognition: Cambridge University Press; 2001.

22. Bialystok E, Craik FI, Luk G. Bilingualism: consequences for mind and brain. Trends in cognitive sciences. 2012; 16(4):240–50. doi: 10.1016/j.tics.2012.03.001 PMID: 22464592

23. Calvo A, Bialystok E. Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. Cognition. 2014; 130(3):278–88. doi: 10.1016/j.cognition.2013.11.015 PMID: 24374020

24. Kapa LL, Colombo J. Attentional control in early and later bilingual children. Cognitive development. 2013; 28(3):233–46. doi: 10.1016/j.cogdev.2013.01.011 PMID: 24910499

25. Portocarrero JS, Burright RG, Donovick PJ. Vocabulary and verbal fluency of bilingual and monolingual college students. Archives of Clinical Neuropsychology. 2007; 22(3):415–22. doi: 10.1016/j.acn.2007.01.015 PMID: 17336036

26. Bialystok E, Luk G, Peets KF, Yang S. Receptive vocabulary differences in monolingual and bilingual children. Bilingualism: Language and Cognition. 2010; 13(04):525–31.

27. Gollan TH, Montoya RI, Fennema-Notestine C, Morris SK. Bilingualism affects picture naming but not picture classification. Memory & Cognition. 2005; 33(7):1220–34.

28. Van Heuven WJ, Dijkstra T, Grainger J. Orthographic neighborhood effects in bilingual word recognition. Journal of memory and language. 1998; 39(3):458–83.

29. Gullifer JW, Kroll JF, Dussias PE. When language switching has no apparent cost: Lexical access in sentence context. Frontiers in psychology. 2013; 4.

30. Blumenfeld HK, Marian V. Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. Language and cognitive processes. 2007; 22(5):633–60.

31. Marian V, Spivey M. Competing activation in bilingual language processing: Within-and between-language competition. Bilingualism: Language and Cognition. 2003; 6(02):97–115.

32. Mayo LH, Florentine M, Buus S. Age of second-language acquisition and perception of speech in noise. Journal of speech, language, and hearing research. 1997; 40(3):686–93. PMID: 9210123

33. Shi L-F. Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. Journal of Speech, Language, and Hearing Research. 2010; 53(4):821–35. doi: 10.1044/1092-4388(2010/09-0081) PMID: 20220026

34. Rogers CL, Lister JJ, Febo DM, Besing JM, Abrams HB. Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. Applied Psycholinguistics. 2006; 27(03):465–85.

35. Bilger RC, Nuetzel J, Rabinowitz W, Rzeczkowski C. Standardization of a test of speech perception in noise. Journal of Speech, Language, and Hearing Research. 1984; 27(1):32–48.

36. Calandruccio L, Zhou H. 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. 2014; 57(3):1089–97. doi: 10.1044/2013_JSLHR-H-12-0378 PMID: 24167230

37. Calandruccio L, Gomez B, Buss E, Leibold LJ. Development and Preliminary Evaluation of a Pediatric Spanish–English Speech Perception Task. American journal of audiology. 2014; 23(2):158–72. doi: 10.1044/2014_AJA-13-0055 PMID: 24686915

38. Shi L-F. How "Proficient" Is Proficient? Bilingual Listeners' Recognition of English Words in Noise. American journal of audiology. 2015; 24(1):53–65. doi: 10.1044/2014_AJA-14-0041 PMID: 25551364

39. Ross LA, Molholm S, Blanco D, Gomez-Ramirez M, Saint-Amour D, Foxe JJ. The development of multisensory speech perception continues into the late childhood years. European Journal of Neuroscience. 2011; 33(12):2329–37. doi: 10.1111/j.1460-9568.2011.07685.x PMID: 21615556

40. Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cerebral Cortex. 2007; 17(5):1147–53. doi: 10.1093/cercor/bhl024 PMID: 16785256

41. Xie Z, Yi H-G, Chandrasekaran B. Nonnative Audiovisual Speech Perception in Noise: Dissociable Effects of the Speaker and Listener. PloS one. 2014; 9(12):e114439. doi: 10.1371/journal.pone.0114439 PMID: 25474650

42. Sebastián-Gallés N, Albareda-Castellot B, Weikum WM, Werker JF. A bilingual advantage in visual language discrimination in infancy. Psychological Science. 2012; 23(9):994–9. doi: 10.1177/0956797612436817 PMID: 22810164

43. Pons F, Bosch L, Lewkowicz DJ. Bilingualism modulates infants' selective attention to the mouth of a talking face. Psychological science. 2015:0956797614568320.

44. Weikum WM, Vouloumanos A, Navarra J, Soto-Faraco S, Sebastián-Gallés N, Werker JF. Visual language discrimination in infancy. Science. 2007; 316(5828):1159–. doi: 10.1126/science.1137686 PMID: 17525331

45. Hazan V, Sennema A, Faulkner A, Ortega-Llebaria M, Iba M, Chung H. The use of visual cues in the perception of non-native consonant contrastsa). The Journal of the Acoustical Society of America. 2006; 119(3):1740–51. PMID: 16583916

46. Hazan V, Kim J, Chen Y. Audiovisual perception in adverse conditions: Language, speaker and listener effects. Speech Communication. 2010; 52(11):996–1009.

47. Chen Y, Hazan V. Developmental factors and the non-native speaker effect in auditory-visual speech perceptiona). The Journal of the Acoustical Society of America. 2009; 126(2):858–65. doi: 10.1121/1.3158823 PMID: 19640050

48. Wang Y, Behne DM, Jiang H. Influence of native language phonetic system on audio-visual speech perception. Journal of Phonetics. 2009; 37(3):344–56.

49. Wang Y, Behne DM, Jiang H. Linguistic experience and audio-visual perception of non-native fricatives. The Journal of the Acoustical Society of America. 2008; 124(3):1716–26. doi: 10.1121/1.2956483 PMID: 19045662

50. Burfin S, Pascalis O, Tada ER, Costa A, Savariaux C, Kandel S. Bilingualism affects audiovisual phoneme identification. Frontiers in psychology. 2014; 5.

51. Werker JF, Byers-Heinlein K, Fennell CT. Bilingual beginnings to learning words. Philosophical Transactions of the Royal Society of London B: Biological Sciences. 2009; 364(1536):3649–63. doi: 10.1098/rstb.2009.0105 PMID: 19933138

52. Bialystok E, Martin MM. Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. Developmental science. 2004; 7(3):325–39. PMID: 15595373

53. Wong PC, Uppunda AK, Parrish TB, Dhar S. Cortical mechanisms of speech perception in noise. Journal of Speech, Language, and Hearing Research. 2008; 51(4):1026–41. doi: 10.1044/1092-4388(2008/075) PMID: 18658069

54. Cunningham J, Nicol T, Zecker SG, Bradlow A, Kraus N. Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. Clinical Neurophysiology. 2001; 112(5):758–67. PMID: 11336890

55. Hoff E. The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. Child development. 2003; 74(5):1368–78. PMID: 14552403

56. Calandruccio L, Smiljanic R. New sentence recognition materials developed using a basic non-native English lexicon. Journal of Speech, Language, and Hearing Research. 2012; 55(5):1342–55. doi: 10.1044/1092-4388(2012/11-0260) PMID: 22411279

57. Kaufman AS, Kaufman NL. Kaufman Brief Intelligence Test-Second Edition. Bloomington, MN: Pearson; 2004.

58. Hollingshead AB. Four Factor Index of Social Status. Yale Journal of Sociology. 2001;(8:):21–52.

59. Gutiérrez-Clellen VF, Kreiter J. Understanding child bilingual acquisition using parent and teacher reports. Applied Psycholinguistics. 2003; 24(02):267–88.

60. Boersma P. Praat, a system for doing phonetics by computer. Glot international. 2002; 5(9/10):341–5.

61. Smayda KE, Van Engen KJ, Maddox WT, Chandrasekaran B. Audio-Visual and Meaningful Semantic Context Enhancements in Older and Younger Adults. PloS one. 2016; 11(3):e0152773. doi: 10.1371/journal.pone.0152773 PMID: 27031343

62. Van Engen KJ, Phelps JE, Smiljanic R, Chandrasekaran B. Enhancing speech intelligibility: interactions among context, modality, speech style, and masker. Journal of Speech, Language, and Hearing Research. 2014; 57(5):1908–18. doi: 10.1044/JSLHR-H-13-0076 PMID: 24687206

63. Van Engen KJ, Baese-Berk M, Baker RE, Choi A, Kim M, Bradlow AR. The Wildcat Corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. Language and speech. 2010; 53(4):510–40.

64. Bradlow AR, Alexander JA. Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. The Journal of the Acoustical Society of America. 2007; 121(4):2339–49. PMID: 17471746

65. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013. Document freely available on the internet at: http://www.r-project.org 2015: ISBN 3-900051-07-0; 2015.

66. Yi H-G, Phelps JE, Smiljanic R, Chandrasekaran B. Reduced efficiency of audiovisual integration for nonnative speech. The Journal of the Acoustical Society of America. 2013; 134(5):EL387–EL93. doi: 10.1121/1.4822320 PMID: 24181980

67. Fallon M, Trehub SE, Schneider BA. Children's perception of speech in multitalker babble. The Journal of the Acoustical Society of America. 2000; 108(6):3023–9. PMID: 11144594

68. Leibold LJ, Buss E. Children's identification of consonants in a speech-shaped noise or a two-talker masker. Journal of Speech, Language, and Hearing Research. 2013; 56(4):1144–55. doi: 10.1044/1092-4388(2012/12-0011) PMID: 23785181

69. Brungart DS, Chang PS, Simpson BD, Wang D. Multitalker speech perception with ideal time-frequency segregation: Effects of voice characteristics and number of talkers. The Journal of the Acoustical Society of America. 2009; 125(6):4006–22. doi: 10.1121/1.3117686 PMID: 19507982

70. Papso CF, Blood IM. Word recognition skills of children and adults in background noise. Ear and Hearing. 1989; 10(4):235–6. PMID: 2776983

71. Plomp R, Mimpen A. Speech-reception threshold for sentences as a function of age and noise level. The Journal of the Acoustical Society of America. 1979; 66(5):1333–42. PMID: 500971

72. Bradley J, Reich R, Norcross S. On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. The Journal of the Acoustical Society of America. 1999; 106(4):1820–8.

73. Schafer EC, Beeler S, Ramos H, Morais M, Monzingo J, Algier K. Developmental effects and spatial hearing in young children with normal-hearing sensitivity. Ear and hearing. 2012; 33(6):e32–e43. doi: 10.1097/AUD.0b013e318258c616 PMID: 22688920

74. Sumby WH, Pollack I. Visual contribution to speech intelligibility in noise. The journal of the acoustical society of america. 1954; 26(2):212–5.

75. Lindblom B. On the communication process: Speaker-listener interaction and the development of speech*. Augmentative and Alternative Communication. 1990; 6(4):220–30.

76. Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. The Journal of the Acoustical Society of America. 1977; 61(5):1337–51. PMID: 881487

77. Thordardottir E. The relationship between bilingual exposure and vocabulary development. International Journal of Bilingualism. 2011; 15(4):426–45.

78. Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T. Early vocabulary growth: Relation to language input and gender. Developmental psychology. 1991; 27(2):236.

79. Paap KR, Johnson HA, Sawi O. Should the search for bilingual advantages in executive functioning continue. cortex. 2016; 74:305–14. doi: 10.1016/j.cortex.2015.09.010 PMID: 26586100

80. Pons F, Lewkowicz DJ, Soto-Faraco S, Sebastián-Gallés N. Narrowing of intersensory speech perception in infancy. Proceedings of the National Academy of Sciences. 2009; 106(26):10598–602.

81. Kraus N, Chandrasekaran B. Music training for the development of auditory skills. Nature Reviews Neuroscience. 2010; 11(8):599–605. doi: 10.1038/nrn2882 PMID: 20648064

82. Schellenberg EG, Peretz I. Music, language and cognition: unresolved issues. Trends in cognitive sciences. 2008; 12(2):45–6. doi: 10.1016/j.tics.2007.11.005 PMID: 18178126

83. Kühnis J, Elmer S, Jäncke L. Auditory evoked responses in musicians during passive vowel listening are modulated by functional connectivity between bilateral auditory-related brain regions. Journal of cognitive neuroscience. 2014.

84. Besson M, Schön D, Moreno S, Santos A, Magne C. Influence of musical expertise and musical training on pitch processing in music and language. Restorative neurology and neuroscience. 2007; 25(3–4):399–410. PMID: 17943015

85. Magne C, Schön D, Besson M. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. Journal of Cognitive Neuroscience. 2006; 18(2):199–211. doi: 10.1162/089892906775783660 PMID: 16494681

86. Strait DL, Kraus N, Skoe E, Ashley R. Musical experience and neural efficiency–effects of training on subcortical processing of vocal expressions of emotion. European Journal of Neuroscience. 2009; 29(3):661–8. doi: 10.1111/j.1460-9568.2009.06617.x PMID: 19222564

87. Swaminathan J, Mason CR, Streeter TM, Best V, Kidd G Jr, Patel AD. Musical training, individual differences and the cocktail party problem. Scientific reports. 2015; 5.

88. Boebinger D, Evans S, Rosen S, Lima CF, Manly T, Scott SK. Musicians and non-musicians are equally adept at perceiving masked speech. The Journal of the Acoustical Society of America. 2015; 137 (1):378–87. doi: 10.1121/1.4904537 PMID: 25618067

89. Ruggles DR, Freyman RL, Oxenham AJ. Influence of musical training on understanding voiced and whispered speech in noise. PloS one. 2014; 9(1):e86980. doi: 10.1371/journal.pone.0086980 PMID: 24489819

90. Parbery-Clark A, Skoe E, Lam C, Kraus N. Musician enhancement for speech-in-noise. Ear and hearing. 2009; 30(6):653–61. doi: 10.1097/AUD.0b013e3181b412e9 PMID: 19734788

91. Fuller CD, Galvin JJ, Maat B, Free RH, Başkent D. The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? Frontiers in neuroscience. 2014; 8:179. doi: 10.3389/fnins.2014.00179 PMID: 25071428

92. Başkent D, Gaudrain E. Musician advantage for speech-on-speech perception. The Journal of the Acoustical Society of America. 2016; 139(3):EL51–EL6. doi: 10.1121/1.4942628 PMID: 27036287