



Published in final edited form as:

Ear Hear. 2020 ; 41(2): 259–267. doi:10.1097/AUD.0000000000000752.

Masking Release for Speech-in-Speech Recognition Due to a Target/Masker Sex Mismatch in Children with Hearing Loss

Lori J. Leibold¹, Jenna M. Browning¹, Emily Buss²

¹Center for Hearing Research, Boys Town National Research Hospital, Omaha, Nebraska, USA

²Department of Otolaryngology/Head and Neck Surgery, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Abstract

Objectives: The goal of the present study was to compare the extent to which children with hearing loss and children with normal hearing benefit from mismatches in target/masker sex in the context of speech-in-speech recognition. It was hypothesized that children with hearing loss experience a smaller target/masker sex mismatch benefit relative to children with normal hearing due to impairments in peripheral encoding, variable access to high-quality auditory input, or both.

Design: Eighteen school-age children with sensorineural hearing loss (7-15 years) and 18 age-matched children with normal hearing participated in this study. Children with hearing loss were bilateral hearing aid users. Severity of hearing loss ranged from mild to severe across participants, but most had mild to moderate hearing loss. Speech recognition thresholds for disyllabic words presented in a two-talker speech masker were estimated in the sound field using an adaptive, forced-choice procedure with a picture-pointing response. Participants were tested in each of four conditions: (1) male target speech/two-male-talker masker; (2) male target speech/two-female-talker masker; (3) female target speech/two-female-talker masker; and (4) female target speech/two-male-talker masker. Children with hearing loss were tested wearing their personal hearing aids at user settings.

Results: Both groups of children showed a sex-mismatch benefit, requiring a more advantageous signal-to-noise ratio when the target and masker were matched in sex than when they were mismatched. However, the magnitude of sex-mismatch benefit was significantly reduced for children with hearing loss relative to age-matched children with normal hearing. There was no effect of child age on the magnitude of sex-mismatch benefit. The sex-mismatch benefit was larger for male target speech than for female target speech. For children with hearing loss, the magnitude of sex-mismatch benefit was not associated with degree of hearing loss or aided audibility.

Conclusions: The findings from the present study indicate that children with sensorineural hearing loss are able to capitalize on acoustic differences between speech produced by male and female talkers when asked to recognize target words in a competing speech masker. However, children with hearing loss experienced a smaller benefit relative to their peers with normal hearing. No association between the sex-mismatch benefit and measures of unaided thresholds or

aided audibility were observed for children with hearing loss, suggesting that reduced peripheral encoding is not the only factor responsible for the smaller sex-mismatch benefit relative to children with normal hearing.

INTRODUCTION

Children with hearing loss are often expected to listen to the speech of one talker in the presence of speech produced by competing talkers. This challenging task depends upon the ability to separate streams of speech into distinct auditory objects and then focus attention on the target speech stream of interest while ignoring the competing streams (Bregman 1990). Due to impairments in peripheral encoding as well as variable access to high-quality auditory input (e.g., Walker et al. 2013; 2015), most children with hearing loss perform more poorly than age-matched children with normal hearing on measures of speech-in-speech recognition, even when children with hearing loss are tested wearing appropriately fitted hearing aids (e.g., Leibold et al. 2013). This performance gap between children with hearing loss and children with normal hearing has both theoretical and practical implications, given that competing speech pervades children's everyday environments (e.g., Ambrose et al. 2014). Nonetheless, the factors responsible for the increased susceptibility to speech-in-speech masking experienced by children with hearing loss relative to children with normal hearing are not well understood. Do children with hearing loss rely on the same acoustic differences between target and masker speech used by children with normal hearing to separate streams of speech, albeit less effectively? Alternatively, does hearing loss impact the acoustic cues that would otherwise support the segregation of target and masker speech or the recognition of degraded speech?

To begin to address these questions, the present study compared the extent to which children with hearing loss and children with normal hearing benefit from a mismatch in target/masker sex in the context of speech-in-speech recognition. Males and females tend to differ with respect to the mass, length, and tension of the vocal folds, as well as the resonant frequencies of the vocal tract and associated cavities (e.g., Fitch and Giedd, 1999). As a result of these structural differences, the average fundamental frequency (F_0 ; corresponding to the rate of vocal fold vibration) is about 210 Hz for adult females compared with 125 Hz for adult males (e.g., Rendall et al. 2005). Adult females tend to have shorter vocal tracts than adult males, influencing both formant frequencies and the overall spectral envelope of speech (e.g., Fitch and Giedd 1999; Smith and Patterson 2005). Prior research has shown that adults with normal hearing benefit from the resulting acoustic differences between speech produced by male and female talkers, showing better speech recognition performance in the presence of one or two streams of competing speech when the target and masker talkers are different sexes compared to when they are the same sex (e.g., Festen and Plomp 1990; Brungart 2001; Helfer and Freyman 2008).

Adults with mild to moderate sensorineural hearing loss also appear to benefit from target/masker sex mismatches (e.g., Festen and Plomp 1990; Humes et al. 2006; Helfer and Freyman 2008). For example, Helfer and Freyman (2008) examined masked sentence recognition in 12 younger adults with normal hearing (mean age = 22.7 years) and 12 older adults (mean age = 71.5 years). While some older adults had pure-tone audiometric

thresholds within normal limits, most had mild to moderate high-frequency hearing loss. Target sentences were produced by an adult female. The maskers included two-female-talker speech (target/masker sex matched) and two-male-talker speech (target/masker sex mismatched). Both groups showed a substantial sex-mismatch benefit. Older adults performed more poorly overall than younger adults, however, particularly in the sex-mismatched condition. The authors posited that the pronounced group difference between older and younger adults in the sex-mismatched condition might have been due to impaired spectral resolution associated with hearing loss.

Although mild to moderate sensorineural hearing loss does not preclude a benefit associated with mismatches in target and masker sex, peripheral impairments could degrade some of the segregation cues that are available to listeners with normal hearing. Mackersie et al. (2011) assessed masked sentence recognition in a group of 13 adults with hearing loss (45-76 years) and a group of 6 adults with normal hearing (25 – 69 years) using the Coordinate Response Measure speech corpus (Bolia et al. 2000). The target and masker sentences were produced by the same male talker, and differences in F_0 and simulated vocal tract length were introduced using digital signal processing. Adults with hearing loss were tested while fitted with individualized amplification. Shifting the target F_0 up in frequency, above the masker F_0 , improved performance in the two groups to a similar degree. Whereas shifting the simulated vocal tract length provided benefit for normal-hearing listeners, most adults with hearing loss failed to benefit.

Like adults, school-age children with normal hearing benefit from mismatches in target/masker sex (e.g., Wightman and Kistler 2005; Leibold et al. 2018). In a recent study, Leibold et al. (2018) assessed whether a target/masker sex mismatch facilitates speech perception in a two-talker masker for children (5-10 years) and adults (18-33 years) with normal hearing. In two experiments, speech recognition thresholds (SRTs) for spondaic words were estimated for children and adults in a continuous masker comprised of two male talkers or two female talkers. Target words were either matched or mismatched in sex to the masker talkers. Although Leibold et al. (2018) reported the sex mismatch benefit as the difference in SRTs obtained for a given masker, the more conventional metric is the difference in SRTs obtained for a given target. Using this more conventional metric, the mean sex-mismatch benefit for child listeners was 9.6 dB (male target) and 5.4 dB (female target), compared to values in adults of 7.6 dB (male target) and 3.1 dB (female target). Given that SRTs were higher for children and adults in the sex-matched conditions, it was posited that greater benefit of a sex mismatch in children could reflect greater amounts of informational masking in the baseline condition.

Although school-age children with normal hearing take advantage of acoustic differences between speech produced by male and female talkers to separate target from masker speech, this effect has not been observed in listeners younger than 30 months of age (e.g., Newman and Morini, 2017; Leibold et al., 2018). These findings suggest that the ability to segregate streams of speech based on acoustic differences between speech produced by males and females does not emerge until the second year of life for individuals with normal hearing.

The question of whether mismatches in target/masker sex improve speech-in-speech recognition for school-age children with hearing loss has not been addressed in the literature. School-age children with normal hearing benefit from mismatches in target/masker sex, but several consequences of sensorineural hearing loss have the potential to limit the extent to which children with hearing loss take advantage of acoustic differences between speech produced by male and female talkers. Sensorineural hearing loss, including congenital sensorineural hearing loss, is most often a byproduct of abnormal functioning of sensory cells within the cochlea (reviewed by Korver et al. 2017). In addition to reduced audibility, these abnormalities may compromise the peripheral encoding of supra-threshold sounds. Deficits in supra-threshold sound encoding are particularly evident for individuals with moderate or greater degrees of hearing loss (e.g., Dubno and Dirks 1989; Henry et al. 2005). These participants exhibit reduced frequency selectivity (e.g., Glasberg and Moore 1986; Moore and Carlyon 2005) and impaired temporal coding (e.g., Buss et al. 2004). Degraded peripheral encoding is associated with poorer speech-in-noise recognition for children with hearing loss fitted with hearing aids relative to age-matched peers with normal hearing (e.g., Finitzo-Hieber and Tillman 1978; Gravel et al. 1999; Nitttrouer et al. 2013; McCreery et al. 2015).

Another important issue to consider when evaluating the extent to which children with hearing loss benefit from a target/masker sex mismatch in the context of speech-in-speech recognition is that many of these children have had reduced and/or inconsistent experience with sound relative to their peers with normal hearing (e.g., Walker et al. 2013; 2015). Both aided audibility and daily hearing aid usage moderate language outcomes for children with hearing loss (e.g., Moeller and Tomblin 2015). Further, it has been hypothesized that reduced auditory experience disrupts the development of perceptual abilities related to the segregation and selection of target from background speech (Leibold et al. 2013; Hillock-Dunn et al. 2015), and may impact general cognitive processing skills (e.g., McCreery et al. 2017). For example, Leibold et al. (2013) evaluated speech recognition in children with hearing loss who wear hearing aids and children with normal hearing, and they found that the detrimental effect of hearing loss was larger in a two-talker speech masker than in a spectrally matched noise masker.

In the present study, we compared the extent to which children with hearing loss and children with normal hearing take advantage of target/masker sex mismatches in the context of word recognition in a background of two competing talkers. Given the asymmetrical benefit of target/masker F_0 differences observed for adults with hearing loss by Mackersie et al. (2011) -- with a benefit observed when the F_0 of the target was higher than the masker but not the converse -- the release from masking associated with a target/masker sex mismatch was evaluated for target speech produced by a male talker *and* for target speech produced by a female talker. Based on previous data on older adults with hearing loss (Helfer and Freyman 2008) and on emerging evidence that many children with hearing loss have variable access to high-quality acoustic input (e.g., Walker et al. 2013; Moeller and Tomblin 2015), it was predicted that the benefit of a target/masker sex mismatch would be smaller for children with hearing loss than for children with normal hearing.

Methods

Participants:

A total of 36 school-age children participated in this study, including 18 children with sensorineural hearing loss and 18 children with normal hearing. Inclusion criteria for all participants were: (1) between the ages of 7 and 16 years; (2) English spoken at home; (3) negative history of unresolved conductive or middle ear issues; and (4) negative history of learning, cognitive, or motor delays as per parental report.

Children with hearing loss presented with sensorineural hearing loss in both ears that varied from mild (audiometric thresholds ranging from 26 to 40 dB HL) to severe (audiometric thresholds ranging from 71 to 90 dB HL) across participants. Note, however, that only one child presented with severe hearing loss. The remaining 17 children had hearing loss that ranged from mild to moderately severe across participants. All of the children with hearing loss wore bilateral hearing aids. These children were recruited from the Human Research Subjects Core database at Boys Town National Research Hospital in Omaha, Nebraska. They ranged in age from 7.3 to 15.7 years ($M = 11.3$ years, $SD = 2.4$). The mean age at the first hearing aid fitting was 3.8 years ($SD = 2.9$), and the mean duration of device use was 7.5 years ($SD = 3.0$). Demographic and device information for individual children with hearing loss is shown in Table 1. Audiometric data were obtained on the day of testing if previous audiometric data were more than six months old. The mean better-ear pure-tone average (PTA) was 45.0 dB HL ($SD = 14.0$). Pure-tone thresholds for both ears are provided for each child with hearing loss in Table 2.

Children with normal hearing were matched within 6 months of chronological age to the children with hearing loss, ranging in age from 7.1 to 15.6 years ($M = 11.3$ years, $SD = 2.5$). Prior to testing, each child with normal hearing passed a hearing screening at 20 dB HL for all octave frequencies between 250 and 8000 Hz (ANSI, 2010).

Stimuli:

Following Calandruccio et al. (2014), target stimuli were 30 disyllabic English words familiar to children as young as 5 years of age. The target words were recorded from one adult male and one adult female. Both talkers were native speakers of American English. Recordings were created in a sound-isolated room using a condenser microphone (AKG-1000S) positioned six inches from the talker's mouth. The mean F_0 was 143 Hz for the male talker and 214 Hz for the female talker. The recordings were amplified (TDT MA3) and digitized (CardDeluxe) using a 44.1 kHz sampling rate (32 bits). Individual words were then scaled to normalize the root-mean-square (rms) amplitude across words.

The masker was either two-male-talker speech or two-female-talker speech, with talkers reading different passages from the children's book, *Jack and the Beanstalk*. The mean F_0 values of the two male masker talkers were 144 Hz and 124 Hz, and the mean F_0 values of the two female talkers were 210 Hz and 170 Hz. All four masker talkers were native speakers of American English. The method for obtaining each masker recording was the same as described above for the target recordings. Silent pauses longer than 300 ms were reduced to approximately 200 ms. The two male and two female streams of speech were

equated for overall rms level, mixed, and edited so that the masker ended at a word boundary for both talkers. The two-female-talker masker sample was 2 min 48 sec, and the two-male-talker masker sample was 2 min 35 sec.

A custom MATLAB script was used to control the selection and presentation of stimuli. Stimuli were played through a 24-bit digital-to-analog converter (Avid, Fast Track Solo), amplified (Applied Research and Technology, SLA4), and presented via a loudspeaker (JBL, Control1 Pro). Participants were tested while seated in a sound-treated booth, facing a loudspeaker that was mounted at a distance of approximately 1 meter away.

Procedure:

Children with hearing loss wore their devices at user settings during testing, as programmed by their clinical audiologist. Prescriptive settings were based on the child's audiometric thresholds and individual real-ear-to-coupler differences (RECDs) using the Desired Sensation Level [i/o] v5.0 method (Scollie et al. 2005). Hearing aid verification was performed in the laboratory prior to testing (Audioscan, Verifit 2) in order to determine (1) the speech intelligibility index for a 65 dB SPL input, and (2) the maximum audible frequency defined as the highest audiometric frequency at which the long-term average speech spectrum reached or exceeded the participant's audiometric threshold. These measurements are provided for each child with hearing loss in Table 1. Hearing aid settings were not adjusted.

Participants were tested in each of four conditions: (1) male target speech/two-male-talker masker; (2) male target speech/two-female-talker masker; (3) female target speech/two-female-talker masker; and (4) female target speech/two-male-talker masker. Thus, there were two sex-matched conditions and two sex-mismatched conditions. Testing order was randomized across participants.

A familiarization task was completed in quiet prior to testing in which participants were asked to point to the picture associated with each of the 30 target words shown in a laminated picture book. All participants completed the familiarization phase with ease. During testing, the task was a four-alternative, forced-choice (4AFC). Participants held a touchscreen monitor (iPad Mini 2, Apple) that was connected wirelessly to the testing computer. One of the 30 target words was randomly selected on each trial. Three additional illustrations were selected at random and without replacement from the remaining 29 words to serve as foils. The four different pictures were displayed in black and white on the touchscreen prior to the onset of the selected target word. Following the presentation of the target word, the pictures turned from black-and-white to color. Participants indicated the word they heard by touching the corresponding image on the touchscreen. After each response, the correct image blinked in isolation to provide the participant with visual feedback.

The masker was played continuously during each threshold estimation track at an overall level of 60 dB SPL (57 dB/stream). The masker always began at the start of the sample and repeated continuously until the track finished. The level of the target words changed adaptively using a 2-down, 1-up stepping rule to achieve a SNR corresponding to 70.7%

correct performance (Levitt 1971). The level of the target words at the beginning of each track was approximately 10 dB above the expected threshold, based on pilot data. If warranted, starting levels were adjusted for individual participants after initial estimates were obtained. The initial step size was 4 dB, reducing to 2 dB after the first two reversals. Each track continued until eight reversals were obtained; the levels at the final six reversals were averaged to compute the SRT. Two threshold estimates were obtained for each participant in each condition. There was generally good agreement between the two SRT estimates within individuals. The average difference in SRT across the two estimates was 1.7 dB for CHH and 2.5 dB for CNH. A difference greater than 4 dB across the first and second SRT was rarely observed; exceptions include SRTs in a single condition for 1 CHH and 6 CNH. The final threshold was the average of both SRTs for each participant. Testing was completed in a single session lasting no longer than two hours.

Data analyses:

Repeated-measures analysis-of-variance (rmANOVA) was used to evaluate group differences in SRTs and in the magnitude of the sex-mismatch benefit associated with the male and female target talkers. A linear mixed-model analysis was conducted to determine the association between the logarithm (base 10) of child age and the magnitude of the sex-mismatch benefit for the two groups of children. The rationale for representing age in log units was to account for reduced maturational effects observed with increasing age during the school-age years (e.g., Buss et al. 2017).

RESULTS

Comparison of SRTs between children with hearing loss and children with normal hearing:

Figure 1 shows SRTs for children with hearing loss (left) and children with normal hearing (right). SRTs are represented in dB SNR. Overall, SRTs were higher for children with hearing loss than for children with normal hearing, although group differences were not equivalent across the four listening conditions. For example, while children with hearing loss required an additional 1.4 dB SNR to achieve comparable performance to peers with normal hearing for the male target/two-male-talker masker condition, this disadvantage was 4.5 dB for the male target/two-female-talker masker condition. SRTs for the one child with severe hearing loss fell within the range observed for the remaining 17 children with mild to moderately severe hearing loss.

A rmANOVA with SRT as the dependent variable was conducted to evaluate the trends observed in Figure 1. The analysis included the between-subjects factor of *group* (children with hearing loss, children with normal hearing), and the within-subjects factors of *target/masker correspondence* (matched, mismatched) and *target sex* (male, female). All three main effects were significant: *group* [$F(1,34)=8.95$; $p<0.001$; $\eta_p^2=0.54$], *target/masker correspondence* [$F(1,34)=191.47$; $p<0.001$; $\eta_p^2=0.85$], and *target sex* [$F(1,34)=50.42$; $p<0.001$; $\eta_p^2=0.60$]. There were significant two-way interactions between *group* and *target/masker correspondence* [$F(1,34)=16.78$; $p<0.001$; $\eta_p^2=0.33$] and between *target sex* and *target/masker correspondence* [$F(1,34)=4.68$; $p=0.04$; $\eta_p^2=0.12$]. Neither the two-way

interaction between *group* and *target sex* [$F(1,34)=1.19$; $p=0.28$; $\eta_p^2=0.03$] nor the three-way interaction term [$F(1,34)=0.56$; $p=0.46$; $\eta_p^2=0.02$] were significant.

The significant *group x target/masker correspondence* interaction indicates that the performance gap between children with hearing loss and children with normal hearing differed for sex-mismatched compared with sex-matched conditions. Specifically, SRTs were more similar across the two groups of children for sex-matched relative to sex-mismatched conditions. Follow-up tests were conducted to evaluate the pairwise differences among the marginal means for each group of children (with Bonferroni adjustment). These tests revealed that children with hearing loss performed more poorly than peers with normal hearing for both matched ($p<0.001$) and mismatched ($p<0.001$) target/masker conditions.

The *target sex x target/masker correspondence* interaction was also significant, indicating that the difference in SRT observed between the male and female target speech was not equivalent for the sex-matched and sex-mismatched conditions. Follow-up pairwise comparisons (with Bonferroni correction) indicated that SRTs were higher for the male target talker than for the female target talker for both matched ($p<0.001$) and mismatched ($p=0.001$) conditions.

Figure 2 shows the mean sex-mismatch benefit (SRT in sex-matched condition minus SRT in sex-mismatched condition) for both groups of children, plotted separately for the male and the female target talker. The average sex-mismatch benefit across both target talkers was 3.2 dB for children with hearing loss and 5.9 dB for children with normal hearing. Both significant interactions are evident in Figure 2. Consistent with the significant *group x target/masker correspondence* interaction, a smaller sex-mismatch benefit was observed for children with hearing loss than for age-matched peers with normal hearing. Consistent with the significant *target sex* and *target/masker correspondence* interaction, a smaller sex-mismatch benefit associated with the female target talker than the male target talker was observed for both groups of children.

Effects of age and group on the sex-mismatch benefit:

Individual estimates of the sex-mismatch benefit are plotted in Figure 3 as a function of child age. Estimates for the male and female target talkers are shown in the left and right panels, respectively. Data points above the dashed line indicate better performance in the sex-mismatched compared with the sex-matched condition. All children showed a sex-mismatch benefit for the male target talker, with the exception of a 15-year-old with hearing loss. Seventeen (out of 18) children with hearing loss and 16 (out of 18) children with normal hearing showed a sex-mismatch benefit for the female target talker.

Fixed effects of *log age*, *group* (children with hearing loss, children with normal hearing), and *target sex* (male, female) on the sex-mismatch benefit were analyzed with a linear mixed-effects model. Subject was included as a random variable. The analysis was conducted using the nlme package for R (Pinheiro et al. 2018). Results are provided in Table 3. A significant main effect of *group* was observed ($p<0.05$), indicating the sex-mismatch benefit was smaller for children with hearing loss than for children with normal hearing.

There was no main effect of *log age* ($p=0.40$) or *talker sex* ($p=0.15$). No interaction terms were significant.

No significant correlations were observed between the sex-mismatch benefit and degree of hearing loss (unaided pure-tone average in the better-hearing ear) or aided audibility (SII in the better ear) for children with hearing loss ($p>0.25$).

DISCUSSION

The goal of this study was to compare the extent to which children with hearing loss and age-matched peers with normal hearing benefit from mismatches in target/masker sex when asked to identify words in the presence of competing speech. A clear effect of target/masker mismatch was observed for both groups of children; SRTs were lower for sex-mismatched relative to sex-matched target/masker conditions. Individual data were consistent with the group trends; most children in each group showed a sex-mismatch benefit with both sets of target stimuli. Collectively, these results indicate that children with sensorineural hearing loss can take advantage of acoustic voice differences between male and female talkers to facilitate speech-in-speech recognition.

Although children with hearing loss were able to capitalize on mismatches in target/masker sex, the sex-mismatch benefit was significantly smaller for children with hearing loss than for age-matched children with normal hearing. This observation is consistent with prior data on older adults with hearing loss (Humes et al. 2006; Helfer and Freyman 2008). For example, Humes et al. (2006) observed that young adults with normal hearing (21-34 years) experienced a larger sex-mismatch benefit than older adults with hearing loss (61-81 years) in the context of monaural sentence recognition in a single competing talker. The present study extends this general finding to school-age children, providing additional evidence that sensorineural hearing loss interferes with the ability to recognize target speech in the presence of competing speech, even when the target and masker talkers are mismatched in sex, and listeners are tested wearing appropriately fitted hearing aids.

One possible explanation for the reduced sex-mismatch benefit experienced by individuals with hearing loss is that hearing loss interferes with the peripheral encoding of acoustic features that differentiate male and female voices (i.e., F_0 and vocal tract length). Data on adult hearing aid users with mild to severe sensorineural hearing loss are limited, and findings are somewhat mixed across studies (e.g., Summers and Leek 1998; Arehart et al. 2005; Mackersie 2011). Of highest relevance to the present study, Mackersie et al. (2011) evaluated the release from masking associated with target/masker differences in F_0 and vocal tract length in adults with normal hearing (mean age = 48 years) and adults with hearing loss (mean age = 61 years). While adults with normal hearing took advantage of target/masker differences in vocal tract length, most adults with hearing loss did not. The two groups benefited to a similar extent from target/masker F_0 differences when the target talker had a higher mean F_0 than the masker talker. Surprisingly, neither group benefitted from target/masker F_0 differences when the target talker had a lower mean F_0 than the masker talker. The authors posited that adults with hearing loss might have failed to utilize acoustic

differences associated with vocal tract length because of impaired frequency resolution associated with sensorineural hearing loss.

The a priori hypothesis of the present study was that children with hearing loss experience a smaller target/masker sex mismatch benefit than age-matched children with normal hearing. This hypothesis was based, in part, on the idea that impaired frequency resolution may reduce the extent to which adults with hearing loss benefit from target/masker differences in F0 and vocal tract length (e.g., Mackersie et al., 2011). In addition, children with hearing loss often have variable access to high-quality acoustic input (e.g., Walker et al. 2013; Moeller and Tomblin 2015). The ability to separate and attend to target speech when multiple people are talking at the same time appears to follow a prolonged time course of development (e.g., Wightman and Kistler 2005; Corbin et. al. 2016; Flaherty et al. 2018), and is thought to require extensive experience with sound (reviewed by Leibold 2017). This time course of development might be prolonged or disrupted in children with hearing loss who may have limited access to the acoustic information that differentiates one talker from another, particularly if their experience with sound is reduced and/or inconsistent relative to their peers with normal hearing (e.g., Walker et al. 2013). Of course, access to high quality acoustic cues and prior experience are likely associated.

The benefit of the sex-mismatch was larger for both groups of children with the male target words than with the female target words. For example, children with normal hearing showed an average sex-mismatch benefit of 6.8 dB for the male target talker, compared with 5.0 dB for the female target talker. A similar asymmetry in sex-mismatch benefit was observed for children and adults with normal hearing tested by Leibold et al. (2018), using speech produced by different talkers than in the present study. In that study, an average sex-mismatch benefit of 9.6 dB was observed for 5- to 10-year-olds with normal hearing in the context of a 4AFC word identification task when target words were produced by a male talker. The corresponding benefit was 5.4 dB when target words were produced by a female talker.

One possible explanation for the reduced sex-mismatch benefit observed with the female target speech relative to the male target speech is that the female target/two-female-talker masker produced less informational masking than the male target/two-male-talker masker condition, limiting the magnitude of sex-mismatch benefit that could be obtained. One approach that has been used in previous studies to estimate the amount of informational masking produced by combinations of target and masker speech is to compare performance between conditions in which the target and masker speech are co-located in space and conditions in which target and masker speech are separated on the horizontal plane (e.g., Freyman et al. 1999; Arbogast et al. 2002). Spatially separating the target and masker speech is thought to facilitate segregation (e.g., Freyman et al. 1999). A larger effect of spatial separation is often interpreted as indicating greater informational masking in the baseline (co-located) condition (Kidd et al. 2016).

To evaluate the relative magnitudes of informational masking associated with the two-talker-male and two-talker-female maskers, supplemental data were collected on a group of six adults with normal hearing (22-32 years). Each adult completed testing in a single visit

lasting approximately one hour. Participants completed testing in the four conditions included in the main experiment, in which the target and masker stimuli were presented through a single loudspeaker at 0° azimuth (co-located). Participants also completed testing in four additional conditions in which target stimuli were presented at 0° azimuth while masker stimuli were presented to the right at 90° azimuth (spatially separated). Testing order was randomized across the eight conditions, with 2-3 runs completed per condition. If differences in informational masking are responsible for differences in performance with the male and female maskers in the main experiment, then providing a segregation cue should reduce variability across conditions. We therefore expected to see variability in thresholds when the target and masker were co-located, but more uniform performance when they were spatially separated.

The supplemental data support the idea that differences in informational masking between sex-matched target/masker conditions may be responsible for the reduced sex-mismatch benefit observed with female versus male target speech tokens in the primary dataset, at least for adults. When the target and masker were co-located, the average SRT differed by 14.3 dB across the four conditions. Considering the co-located, sex-matched conditions, the average SRT was -4.8 dB for the male target/two-male-talker masker condition (SD=2.7) compared with -17.3 dB for the female target/two-female-talker masker condition (SD=4.5). Considering the co-located, sex-mismatched conditions, the average SRT was -18.3 dB for the male target/two-female-talker masker condition (SD=3.3) compared with -19.1 dB for the female target/two-male-talker masker condition (SD=3.3). In contrast to variability in SRTs across co-located conditions, the average SRT differed by only 4.5 dB across the four conditions when the target and masker were spatially separated. The average SRT for spatially separated conditions was -21.5 dB for the male target/two-female-talker masker condition (SD=1.8), -26.8 dB for the female target/two-female-talker masker condition (SD=3.5), -26.0 dB for the male target/two-female-talker masker condition (SD=1.8), and -24.7 dB for the female target/two-male-talker masker condition (SD=3.0).

As expected, variability across conditions was reduced in the spatial separation condition, but it was not eliminated; the remaining 4.5-dB difference in SRTs could reflect modest differences in energetic masking, residual effects of informational masking, or differences in head shadow associated with differences in the spectral content of the male and female voices. These results provide support for the idea that higher SRTs for the male target/two-male-talker masker condition than the female target/two-female-talker masker in the primary dataset are due to greater informational masking in the male target/two-male-talker masker condition.

Another consideration when evaluating the magnitude of the sex-mismatch benefit in the primary dataset is the observation that SRTs tended to hover near 0 dB SNR, particularly for children with hearing loss. Previous studies investigating speech-in-speech recognition have similarly observed that SRTs rarely exceed more than 1-2 dB SNR for listening conditions expected to provide substantial informational masking for both adults and children (e.g., Swaminathan et al. 2015; Corbin et al. 2016). It has been argued that the psychometric function approaches the upper asymptote once the level of the target speech exceeds the level of the masker speech (e.g., Brungart 2001; Wightman and Kistler 2005; Swaminathan

et al. 2015), which would limit the magnitude of sex-mismatch benefit that can be achieved. Future experiments are planned to evaluate the influence of relative level differences between target and masker speech in children.

Our sample of children with hearing loss spanned a wide range of ages, and varied with respect to degree of hearing loss and the age at which they first received hearing aids. While no obvious association between these factors and the magnitude of sex-mismatch benefit was observed, systematic evaluation of the influence of listener factors on performance requires a larger sample of children with hearing loss. Another potential limitation is that speech recognition performance was assessed for disyllabic word recognition using a forced-choice paradigm and a limited number of target and masker talkers. To increase the generalizability of these findings, future studies could evaluate open-set word or sentence recognition using speech materials recorded from a larger number of target and masker talkers.

In summary, our results provide additional evidence that children with hearing loss are at a disadvantage relative to children with normal hearing when asked to recognize speech in the presence of a small number of competing talkers. The present findings extend prior research by demonstrating a larger performance gap between children with hearing loss and children with normal hearing for sex-mismatched compared with sex-matched target/masker configurations. One implication of our findings is that hearing loss appears to impact the extent to which children are able to take advantage of acoustic differences between male and female voices. The observation of larger effects for male than female target speech for both groups of children highlights the clinical importance of considering stimulus factors when evaluating masked speech recognition. Further testing of speech-in-speech recognition in children with hearing loss is important to determine whether the present results reflect cascading effects of variable auditory experience, peripheral encoding deficits associated with hearing loss, or a combination of factors.

ACKNOWLEDGEMENTS

Research reported in this publication was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under award number R01 DC011038 (LJL) and by the National Institute of General Medical Sciences of the National Institutes of Health under award number P20GM109023.

Financial Disclosures/Conflicts of Interest

This research was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under award number R01 DC011038. Subject recruitment was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number P30DC004662. The authors have no conflicts of interest to disclose.

REFERENCES

- Ambrose SE, VanDam M, Moeller MP (2014). Linguistic input, electronic media, and communication outcomes of toddlers with hearing loss. *Ear Hear*, 35(2), 139–147. [PubMed: 24441740]
- ANSI (2010). ANSI S3.6-2010, Specifications for Audiometers (American National Standards Institute, New York).
- Arbogast TL, Mason CR, Kidd G Jr (2002). The effect of spatial separation on informational and energetic masking of speech. *J Acoust Soc Am*, 112(5), 2086–2098. [PubMed: 12430820]

- Arehart KH, Rossi-Katz J, Swensson-Prutsman J (2005). Double-vowel perception in listeners with cochlear hearing loss: differences in fundamental frequency, ear of presentation, and relative amplitude. *J Speech Lang Hear Res*, 48(1), 236–252. [PubMed: 15938067]
- Bolia RS, Nelson WT, Ericson MA, Simpson BD (2000). A speech corpus for multitalker communications research. *J Acoust Soc Am*, 107(2), 1065–1066. [PubMed: 10687719]
- Bregman AS (1990). *Auditory Scene Analysis: The perceptual organization of sound*. MIT Press: Cambridge.
- Brungart DS (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am*, 109(3), 1101–1109. [PubMed: 11303924]
- Buss E, Hall III JW, Grose JH (2004). Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear Hear*, 25(3), 242–250. [PubMed: 15179115]
- Buss E, Leibold LJ, Porter HL, Grose JH (2017) Speech recognition in one- and two-talker maskers in school-age children and adults: Development of perceptual masking and glimpsing. *J Acoust Soc Am* 141(4), 2650–2660. [PubMed: 28464682]
- Calandruccio L, Gomez B, Buss E, Leibold LJ (2014). Development and preliminary evaluation of a pediatric Spanish–English speech perception task. *Am J Audiol*, 23(2), 158–172. [PubMed: 24686915]
- Corbin NE, Bonino AY, Buss E, Leibold LJ (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear Hear*, 37(1), 55–63. [PubMed: 26226605]
- Dubno JR, Dirks DD (1989). Auditory filter characteristics and consonant recognition for hearing-impaired listeners. *J Acoust Soc Am*, 85(4), 1666–1675. [PubMed: 2708682]
- Festen JM, Plomp R (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am*, 88(4), 1725–1736. [PubMed: 2262629]
- Finitzo-Hieber T, Tillman TW (1978). Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children. *J Speech Lang Hear Res*, 21(3), 440–458.
- Fitch WT, Giedd J (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *J Acoust Soc Am*, 106(3), 1511–1522. [PubMed: 10489707]
- Flaherty MM, Leibold LJ, Buss E (2018). Developmental effects in the ability to benefit from F0 differences between target and masker speech. *Ear Hear*, publish ahead of print, doi: 10.1097/AUD.0000000000000673
- Freyman RL, Helfer KS, McCall DD, Clifton RK (1999). The role of perceived spatial separation in the unmasking of speech. *J Acoust Soc Am*, 106(6), 3578–3588. [PubMed: 10615698]
- Gaudrain E, Ba kent D (2018). Discrimination of voice pitch and vocal-tract length in cochlear implant users. *Ear Hear*, 39(2), 226–237. [PubMed: 28799983]
- Glasberg BR, Moore BCJ (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J Acoust Soc Am*, 79(4), 1020–1033. [PubMed: 3700857]
- Gravel JS, Fausel N, Liskow C, Chobot J (1999). Children’s Speech Recognition in Noise Using Omni-Directional and Dual-Microphone Hearing Aid Technology. *Ear Hear*, 20(1), 1–11. [PubMed: 10037061]
- Helfer KS, Freyman RL (2008). Aging and speech-on-speech masking. *Ear Hear*, 29(1), 87–98. [PubMed: 18091104]
- Henry BA, Turner CW, Behrens A (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *J Acoust Soc Am*, 118(2), 1111–1121. [PubMed: 16158665]
- Hillock-Dunn A, Taylor C, Buss E, Leibold LJ (2015). Assessing speech perception in children with hearing loss: What conventional clinical tools may miss. *Ear Hear*, 36(2), e57–e60. [PubMed: 25329371]
- Humes LE, Lee JH, Coughlin MP (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *J Acoust Soc Am*, 120(5), 2926–2937. [PubMed: 17139749]

- Kidd G Jr., Mason CR, Swaminathan J, Roverud E, Clayton KK, Best V (2016). Determining the energetic and informational components of speech-on-speech masking. *J Acoust Soc Am*, 140(1), 132–144. [PubMed: 27475139]
- Korver AM, Smith RJ, Van Camp G, Schleiss MR, Bitner-Glindzicz MA, Lustig LR, Usamin SI, Boudewyns AN (2017). Congenital hearing loss. *Nat Rev Dis Primers*, 3, 16094. [PubMed: 28079113]
- Leibold LJ (2017). Speech perception in complex acoustic environments: Developmental effects. *J Speech Lang Hear Res*, 60(10), 3001–3008. [PubMed: 29049600]
- Leibold LJ, Buss E, Calandruccio L (2018). Developmental Effects in Masking Release for Speech-in-Speech Perception Due to a Target/Masker Sex Mismatch. *Ear Hear*, 39(5), 935–945. [PubMed: 29369288]
- Leibold LJ, Hillock-Dunn A, Duncan N, Roush PA Buss E (2013). Influence of hearing loss on children’s identification of spondee words in a speech-shaped noise or a two-talker masker. *Ear Hear*, 34(5), 575–584. [PubMed: 23492919]
- Levitt HC (1971). Transformed up-down methods in psychoacoustics. *J Acoust Soc Am*, 49(2B), 467–477.
- Mackersie CL, Dewey J, Guthrie LA (2011). Effects of fundamental frequency and vocal-tract length cues on sentence segregation by listeners with hearing loss. *J Acoust Soc Am*, 130(2), 1006–1019. [PubMed: 21877813]
- McCreery RW, Walker EA, Spratford M, Oleson J, Bentler R, Holte L, Roush P (2015). Speech recognition and parent-ratings from auditory development questionnaires in children who are hard of hearing. *Ear Hear*, 36(0 1), 60S–75S. [PubMed: 26731160]
- McCreery RW, Spratford M, Kirby B, Brennan M (2017). Individual differences in language and working memory affect children’s speech recognition in noise. *Int J Audiol*, 56(5), 306–315. [PubMed: 27981855]
- Moeller MP, Tomblin JB (2015). An introduction to the Outcomes of Children with Hearing Loss study. *Ear Hear*, 36(0 1), 4S–13S. [PubMed: 26731159]
- Moore BC, Carlyon RP (2005). Perception of pitch by people with cochlear hearing loss and by cochlear implant users In *Pitch* (pp. 234–277). Springer, New York, NY.
- Newman RS, Morini G (2017). Effect of the relationship between target and masker sex on infants’ recognition of speech. *J Acoust Soc Am*, 141(2), EL164–EL169. [PubMed: 28253666]
- Nittrouer S, Caldwell-Tarr A, Tarr E, Lowenstein JH, Rice C, Moberly AC (2013). Improving speech-in-noise recognition for children with hearing loss: Potential effects of language abilities, binaural summation, and head shadow. *Int J Audiol*, 52(8), 513–525. [PubMed: 23834373]
- Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team (2018). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-131.1, <https://CRAN.R-project.org/package=nlme>.
- Rendall D, Kollias S, Ney C, Lloyd P (2005). Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. *J Acoust Soc Am*, 117(2), 944–955. [PubMed: 15759713]
- Scollie S, Seewald R, Cornelisse L, Moodie S, Bagatto M, Larnagaray D, Pumford J (2005). The desired sensation level multistage input/output algorithm. *Trends Amp*, 9(4), 159–197.
- Smith DR, Patterson RD (2005). The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *J Acoust Soc Am*, 118(5), 3177–3186. [PubMed: 16334696]
- Summers V, Leek MR (1998). FO processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss. *J Speech Lang Hear Res*, 41(6), 1294–1306. [PubMed: 9859885]
- Swaminathan J, Mason CR, Streeter TM, Best V, Kidd G Jr, Patel AD (2015). Musical training, individual differences and the cocktail party problem. *Sci Rep*, 5, 11628. [PubMed: 26112910]
- Walker EA, Spratford M, Moeller MP, Oleson J, Ou H, Roush P, Jacobs S (2013). Predictors of hearing aid use time in children with mild-to-severe hearing loss. *Lang Speech Hear Serv Schools*, 44(1), 73–88.

- Walker EA, McCreery RW, Spratford M, Oleson J, Van Buren J, Bentler R, Roush P, Moeller MP (2015). Trends and predictors of longitudinal hearing aid use for children who are hard of hearing. *Ear Hear*, 36(0 1), 38S–47S. [PubMed: 26731157]
- Wightman FL, Kistler DJ (2005). Informational masking of speech in children: effects of ipsilateral and contralateral distracters. *J Acoust Soc Am*, 118(5), 3164–3176. [PubMed: 16334898]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

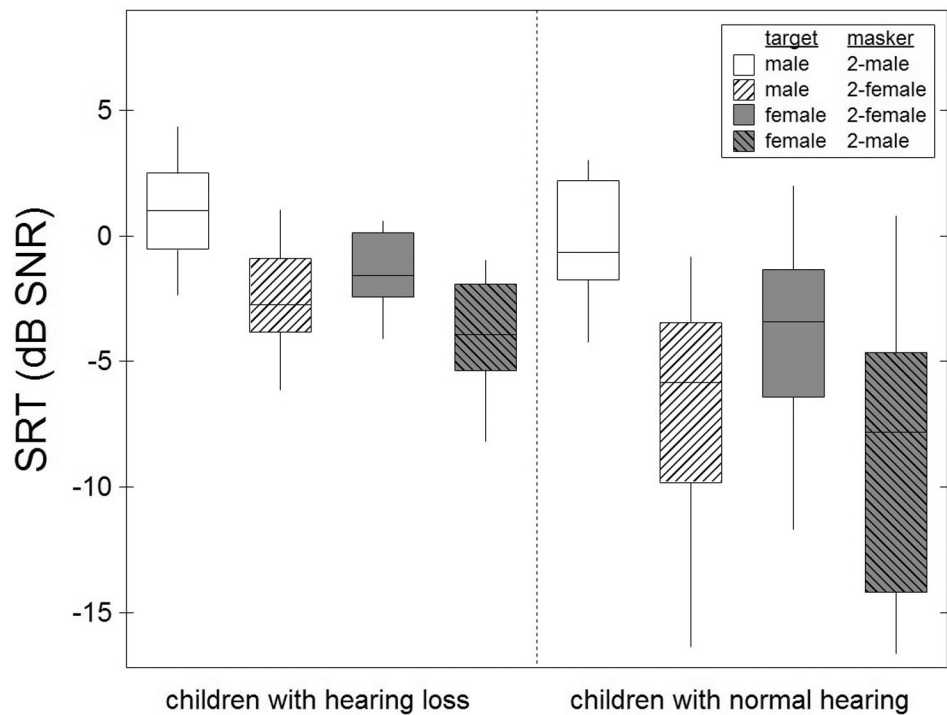


Figure 1.

The boxplots show speech recognition thresholds (in dB SNR) for children with hearing loss (left panel) and children with normal hearing (right panel). Box shading reflects the target and masker conditions, as defined in the legend. The range of performance for sex-matched and sex-mismatched conditions is shown in boxes without and with diagonal lines, respectively. Median scores are shown by the horizontal lines inside each box. The 10th and 90th percentiles are shown by the vertical lines.

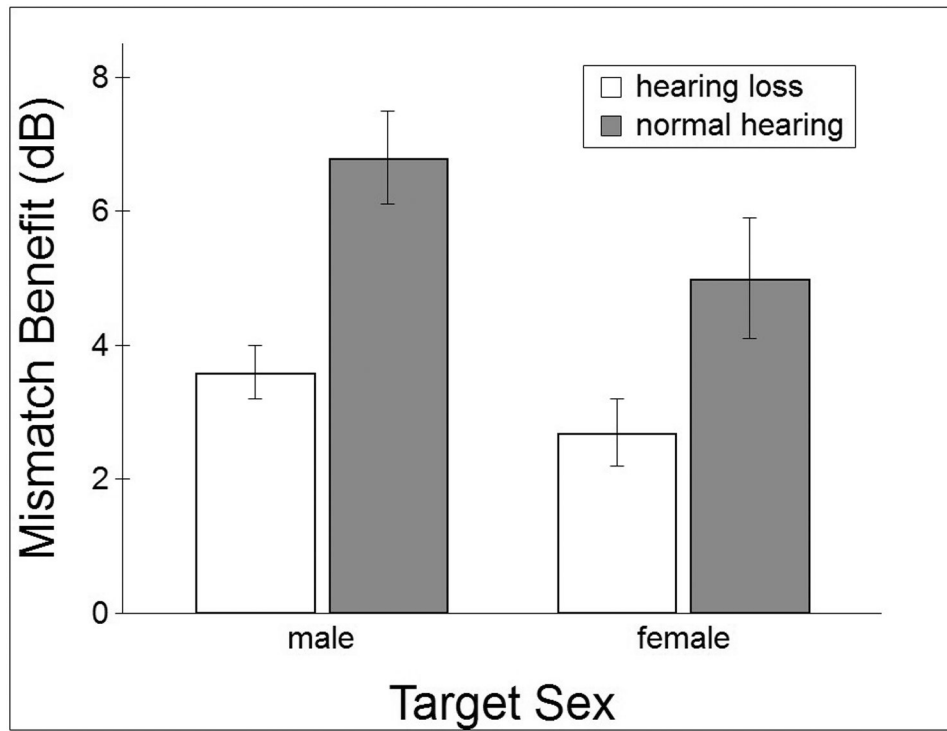


Figure 2. The group average sex-mismatch benefit (SRT in sex-matched condition minus SRT in sex-mismatched condition) is shown for children with hearing loss (open bars) and children with normal hearing (filled bars) for both the male and female target speech, indicated on the abscissa. Error bars show ± 1 standard deviation.

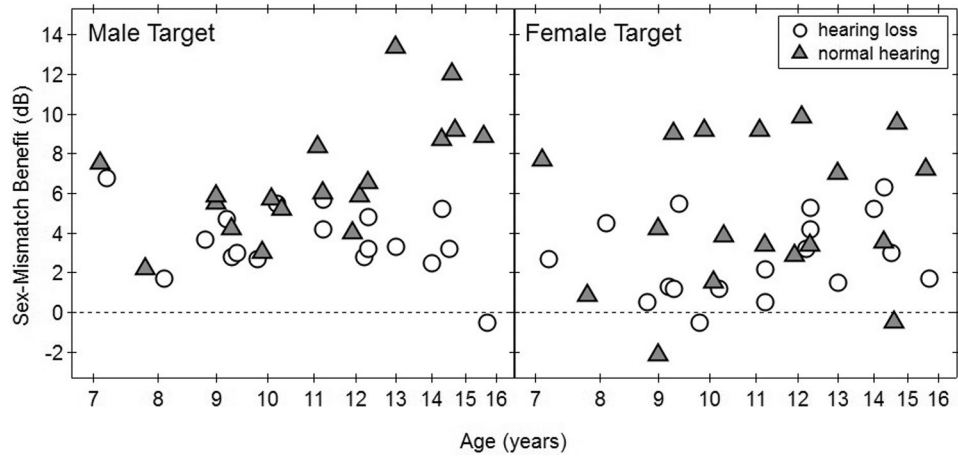


Figure 3. Individual estimates of the sex-mismatch benefit in dB are shown as a function of age for children with hearing loss (open circles) and children with normal hearing (filled triangles). Estimates for the male and female target speech are shown in the left and right panels, respectively.

Table 1.

Demographic and device information for children with hearing loss.

Listener	Sex	Etiology	Age at first HA fitting (years)	Duration of HA use (years)	Age at testing (years)	Current Amplification	Aided SII ^b at 65 dB SPL input (left ear)	Aided SII ^b at 65 dB SPL input (right ear)	MAF ^c (Hz; left ear)	MAF ^c (Hz; right ear)
HA-1	M	Genetic	8.0	3.2	11.2	Phonak Sky Q70 M13	84	87	8000	8000
HA-2	M	Genetic	8.0	3.2	11.2	Phonak Sky Q70 M13	91	92	8000	8000
HA-3	M	Genetic	4.8	4.3	9.2	Phonak Nios S H20 III	89	89	8000	8000
HA-4	F	Cochlear Malformation	4.5	7.7	12.2	Oticon Sensei Pro SP	69	74	5600	5700
HA-5	M	Unknown	0.2	9.1	9.3	Phonak Certena Micro	86	91	8000	8000
HA-6	F	Unknown	0.2	7.0	7.3	Phonak Sky Q50 M13	88	89	8000	8000
HA-7	M	Unknown	10.0	5.7	15.7	Phonak Audeo S SMART III	81	72	8000	8000
HA-8	M	EVA ^a	0.6	9.2	9.8	Phonak Sky Q70 SP	39	76	3000	5000
HA-9	M	EVA ^a	2.0	12.5	14.5	Phonak Bolero V70-P	72	71	8000	8000
HA-10	M	Unknown	0.3	12.1	12.3	Oticon Safari 300	82	82	8000	8000
HA-11	M	Unknown	4.0	10.3	14.3	Oticon Nero Pro RITE	61	68	8000	8000
HA-12	F	Unknown	3.3	6.9	10.2	Phonak Sky Q70 SP	32	30	5100	5600
HA-13	M	Unknown	5.0	4.4	9.4	Phonak Nios S H20 III	90	88	8000	8000
HA-14	F	Unknown	5.0	3.1	8.1	Phonak Nios S H20 III	80	78	6000	8000
HA-15	M	Unknown	2.0	11.0	13.0	Phonak Nios Micro III	76	77	5000	5000
HA-16	M	Genetic	5.0	7.3	12.3	Oticon Sensei BTE 13 90	88	83	8000	8000
HA-17	F	Unknown	4.5	9.5	14.0	Phonak Nios S H20 III	87	86	8000	8000
HA-18	F	Unknown	0.3	8.6	8.8	Phonak Nios S H20 III	84	89	8000	8000
<i>Mean</i>			3.8	7.5	11.3		77	79	7150	7406
<i>SD</i>			2.9	3.0	2.4		17	14	1518	1155

^aEVA = Enlarged Vestibular Aqueduct

^bSII = Speech Intelligibility Index

^cMAF = Maximum Audible Frequency

Table 2.

Pure-tone thresholds (dB HL) for children with hearing loss.

Listener	Frequency (Hz)													
	Left ear							Right ear						
	250	500	1000	2000	4000	8000	PTA*	250	500	1000	2000	4000	8000	PTA*
HA-1	30	30	30	35	40	40	31.7	25	25	25	35	40	50	28.3
HA-2	25	25	25	35	40	50	28.3	30	25	30	35	35	50	30.0
HA-3	30	35	45	45	45	40	41.7	30	35	40	45	45	40	40.0
HA-4	25	35	45	60	80	NR**	46.7	10	20	30	55	80	80	35.0
HA-5	25	30	45	40	40	35	38.3	25	35	40	40	30	30	38.3
HA-6	30	30	40	45	45	35	38.3	30	35	45	45	45	35	41.7
HA-7	25	40	55	50	40	30	48.3	20	40	55	50	40	40	48.3
HA-8	75	80	80	80	90	80	80.0	50	55	65	55	50	55	58.3
HA-9	45	65	70	60	55	65	65.0	45	70	70	60	55	70	66.7
HA-10	35	40	50	50	55	50	46.7	35	45	55	55	55	45	51.7
HA-11	25	45	55	70	65	55	56.7	30	45	55	70	65	55	56.7
HA-12	70	85	85	80	75	60	83.3	70	85	85	85	70	55	85.0
HA-13	20	35	45	40	35	60	40.0	20	30	40	45	30	55	38.3
HA-14	15	25	50	60	70	75	45.0	10	10	35	65	60	80	36.7
HA-15	35	40	55	55	60	55	50.0	40	40	60	55	60	35	51.7
HA-16	30	40	50	50	50	60	46.7	35	40	50	50	55	50	46.7
HA-17	25	35	40	50	50	50	41.7	25	35	45	50	50	45	43.3
HA-18	25	30	35	55	55	40	40.0	25	20	25	45	50	35	30.0

* PTA = the three-frequency Pure-Tone Average (500, 1000, 2000 Hz)

** NR = no response

Table 3.

Parameter estimates for the mixed-effects regression model analyzing data from children with hearing loss and children with normal hearing.

	Estimate (β)	Standard Error (SE)	df	<i>t</i> -value	<i>p</i> -value
intercept	2.50	0.67	32	3.75	< 0.001
target sex (male)	1.38	0.94	32	1.48	0.149
group (CNH)	2.25	0.95	32	2.38	0.023
log age	2.46	2.85	32	0.86	0.395
target sex (male) x correspondence (matched)	1.38	0.93	32	1.48	0.143
target sex (male) x group (CNH)	-0.17	1.32	32	-0.13	0.896
target sex (male) x log age	-5.16	4.00	32	-1.29	0.206
group (CNH) x log age	-0.47	4.00	32	-0.12	0.907
target sex (male) x group (CNH) x log age	11.26	5.60	32	20.1	0.053