Voice Emotion Recognition by Children with Mild to Moderate Hearing Loss

Shauntelle A. Cannon[1,2] and **Monita Chatterjee**[2]

[1]Department of Speech and Hearing Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[2]Auditory Prostheses & Perception Lab, Center for Hearing Research, Boys Town National Research Hospital, Omaha, NE, USA

Abstract

Objective—Emotional communication is important in children's social development. Previous studies have shown deficits in voice emotion recognition by children with moderate to severe hearing loss or with cochlear implants. Little, however, is known about emotion recognition in children with mild to moderate hearing loss. The objective of this study was to compare voice emotion recognition by children with mild to moderate hearing loss relative to their peers with normal hearing, under conditions in which the emotional prosody was either more or less exaggerated (child-directed or adult-directed speech, respectively). We hypothesized that the performance of children with mild to moderate hearing loss would be comparable to their normally hearing peers when tested with child-directed materials but would show significant deficits in emotion recognition when tested with adult-directed materials, which have reduced prosodic cues.

Design—Nineteen school-aged children (8–14 years of age) with mild to moderate hearing loss and 20 children with normal hearing aged 6–17 years participated in the study. A group of 11 young, normally hearing adults was also tested. Stimuli comprised sentences spoken in one of five emotions (*angry, happy, sad, neutral* and *scared*), either in a child-directed or in an adult-directed manner. The task was a single-interval, five-alternative forced-choice paradigm, in which the participants heard each sentence in turn and indicated which of the five emotions was associated with that sentence. Reaction time was also recorded as a measure of cognitive load.

Results—Acoustic analyses confirmed the exaggerated prosodic cues in the child-directed materials relative to the adult-directed materials. Results showed significant effects of age, specific emotion (happy, sad, etc.) and test materials (better performance with child-directed materials) in both groups of children, as well as susceptibility to talker variability. Contrary to our hypothesis, no significant differences were observed between the two groups of children in either emotion

Address correspondence to Monita Chatterjee, Auditory Prostheses & Perception Lab, Boys Town National Research Hospital, 425 N 30th St, Omaha, NE 68131, USA. Monita.Chatterjee@boystown.org. Phone: 402-452-5069, Fax: 402-452-5027.

recognition (percent correct or d′ values) or in reaction time, with either child- or adult-directed materials. Among children with hearing loss, degree of hearing loss (mild or moderate) did not predict performance. In children with hearing loss, interactions between vocabulary, materials and age were observed, such that older children with stronger vocabulary showed better performance with child-directed speech. Such interactions were not observed in children with normal hearing. The pattern of results was broadly consistent across the different measures of accuracy, d′ and reaction time.

Conclusions—Children with mild to moderate hearing loss do not have significant deficits in overall voice emotion recognition compared to their normally hearing peers, but mechanisms involved may be different between the two groups. The results suggest a stronger role for linguistic ability in emotion recognition by children with normal hearing than by children with hearing loss.

Keywords

children; hearing loss; emotion recognition

Introduction

The speech signal provides both linguistic and indexical information. Linguistic cues pertain to the lexical content of the speech signal, conveying the message via phonological, morphological, syntactic, and semantic information within the utterance. Indexical cues provide information about the talker, including cues to speaker identity, gender, and emotional state (Borrie et al. 2013). There is evidence that processing of the two components of speech occurs interdependently (Pisoni 1997), even in children and adults with cochlear implants (Li & Fu 2011; Geers et al. 2013). Other studies indicate that the use of indexical cues and the ability to attend to speech prosody are influenced by linguistic exposure early in infancy, especially when listening to infant- or child-directed speech (Jusczyk et al. 1993; Johnson et al. 2011). The objective of the present study was to investigate voice emotion recognition by children with hearing loss, who have limited access to both linguistic and indexical cues in speech.

Recognition of voice emotion requires a listener to have good access to acoustic cues relevant to emotional prosody, such as pitch and its associated changes, intensity, speaking rate, and vocal timbre. However, in hearing loss, the speech signal is inherently degraded and access to these acoustic cues may be limited. In children who are developing with hearing loss, the ability to map a limited set of acoustic cues to the correct vocal emotion may be further constrained by limitations in their cognitive and linguistic resources.

Voice Emotion Recognition and Social Development in Children with Hearing Loss

Voice emotion recognition is important in the development of both social competence and linguistic knowledge as it enables children to maximize the information obtained from the incoming speech stream and to react appropriately to their peers in different social settings. Geers et al. (2013) reported that better perception of the indexical components of speech,

such as the ability to discriminate voice emotion, is highly associated with more developed social skills in children with cochlear implants. Further, Eisenberg et al. (2010) reported links between social function/psychopathological symptoms and emotion understanding in normal hearing children and adults. Additionally, studies investigating the effects of deafness or severe hearing loss in children have shown deficits in social competence. Results from Wauters & Knoor (2008) indicated that deaf children scored lower on a questionnaire about prosocial behaviors (e.g. deaf children were seen as less likely to help other students or be cooperative in class) and scored higher on socially withdrawn behaviors (e.g. more likely to be bullied) compared to their normally hearing peers. Children with more severe hearing losses have also been shown to have deficits in psychosocial function (Laugen et al. 2016). Poorer performance in voice emotion identification has been related to lower self-reported quality of life in children with cochlear implants (Schorr et al. 2009).

Alongside deficits in psychosocial development, children with mild to profound hearing loss show significant deficits in auditory and facial emotion recognition relative to their typically developing peers (Most et al. 1993; Dyck et al. 2004; Ludlow et al. 2010; Most & Michaelis 2012). Furthermore, childhood hearing loss has been linked to children's development of "theory of mind," empathy and/or emotion understanding (e.g., Peterson & Siegel 1995 e.g., Peterson & Siegel 1998; Steeds et al. 1997; Russell et al. 1998; Rieffe & Terwogt 2000; Schick et al. 2007; Ketelaar et al. 2012a,b). Many of the studies addressing these issues, however, were conducted in children identified with severe to profound hearing loss and fitted with hearing aids before advances such as universal hearing screening, early diagnosis and intervention. Comprehensive studies in today's pediatric population using modern methodologies are much needed (Moeller 2007). In a recent study of false belief understanding using a large cohort of participants, Walker et al (2016) found that pre-school aged children with hearing loss have deficits in theory of mind, but generally caught up with normally hearing peers by the time they were in second grade. Some children with hearing loss showed considerable deficits in false belief understanding even at six years of age.

Voice Emotion Recognition and Hearing Loss

Emotion recognition is relatively straightforward when visual cues are present along with the speech signal; however in an auditory-only condition, the task is more challenging when key acoustic cues differentiating vocal emotion are degraded due to hearing loss. Studies of auditory-only emotion recognition in children with hearing loss, including children who are cochlear implant users, suggest significant deficits in vocal emotion recognition (e.g., Most 1993; Dyck et al. 2004; Most 2009; Most & Aviner 2009; Ludlow et al. 2010; Wiefferink et al. 2013; Chatterjee et al. 2015). Most of these studies have focused on children with more severe degrees of hearing loss. Information about the effect of lesser degrees of hearing loss (mild to moderate) on children's emotion recognition in speech is relatively limited. For example, only two child participants were categorized with mild hearing loss in the Dyck et al. (2004) study. Children with mild to moderate hearing loss should have enough low-frequency residual hearing to access the spectro-temporal information needed to perceive complex pitch variations in speech and make use of indexical cues (Ling 1976), but the extent to which they are actually able to identify emotional prosody remains unclear. Studies focusing on the effects of mild to moderate hearing loss have shown that these children have

deficits when compared to their normally hearing peers in areas of comprehension tasks in a classroom-simulated environment (Lewis et al. 2015), speech perception in quiet (Pittman et al. 2009), speech perception in noise (Crandell 1993), psychosocial development (Moeller 2007) and perceptual processing of auditory cues (Jerger 2007). Given that speech and language development occurs simultaneously with the development of social cognition, the ability of children with mild to moderate hearing loss to identify vocal emotion warrants investigation.

Acoustic Cues in Child- and Adult-directed Speech

Acoustic cues signaling vocal emotions have been extensively studied (e.g., Murray & Arnott, 1993; Banse & Scherer, 1996). Despite inter-talker variability, primary cues include mean voice pitch and its fluctuations, intensity, vocal timbre, and speaking rate. Although hearing loss preserves basic pitch cues, associated cues such as intensity and vocal timbre may be impacted by altered dynamic range and spectral shape coding. In addition, hearing aid processing (primarily designed to benefit speech perception with audibility in mind) may distort spectro-temporal information to some extent. Further, deficits in speech perception and linguistic access, combined with limited cognitive and linguistic development in young children, may result in increased cognitive load requirements for speech recognition tasks in general. It therefore seems reasonable to hypothesize that children with even mild to moderate hearing loss may have difficulties in vocal emotion recognition. In particular, these difficulties may be exacerbated by the type of speech they are being exposed to. For instance, speech with more exaggerated prosody may be easier for children with hearing loss to process than speech with less exaggerated prosody, which would carry fewer distinctive acoustic cues for different emotions. Child-directed speech is characterized by more exaggerated affective cues than afforded by adult-directed speech; including slower speaking rate, hyper-articulated vowels, and a wider pitch range (Fernald & Simon, 1984; Song et al 2010; Wang et al 2015). Child-directed speech has been shown to be more effective in maintaining the attention of infant and child listeners; there is evidence that exaggerated acoustic cues in child-directed speech serve to highlight important linguistic and prosodic elements of speech which are crucial to language development (Nelson et al, 1989; Singh et al, 2002). It is reasonable to speculate that the greater attention to speech with exaggerated prosody may also help young children decipher vocal emotional cues in speech. Adults have been also shown to benefit from infant-directed speech in an emotion-recognition task (Fernald, 1989).

However, children are not always exposed to speech under ideal conditions with affective or prosodic cues most appropriate for their developmental age. Instead, they are exposed to a variety of talkers with individual speaking styles, and are often present when adult-adult or adult to older child, such as an older sibling, interactions occur. Adults are able to make use of the statistical properties of speech phonetics occurring in their environment to enhance the implicit learning of speech (Hasher et al. 1987). The literature also demonstrates that children develop language through implicit or incidental learning (Reber 1967) and are adult-like in performance by first grade (Saffran et al. 1997). Thus, by the time they are school-aged, children have considerable exposure to adult-directed speech in addition to child-directed speech, through direct and indirect interactions in their daily lives. To ensure

that our study is ecologically relevant to children's everyday experiences with different styles of speech, we used both child- and adult-directed materials.

The Recognition of Individual Emotions by Children and Adults

The literature on emotion recognition has thoroughly documented, in normally hearing adults, that individual emotions are not recognized with the same accuracy. Previous studies of emotion recognition have used a wide range of vocal emotions. From these studies it is known that anger, sadness, fear, and to a lesser extent happiness/joy are more easily identified by participants, (Scherer, 1991; Juslin & Lukka, 2001; Linnankoski et al, 2005; Paulmann et al, 2008; Rodero, 2011; Sauter et al, 2010) while disgust, scorn, relief, and surprise are more difficult to recognize (Scherer, 1991; Linnankoski et al, 2005; Paulmann et al, 2008; Sauter et al, 2010). Data from Luo et al. (2007) showed that for adults with cochlear implants, recognition was best for sadness and the neutral emotion; recognition of anxious and happy emotions were most difficult. The literature investigating the recognition of individual emotions by children with more severe to profound hearing loss, including cochlear implant users, show similar patterns. Sadness was more easily recognized by children with cochlear implants while fear, surprise, and disgust were more difficult to identify(Most et al., 1993; Most & Aviner, 2009; Most & Michaelis, 2012; Mildner & Koska, 2014). In children with moderate to profound hearing loss, the pattern was somewhat different, with sadness, happiness, and anger being more easily recognized than fear, surprise and disgust (Most, 1993; Most, 2009; Most 2012). We therefore expect that children in our study, both normally hearing and with hearing loss will have differing levels of accuracy in recognition of individual emotions. However, the focus of the present study is on interactions between emotion types and children's hearing status (i.e., whether children with hearing loss have more relative difficulty with specific emotions than children with normal hearing).

Voice Emotion Recognition and Cognition

There are some indications in the literature of links between general cognition and/or linguistic abilities and measures of social cognition such as emotion recognition, at least in adults. Orbelo et al (2005) found that cognitive tests were marginally predictive of emotion processing in older adults with clinically normal age-related hearing loss. Lambrecht et al (2013) reported a significant relation between cognitive measures of attention and working memory and accuracy in an emotion recognition task. However, Lima et al (2014) did not observe any links between cognitive status and emotion recognition accuracy in older and younger adults. In a recent study of normally hearing children, Tinnemore et al (2017) found that the non-verbal intelligence quotient (NVIQ) predicts their recognition of voice emotion in spectrally-degraded speech, but did not find that vocabulary was a predictor. However, Geers and colleagues (2013) have found links between processing of indexical and linguistic elements of speech in children with cochlear implants, suggesting that verbal and linguistic skills are connected to their social cognition skills. It has been also shown that both vocabulary and working memory are predictive of speech recognition outcomes in children with normal hearing and hearing loss (Klein et al, 2017; McCreery et al., 2017). Together with the potential links between linguistic and indexical information processing of speech in

children with cochlear implants (Geers et al, 2013), it seems reasonable to speculate that vocabulary may be predictive of emotion recognition in children with hearing loss. Thus, these findings in the literature point to a potential role for both cognition and language in the development of emotion recognition in children.

Voice Emotion Recognition and Reaction Time

Reaction time is thought to be associated with measures of processing time, cognitive load, etc. Reaction time increases with age in adults with hearing loss compared to normally hearing adults (Feldman 1967; Husain et al. 2014) and may be related to changes in the brain used for emotional processing in adults with mild to moderate hearing loss (Husain et al. 2014). Slower reaction times in children and adults with mild to moderate hearing loss suggests an increase in listening effort and fatigue most likely due to the additional use of higher cognitive resources to attend to a task, e.g. speech understanding or emotion recognition (Hicks & Tharpe 2002; Tun et al. 2009). Further, Pals et al. (2015) showed that in a speech recognition task that focused on response accuracy rather than speed (as in the present study), response times reflected listening effort. Slower reaction times when recognizing and understanding voice emotion in real time in children with hearing loss could potentially place them at a disadvantage in social interactions with their peers. Therefore, reaction time was measured in the present study, alongside measures of accuracy (percent correct scores) and sensitivity (d′).

Objectives of the present study

The objective of the present study was to investigate the effects of mild to moderate hearing loss in children on their ability to perceive voice emotion using both adult-directed and child-directed speech materials. Specifically, we were interested in the predictive power of degree of hearing loss, age, cognitive and linguistic skills, and type of speech material (adult-directed or child-directed) in emotion recognition. A group of young, normally hearing adults were also tested to provide information about performance in our tasks at the end of the normal developmental trajectory. We hypothesized the following:

- Children with mild to moderate hearing loss would have emotion recognition accuracy scores and reaction times comparable to their normally hearing peers with child-directed materials, but would demonstrate deficits relative to their normally hearing peers with adult-directed materials.

- Accuracy in emotion recognition would be poorer and reaction times would be longer with adult-directed materials than child-directed materials in both children with normal hearing and children with mild to moderate hearing loss, and improvements would be observed as children progressed in age.

- Relative to normally hearing peers and older children with hearing loss, younger children with mild to moderate hearing loss would have greater difficulties in our tasks, resulting from combined effects of degraded/limited auditory input and age-limited cognitive and linguistic resources. Thus, we hypothesized

interactions between age, hearing status, and speech material (adult-directed and child-directed).

- A test of vocabulary and a test of NVIQ would predict performance in both children with normal hearing and in children with mild to moderate hearing loss, but more so in the latter group (as the reconstruction of the intended emotion from a degraded acoustic signal may require more linguistic and cognitive resources).

- Reaction times would be slower in younger children than in older children as well as for children with mild to moderate hearing loss compared to children with normal hearing; the interactions between hearing loss, age and material would also be reflected in reaction times.

- The young adults with normal hearing would outperform children, but would show poorer performance with adult-directed than with child-directed materials.

Materials and Methods

Participants

The results presented here were obtained from two groups of normally hearing children, one group of children with hearing loss, and one group of adults with normal hearing. One of the groups of normally hearing children was tested as part of a previously published study (Chatterjee et al., 2015). This group of participants included 31 children with normal hearing (15 boys, 16 girls, age range: 6.38–18.76 years, mean age: 10.76 years, s.d. 3.065 years). These participants only performed the emotion recognition task with child-directed materials (details are provided below).

The two other groups of children were tested specifically for the present study. These included 20 normally hearing children (9 boys, 11 girls, age range: 6.30–17.76 years, mean age: 11.94 years, s.d. 2.95 years) and 19 children with hearing loss (12 boys, 7 girls, age range: 8.29 – 14.61 years, mean age 9.76 years, s.d. 0.45 years). The group of normally hearing adults included 11 participants (3 men, 8 women, age range: 19.39–24.5 years, mean age: 21.98 years, s.d. 3.65 years). All participants were tested at Boys Town National Research Hospital in Omaha, NE. Participants were recruited from the Human Research Subjects Core database. Normal hearing sensitivity was confirmed for all participants by performing a hearing screening using an Interacoustics Diagnostic Audiometer AD226 (Interacoustics, Middlefart, Denmark). Using TDH-39 headphones, the participants were screened in both ears (20 dB HL for children and 25 dB HL for adults) for the frequencies 250 – 8000 Hz. For participants with hearing loss, air conduction and bone conduction thresholds were obtained. Thresholds from previous audiometric testing were used if the participant had been tested within six months of the test date. Participants were excluded if an air-bone gap greater than 10 dB was discovered. All children with hearing loss had bilateral, symmetric losses. They were separated into two groups (mild and moderate) based on the four frequency (500, 1000500, 2000, and 4000 Hz) pure tone average (PTA) of their better hearing ear. Further specifics on the group of children with hearing loss are provided

in Table 1. Note that participant HI_12 had normal thresholds at low frequencies, but mild to severe hearing loss at high frequencies in both ears, and used hearing aids.

Stimuli

The twelve semantically neutral sentences used in this study were selected from the HINT database and were the same stimuli produced in a child-directed manner described in Chatterjee et al. (2015). The talkers were asked to think of a scenario that would elicit one of five emotions (happy, sad, angry, scared, or neutral) as if they were talking to a child as young as age six for the child-directed materials and as if they were speaking to an adult for the adult-directed materials. The adult directed stimuli were selected from recent recordings completed by a number of male and female talkers. The two talkers for the adult-directed stimuli were the male and female talker from a cohort of ten whose recordings resulted in the highest performance on the voice emotion recognition task by a group of four normally hearing adults. In total, each talker produced 60 stimuli (12 sentences x 5 emotions). In addition, two different sentences were recorded in the same five emotions and used as practice materials (described below).

The sentences were presented to the participants in the sound field. To retain the natural intensity cues for emotion, the sentences were not equalized in intensity but rather, presented at the same level as a calibration tone (1 kHz) that had a root mean square (rms) level equal to the mean rms computed across sentences for each individual talker. The stimuli were routed through an external soundcard (Edirol-25EX) and presented from a single loudspeaker (Grason Stadler Inc.) located approximately 1m from the participants at a mean level of 65 dB SPL.

Procedure

Participants completed cognitive and vocabulary measures as well as a voice emotion recognition task. The order of testing was randomized for both children with normal hearing and children with hearing loss.

Cognitive measures included the matrix reasoning and block design subtests from the Weschler Abbreviated Scale of Intelligence Second Edition (WASI-II) (Wechsler & Psychological Corporation 2011). These were completed to assess the participant's NVIQ Vocabulary level was measured using the Peabody Picture Vocabulary Test Form B (PPVT) (Dunn et al. 2007). Parents completed the Behavior Rating Inventory of Executive Function (BRIEF) questionnaire (Gioia et al., 2000), which provides a measure of executive function in children.

A custom Matlab-based program (Emognition v 2.0.3) provided experimental control for the voice emotion recognition task. Test sentences were presented in sets, with each set consisting of materials recorded by a specific talker in a child- or adult-directed manner. Within each set, the sentences were presented in randomized order.

The children with hearing loss participated in conditions with both child-directed and adult-directed materials. They completed the task with or without their hearing aids depending on their natural state of listening at home (i.e. those who were consistent device users in their

day-to-day lives wore hearing aids during testing while those who were inconsistent device users or did not own hearing aids did not wear them during testing). The emotion recognition tasks were combined with cognitive and vocabulary measurements, and the tests were administered in fully randomized order.

The adults with normal hearing and children with normal hearing were only tested with the adult-directed materials; if the participants had recent (6-month) cognitive and vocabulary test results on file, emotion test sets were counterbalanced between male and female talkers using the adult-directed materials. Adults did not participate in cognitive or vocabulary testing. If the children with normal hearing had not participated in recent cognitive/ vocabulary tests, these were administered at the time of emotion testing, and the order of tests was fully randomized (as with children with hearing loss). The children with normal hearing who participated in the previous study had not been tested on cognitive/vocabulary skills, and those results were not available for analysis.

In a single-interval, five-alternative forced-choice procedure, participants listened to the test set of 60 sentences (5 emotions, 12 sentences) for each talker (male or female) and each set of materials (adult-directed or child-directed). Participants were encouraged to take breaks between blocks. Each condition was repeated twice, and the mean percent correct of the two runs was calculated to obtain the final accuracy score for each set. The software program recorded the reaction time (in seconds) for each sentence, and the mean was calculated to obtain the final average reaction time for each set. In addition, confusion matrices were recorded for each test set. Before each test set participants were given two passive training sessions that consisted of 10 sentences (5 emotions, 2 sentences). The sentences in the training sets were not included in the test set. During the training session the listener would hear each sentence in a given emotion, and the correct choice of emotion would be indicated on the screen. The purpose of the passive training was to familiarize the participant with the talker's manner of producing the different emotions. After the training sets, the participants listened to each test sentence in turn (randomized order) and responded by clicking on the appropriate button on the screen. No feedback from the experimenter was provided during the formal test and there were no repetitions of the sentences in each block. Due to time constraints, one of the children with hearing loss was unable to complete the task with child-directed.

To reiterate, the normally hearing children in the present study were only presented with adult-directed materials. The normally hearing children from Chatterjee, et al. (2015) were only presented with child-directed materials. The purpose of their inclusion was to provide comparative analyses for the present dataset.

Statistical Analyses

Acoustic analyses of the stimuli were conducted using the Praat software program (Boersma, 2001; Boersma & Weenink, 2017). The percent correct scores were transformed prior to data analyses using the rationalized arcsine transform (Studebaker, 1985). Statistical analyses were completed in R version 3.12 (R Core Team 2015) using *nlme* (Pinheiro et al. 2015) for the linear mixed-effects model implementation and *ggplot2* (Wickham, 2009) to plot the data https://owa.boystown.org/owa/-_msocom_2. Prior to analyses, outliers were

identified and removed using "Tukey fences" (Tukey, 1977). Points that were above the upper fence (third quartile + 1.5*interquartile range) or below the lower fence (first quartile – 1.5*interquartile range) were considered outliers. The RAU-transformed scores showed no outliers. The reaction time data with adult-directed speech had 4 outliers (comprising 5.4% of the dataset), which were removed.

Results

Acoustic analyses of the stimuli

Figure 1 compares results of basic acoustic analyses performed on the adult- and child-directed materials. Analyses focused on pitch and intensity, two fundamental aspects of prosody. On all measures, the child-directed materials showed greater variations across emotions than the adult-directed materials (Fig 1A,B). To statistically confirm this, the ratio between the mean estimated fundamental frequency (F0) for each sentence spoken in each of the *happy, scared, sad,* and *angry* emotions to the mean F0 estimated for the same sentence spoken in the *neutral* emotion, was computed ($F0_{ratio}$; Fig. 1C). Similarly, the difference in estimated intensity in each of the emotions to the mean intensity of the individual sentence spoken in the *neutral* emotion, was computed ($Intensity_{difference}$; Fig 1D). A linear mixed-effects analysis was conducted with $F0_{ratio}$ as the dependent variable, Emotion (four levels) and Material (adult- or child-directed) as fixed effects, and sentence-based random intercepts. Results showed a significant interaction between Emotion and Material ($F(1,177)=8.01, p=0.005$), and a main effect of Material ($F(1,177)=7.36, p=0.007$). The interaction is evident in the greater $F0_{ratio}$ for happy and scared in the child-directed materials relative to the adult-directed materials. A similar linear mixed-effects analysis conducted with $Intensity_{difference}$ as the dependent variable in place of $F0_{ratio}$ showed significant main effects of Emotion ($F(1,177)=8.18, p=0.005$) and Material ($F(1177)=9.47, p=0.002$), and a significant interaction ($F(1,177)=29.53, p<0.001$). The interaction is evident in the exaggerated differences between emotions in Fig 1D. Although the focus of the present study is not on the acoustic analyses, it is evident that the child-directed materials involve greater across-emotion differences in primary acoustic cues than the adult-directed materials.

Group differences in performance with adult-directed materials

Figure 2A shows boxplots of the rationalized arcsine unit (RAU) scores obtained with the adult-directed materials in the three groups (children with hearing loss, children with normal hearing, and adults with normal hearing). Results of a linear mixed-effects model with Group and Talker (male/female) as fixed effects and subject-based random intercepts showed significant effects of Group ($F(2,44)=6.52, p=0.003$) and Talker ($F(1,44)= 9.10, p=0.004$), with no interactions. Post-hoc pairwise t-tests with Bonferroni correction showed that the scores for children with normal hearing and children with hearing loss were significantly lower than for adults with normal hearing ($p<0.001$ in both cases), but were not different from one another ($p=0.112$). The difference in performance between adults and children (both groups of children considered together) was 0.95 times the mean standard deviation (i.e., the effect size (Cohen's d (Cohen, 1992) was 0.96, a large effect). Paired t-

tests within groups showed that all groups performed better with the female talker than with the male talker (p<0.001 for all groups).

Reaction time data were analyzed using, a linear mixed-effects model with Group and Talker as fixed effects and subject-based random intercepts, showed a weak effect of Group (F(2,42)=3.31, p=0.046), but no effects of Talker and no interactions. Post-hoc pairwise t-tests with Bonferroni correction showed that reaction times for children with normal hearing were significantly longer than those of adults with normal hearing (p=0.001) and reaction times for children with hearing loss were also significantly longer than those of adults with normal hearing (p<0.001), but the reaction times were not different between the two groups of children (Fig 2B). The effect size of group (adults vs. children) on the reaction time was large (Cohen's d=0.95).

Figure 3 shows the RAU scores obtained in each group, plotted against the five emotions, for the adult-directed materials. It is apparent that the pattern of results was similarly dependent on emotion across the different groups. This was confirmed by a linear mixed-effects analysis with RAU scores as the dependent variable, Group, Emotion and Talker as fixed effects and subject-based random intercepts. Results showed significant effects of Emotion (F(1, 417) = 32.91, p<0.001) and Talker (F(1,417)=11.99, p<0.001; performance with the female talker > the male talker), but no interactions. A pairwise t-test with Bonferroni corrections for multiple comparisons showed significant differences between *happy* and *angry* (p=0.009), *happy* and *neutral* (p<0.001), and a marginal difference between *happy* and *sad* (p=0.052). Significant differences were also found between *scared* and *angry* (p=0.005), *scared* and *neutral* (p<0.001) and *scared* and *sad* (p=0.032). As interactions were not observed between emotions and groups or talkers, the remaining analyses focused on overall accuracy. In the last section of Results, in which we consider d′ as a measure of sensitivity rather than accuracy, we confirm the significant effect of emotions, as well as the lack of interactions between emotions and groups.

Comparison of performance by children with normal hearing and children with hearing loss listening to adult-directed materials

No significant difference was observed in executive function between children with hearing loss and children with normal hearing (Student's t-test, p=0.88), and no significant correlations were observed between executive function measures and RAUs or reaction times. There were also no significant differences between the children with hearing loss and children with normal hearing in measures of vocabulary (Student's t-test, p=0.15) or NVIQ (Student's t-test, p=0.14). Figure 4(A–C) shows accuracy and reaction time results obtained in the two groups of children, with adult-directed materials. A linear mixed-effects analysis of RAU scores obtained with the adult-directed materials in children with normal hearing and children with hearing loss with subject-based random intercepts showed significant effects of Age (F(1,28)=8.79, p=0.006), Talker (Male/Female) (F(1,32)=12.29, p=0.001), and Vocabulary (F(1, 28)=5.23, p=0.030). These effects are shown in Figures A and B. No significant effects of Group or NVIQ were found.

Reaction time was significantly correlated with accuracy (r=−0.265, p=0.027). A linear mixed-effects analysis with Age, Group (normal hearing/hearing loss), Vocabulary, NVIQ

and Material (adult-directed male/female) as fixed predictors of reaction time and subject-based random intercepts showed a significant decline with age (F(1,28)=6.67, p=0.015) (Fig 4C), but as with the accuracy scores, reaction times were not significantly different between groups. Reaction times were not predicted by Vocabulary, NVIQ or Material.

Comparison of performance by children with normal hearing and children with hearing loss listening to child-directed materials

We also compared emotion recognition performance with child-directed materials by the group of children with hearing loss in the present study against the performance of the group of children with normal hearing in Chatterjee et al. (2015). Figure 5 shows RAU scores obtained in the two groups. A linear mixed-effects analysis comparing RAU scores obtained by the two groups of children listening to child-directed materials showed no significant effects of Age or Group, and no interactions (Fig 5A). However, the distributions of the data were different between the groups, with the lower quartile extending to much poorer scores for children with hearing loss than for children with normal hearing (Fig 5B).

Comparison of the performance by children with normal hearing listening to adult-directed and child-directed materials

We expected that performance by children with normal hearing would be poorer in the present study with adult-directed materials than in the Chatterjee et al (2015) study with child-directed materials. A linear mixed effects analysis comparing the performance of children with normal hearing with adult-directed materials in the present study with the group of children with normal hearing tested with child-directed materials from Chatterjee et al. (2015) showed a significant effect of Age (F(1,47)=25.79, p<0.0001) and Material (adult-directed/child-directed) (F(1,47)=28.87, p<0.0001), and no interactions (Fig 6A). The significant effect of Material was further confirmed by a large effect size (Cohen's d=2.45). Thus, the group of children with normal hearing in the present study showed much poorer performance with adult-directed materials than did their counterparts in Chatterjee et al (2015) with child-directed materials (Fig 6B).

Analyses of predictors of performance by children with hearing loss

Unlike the children with normal hearing, the children with hearing loss showed a significant correlation between vocabulary and cognition measures (PPVT and NVIQ; r=0.73, p<0.001). To avoid problems of interpretation associated with collinearity, NVIQ was excluded from subsequent analyses. The RAU scores obtained in the children with hearing loss were analyzed for predictive effects of Age, Degree of hearing loss (mild or moderate loss), Vocabulary, and the four levels of Material (adult- or child-directed, male and female talkers). The results indicated a significant effect of Age (F(1,10)=6.01, p=0.034), a significant interaction between Age and Vocabulary (F(1,10)=6.54, p=0.029), and a significant interaction between Material and Vocabulary (F(1,41)=8.11, p=0.007). No effect of Degree of hearing loss was found. The main effects of Age and the interaction between Vocabulary and Material are shown in Figure 7A and 7B respectively. To better understand the interaction between Age and Vocabulary, the participants were divided into older and younger age groups (younger than or older than 10 years of age). Figure 7C plots the RAU scores obtained by each group against their vocabulary scores, for the four different sets of

materials. It is apparent that the slope of the function relating RAUs to vocabulary scores is steeper for the older age group, particularly for the adult-directed materials. Separate multiple linear regression analyses were performed to estimate the slope relating RAU and vocabulary for each age group, taking material into account. The estimated slopes for the older and younger age groups were both significant (p<0.001 in each case), but the slope was steeper for the older age group (estimated slope: 0.533, s.e. 0.076) than for the younger age group (estimated slope: 0.31, s.e. 0.076). The difference is slopes was marginally significant (t(28)=2.08, p=0.047). These results suggest that older children gained more from stronger vocabulary in identifying emotions than did younger children.

To better understand the interaction between Material and Vocabulary, linear regression analyses to estimate the slope relating RAU and vocabulary for each of the four levels of Material (male or female talker, adult- or child-directed) showed shallower estimated slopes for the adult-directed materials than for the child-directed materials. Estimated slopes were: 0.266(s.e.=0.091) for the adult-directed, female talker materials, 0.401(s.e.=0.086) for the adult-direct, male talker materials, 0.501(s.e.=0.074) for the child-directed, female talker materials, and 0.555(s.e.=0.109) for the child-directed, male talker materials. Further analyses showed that the slope obtained with the child-directed, male talker was significantly steeper than the slope obtained with the adult-directed, male talker (t(28)=2.04, p=0.05), and the slope obtained with the child-directed, female talker was marginally significantly steeper than the slope of the adult-directed, female talker (p=0.055). Thus, in children with hearing loss, vocabulary was more strongly linked to performance with child-directed materials than adult-directed materials.

Reaction times were found to be significantly correlated with accuracy scores in the children with hearing loss (r=−0.35, p=0.004), with better scores being associated with faster reaction times (Fig 8A). A linear mixed-effects analysis with fixed effects of Age, Degree of hearing loss, Vocabulary and the four levels of Material (male or female talker, adult- or child-directed), and subject-based random intercepts showed significant three-way interaction between Age, Vocabulary and Material (F(1,37)=5.54, p=0.024), a two-way interaction between Age and Material (F(1,37)=5.70, p=0.022) and a main effect of Material (F(1,37)=5.44, p=0.025) but no other significant effects. To better understand the interactions, the subjects were divided into two groups by age (younger and older than 10), and the vocabulary scores were divided into two categories, lower than the median and higher than the median score. Fig 8B shows boxplots of the reaction times of children falling into each vocabulary category (abscissa), for the older and younger children (left and right panels) and for the different materials (colors). It is evident that the mean reaction times were shorter for the older children falling in the higher vocabulary group, but that vocabulary did not make a difference to the younger children's reaction times. The effect of vocabulary on the older children's reaction times was calculated to be moderate in size (Cohen's d=0.65). These relationships are further shown in the scatter plot in Fig 9. The mean of the adult-directed (male and female talker) and child-directed (male and female talker) reaction times were calculated, and plotted against age for the two groups of vocabulary scores (Fig 9A, B). A linear multiple regression model showed no relation between Reaction times and Age and no interaction with the Vocabulary categories for the adult-directed materials, but a marginally significant effect of Vocabulary category

(t(1.25)=1.77, p=0.098) and a marginally significant interaction of Age and the Vocabulary categories (t(0.12)= −1.88, p=0.08) for the child-directed materials (i.e. the slopes in the left and right panels in Fig 9B were marginally different. These results suggest that, with child-directed speech and for older children with higher vocabulary scores, reaction times were somewhat shorter.

To summarize, accuracy scores of the children with hearing loss suggest an interaction between age and vocabulary, indicating that among children with hearing loss, older children's performance may be more strongly mediated by language than younger children's. A second interaction was observed between vocabulary and material, indicating that children with stronger language skills showed better performance with the child-directed materials than children with weaker language skills. This difference was reduced in the case of adult-directed materials (shallower slope in Figure 7C). Reaction time data showed interactions between these variables as well. Reaction times with child-directed materials showed a greater dependence on vocabulary and age than reaction times with adult-directed materials. Overall, older children with higher vocabulary scores showed shorter reaction times. As NVIQ was correlated with vocabulary, we cannot separate the effects of language from those of general cognition in this analysis.

Analyses of results obtained in children with normal hearing listening to adult-directed speech

Given the prominent role of vocabulary as a predictor of performance by children with hearing loss, a separate analysis was conducted with the data obtained in children with normal hearing to examine the effects of vocabulary, age and material (male or female talker) in this group. Only the data from the present study with adult-directed materials were considered, as vocabulary was not measured in the previous study (Chatterjee et al, 2015) with child-directed materials. A linear mixed-effects analysis with the RAU score as the dependent variable, Vocabulary, Age and Material (male or female talker) as fixed effects, and subject-based random intercepts, showed a marginally significant effect of Age (F(1,16)=4.34, p=0.054) and a marginal interaction between Age and Material (F(1, 17)=4.322, p=0.053), but no effects of Vocabulary. The interaction between Age and Material was followed up by separate linear mixed effects analyses on the data obtained with the female talker's and the male talker's materials, respectively. In these analyses, the RAU score was the dependent variable, and predictors were Age and Vocabulary. The analyses with the data obtained with the female talker's materials showed no effects of either Age or Vocabulary. However, the analysis of the data with the male talker's materials showed a significant effect of Age (F(1,16)=7.35, p=0.015) and a marginal, non-significant effect of Vocabulary (F(1,16)=3.94, p=0.06). Thus, the interaction between Age and Material observed earlier was driven by an age-related improvement in performance with the male talker's materials that is not observed with the female talker's materials. These analyses indicate that, unlike the children with hearing loss, the children with normal hearing show talker-dependent effects of age, and no significant effects of vocabulary. A second analysis of reaction times showed no effects of Age, Material or Vocabulary on reaction times when this group is considered alone.

Analyses of d's

To test the possibility that children with hearing loss might have different emotion-specific deficits re: children with normal hearing that might be revealed in an analysis taking both accuracy (hits) and false alarms into account, confusion matrices obtained in the children were analyzed to derive hit rates and false alarm rates for each emotion. Figure 10A shows boxplots of the resulting d' values plotted for each emotion for the two groups for adult-directed materials. Consistent with the findings described above with accuracy as the outcome measure, a linear mixed-effects analysis with subject-based random intercepts and with d' as the independent variable, showed significant effects of Emotion ($F(1,266)=7.06$, $p=0.008$), Age ($F(1,27)=24.50, p<0.001$) and Vocabulary ($F(1,27)=12.60$, $p=0.001$), but no effects of Group or Material and no interactions. Following up on the effect of Emotion, post-hoc pairwise t-tests showed significant differences between *angry* and *neutral* ($p=0.024$), *angry* and *sad* ($p<0.001$), and *angry* and *scared* ($p<0.001$), and a significant difference between *happy* and *scared* ($p=0.029$).

In a second analysis, the d' scores obtained by Chatterjee et al (2015) in the children with normal hearing were compared with the d' scores obtained in the children with hearing loss in the present study, both with the same child-directed materials. Figure 10B shows boxplots of d' values obtained in the two groups of children with child-directed speech, for each emotion and each talker. It is evident that ceiling effects were strong in the children with normal hearing. Vocabulary scores were not obtained in the Chatterjee et al (2015) study. A linear mixed-effects analysis with Group (normal hearing or hearing loss), Emotion, Material (adult- or child-directed), and Age showed significant effects of Age ($F(1,46)=9.50$, $p<0.001$), Group ($F(1,46)=17.17$, $p<0.001$), and Material ($F(1,436)=47.18$, $p<0.001$), and an interaction between Emotion and Material ($F(1,436)=13.22$, $p<0.001$). Because of the ceiling effects in this dataset, it was not clear whether the interaction between Emotion and Material was meaningful and it was not followed up on.

Finally, a linear mixed-effects analysis of d' values obtained in only the children with hearing loss with the child-directed materials showed significant effects of Vocabulary ($F(1,12)=37.29$, $p<0.001$) and Emotion ($F(1,128)=4.52$, $p=0.035$). Effects of Age ($F(1,12)=3.66$, $p=0.080$) and Material ($F(1,128)=3.74$, $p=0.055$) did not reach significance. Post-hoc pairwise t-tests on the effect of Emotion showed only significant differences between *angry* and *neutral* ($p=0.049$) and *angry* and *scared* ($p=0.007$).

Overall, the patterns of results were generally consistent with those observed using accuracy scores. Effects of age, vocabulary and material were consistently present. Effects of emotion were also present, as observed in the percent correct scores. The specific effects of emotion were generally consistent across measures. However, given the talker-variability in acoustic cues for emotion and the limited number of talkers in our present stimulus set, these effects of emotion, or the observed interaction between emotion and material in one of the analyses, are unlikely to generalize to other talkers and materials, and are not of central interest to this study. The primary finding of interest is that no differences were observed between the two groups in emotion-specific patterns of performance.

Discussion

This study sought to examine how children with mild to moderate hearing loss performed on an emotion recognition task compared to children with normal hearing using speech stimuli containing more (child-directed) or less (adult-directed) exaggerated prosody. Performance measures in both groups were compared against those of a group of young, normally hearing adults. Our primary hypotheses were: 1) that the adult-directed materials would present greater difficulties for younger children with normal hearing and for children with hearing loss, and 2) that both performance accuracy scores and reaction times would reflect these differences. Additional research questions centered on the roles of the degree of hearing loss (for children with hearing loss), general cognition (executive function and NVIQ) and vocabulary in predicting the children's performance. One limitation of the study was the more limited age range of the children with hearing loss compared to the children with normal hearing. The results, however, showed no interactions between children's hearing status and age, suggesting overall similar age-related improvements in performance both groups of children.

Comparisons between groups

Our hypothesis that children with mild to moderate hearing loss would perform comparably to their normally hearing counterparts with child-directed materials, but would have greater difficulty with adult-directed materials, was only partially supported by our findings. The performance of children with hearing loss was not significantly different from children with normal hearing (present and previous study) using either child-directed or adult-directed materials; however, all groups showed poorer performance with adult-directed materials.

Our results also revealed a significant overall effect of developmental age on performance in the emotion recognition task. No interaction was observed between the children's hearing status and age. The adults with normal hearing outperformed both groups of children in all analyses with adult-directed materials. Vocabulary was a significant predictor of the children's performance, but executive function and NVIQ were not predictors.

Significant talker effects were observed, with both groups of children performing better with the female talker than the male talker in both sets of materials. Given that the male and female talkers were different for adult-directed and child-directed materials and that there were only two talkers in each set, this cannot be interpreted as a talker-sex effect. Rather, we interpret the talker effects as an indicator of talker variability in emotion communication. Similarly, significant effects of the specific emotions were observed, but these effects were not the focus of the present study. The primary question of interest here was whether there were differences between the children with hearing loss and children with normal hearing in the pattern of performance across emotions. Such an interaction between hearing status and emotion was not observed in the present study.

Analyses of d′ values derived from the confusion matrices were broadly consistent with the patterns of results obtained with accuracy scores, showing effects of developmental age, material and emotion, as well as a predictive role of vocabulary. The pattern of results obtained with the different emotions was also generally consistent with the pattern observed

with the accuracy scores. Also consistent with the accuracy scores, no interaction was observed between the specific emotion types and children's hearing status.

Within the groups of children, the patterns of results were somewhat different, with vocabulary (and/or cognition) playing a more dominant role in the children with hearing loss than in the children with normal hearing. The interactions observed in the group with hearing loss were of particular interest, and discussed further below.

The greater role of vocabulary in children with hearing loss

Our hypothesis that for children with mild to moderate hearing loss interactions between age, degree of hearing loss and material would be observed was supported only by a significant interaction between age and material. When separated into two groups, mild hearing loss and moderate hearing loss, degree of hearing loss had no effect on performance. This is consistent with findings in previous studies investigating children with moderate to severe hearing loss (Davis et al. 1986; Moeller 2000; Most & Michaelis 2012). Most and Michaelis (2012) reported that only children with profound hearing loss differed significantly from children with normal hearing in voice-emotion recognition. Since fundamental frequency is one of the dominant cues for emotion recognition (e.g. Oster & Risberg 1986; Murray & Arnott 1993), individuals with mild to moderate hearing loss are likely to have sufficient residual hearing in the lower frequency range to be able to perceive the changes in fundamental frequency necessary to distinguish one emotion from another. Our findings indicate that, despite a lack of consistent access to the full complement of speech and language cues (Tomblin et al. 2015), children with mild to moderate hearing loss are able to capitalize on this information, even in a challenging emotion recognition task (i.e., with adult-directed stimuli). The observed interaction between age and material in this population suggests a specific vulnerability to speaking style in younger children: specifically, younger children with hearing loss appeared to benefit less from the exaggerated prosody of child-directed than did older children with hearing loss. Such an interaction was not observed in the children with normal hearing groups. Further, interactions were observed between vocabulary and material in the analysis of accuracy scores, and interactions between age, vocabulary and material were observed in the analysis of reaction times. These interactions, when investigated further, showed that unlike younger children with hearing loss, vocabulary was linked to greater benefit from child-directed speech in older children with hearing loss. These analyses contrasted with those in normally hearing children, and suggests that for children with hearing loss, linguistic knowledge might need to be harnessed to a greater degree to achieve the same level of performance as their normally hearing peers.

Consistent with the present findings in children with hearing loss, Geers et al. (2013) reported correlations between indexical and linguistic information processing in children with cochlear implants. Previous studies on children with cochlear implants have also shown that NVIQ is a predictor of language progress and outcomes (Davis et al. 1986; Geers & Sedey 2011; Geers & Nicholas 2013). Further, Tinnemore et al. (2018) reported that NVIQ was a predictor of emotion recognition by children with normal hearing attending to cochlear implant simulations. Previous studies on emotion recognition by children with

hearing loss who do not use cochlear implants have not included measures of language or cognition. Given that NVIQ was correlated with vocabulary in the cohort of children with mild to moderate hearing loss in the present study (but not in the children with normal hearing), the possibility that cognitive as well as linguistic processes might be more involved in how children with hearing loss address a vocal emotion recognition task, cannot be ruled out.

Reaction time in children with normal hearing and children with mild to moderate hearing loss

Our hypothesis that longer reaction times would be observed in children with mild to moderate hearing loss, younger children, and with adult-directed speech materials was only partially supported. Generally, reaction times decreased as age increased, but no interactions were observed between reaction time and group. Our findings were in contrast to those of Husain et al. (2014) in adults indicating that participants with mild to moderate hearing loss had significantly slower reaction times on their emotion processing task than normally hearing adults. However, their task and stimuli were different from those in the present study. Reaction time data also showed interactions between vocabulary, material and age in the children with hearing loss. These findings were generally consistent with accuracy scores and d′ measures. These consistencies are promising for future research, as they indicate that response times may be used as a proxy for more time-consuming or difficult measures of emotion identification, particularly in difficult-to-test populations such as very young children. Although our instructions did not emphasize speed, results of Pals et al. (2015) suggest that reaction time measures in the present study may additionally reflect cognitive load/listening effort.

Implications for children with hearing loss

Children with mild to moderate hearing loss have been relatively under-represented in studies of voice emotion recognition. The present results point to greater vulnerability to materials/speaking style in younger children with mild to moderate hearing loss. Given the literature on psychosocial problems in children with hearing loss, possible connections between emotion understanding, theory of mind, and the development of social cognition in children, it is critical to achieve a deeper understanding of these issues.

Our results indicate that, unlike children with normal hearing, vocabulary is a significant predictor of performance and in children with hearing loss, interacting with age and with type of material (child- or adult-directed). The effect of vocabulary was supported in reaction time measures and in d′ measures as well. It is known that ensuring consistent and appropriate use and fitting of hearing aids is an important contributor to success in language development in this population (McCreery et al 2015; Tomblin et al, 2015). Thus, taking steps to ensure well-fitted hearing aids early in the child's life, together with ensuring that the child wears the hearing aid consistently and for longer hours of the day, as well as other interventions that might support language development, are all likely to contribute to improved processing of emotional prosody, particularly in younger children with hearing loss.

Conclusions

Previous studies in this area focusing on children with moderate to profound hearing loss and children with cochlear implants have reported significant deficits in voice emotion recognition when compared to normally hearing peers. Results from the present study suggest that at least in the population of children with mild to moderate hearing loss, a significant deficit in voice emotion perception does not exist, either in accuracy or in reaction time. However, a different pattern of interactions observed in accuracy scores, d's, and reaction times in the children with hearing loss suggest a greater role for language/cognition and age in their ability to benefit from the exaggerated prosody of child-directed speech.

Acknowledgments

References

Banse R, Scherer KR. Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology. 70(3):614–636.1996; [PubMed: 8851745]

Borrie SA, Mcauliffe MJ, Liss JM, O'beirne GA, Anderson TJ. 2013; The role of linguistic and indexical information in improved recognition of dysarthric speech. The Journal of the Acoustical Society of America. 133(1):474–482. DOI: 10.1121/1.4770239 [PubMed: 23297919]

Paul, Boersma. 2001; Praat, a system for doing phonetics by computer. Glot International. 5(9/10): 341–345.

Chatterjee, M; Zion, DJ; Deroche, ML; Burianek, BA; Limb, CJ; Goren, AP; … Christensen, J. Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. Hearing Research. 2015.

Cohen J. 1992; A Power Primer. Psychological Bulletin. 122:155–159.

Crandell CC. 1993; Speech recognition in noise by children with minimal degrees of sensorineural hearing loss. Ear and Hearing. 14(3):210–216. DOI: 10.1097/00003446-199306000-00008 [PubMed: 8344478]

Davis JM, Elfenbein J, Schum R, Bentler RA. 1986; Effects of mild and moderate hearing impairments on language, educational, and pychological behavior of children. Journal of Speech and Hearing Disorders. 51:53–62. [PubMed: 3945060]

Dunn, LM, Dunn, DM. Pearson Assessments. PPVT-4: Peabody picture vocabulary test. Minneapolis, MN: Pearson Assessments; 2007.

Dyck MJ, Farrugia C, Shochet IM, Holmes-Brown M. 2004; Emotion recognition/understanding ability in hearing or vision-impaired children: Do sounds, sights, or words make the difference? Journal of Child Psychology and Psychiatry and Allied Disciplines. 45(4):789–800. DOI: 10.1111/j.1469-7610.2004.00272.x

Eisenberg N, Spinrad TL, Eggum ND. 2010; Emotion-related self-regulation and its relation to children's maladjustment. Annu Rev Clin Psychol. 6:495–525. DOI: 10.1146/annurev.clinpsy.121208.131208.Emotion-Related [PubMed: 20192797]

Feldman M. 1967; Relations reaction among hearing, time, and age. Journal Of Speech And Hearing Research. 10(3):479–495. [PubMed: 6081931]

Fernald A, Simon T. 1984; Expanded intonation contours in mothers' speech to newborns. Developmental Psychology. 20(1):104–113. DOI: 10.1037/0012-1649.20.1.104

Fernald A. 1989; Intonation and Communicative Intent In Mothers' Speech to Infants: Is the Melody the Message? Child Development. 60:1497–1510. [PubMed: 2612255]

Geers AE, Sedey AL. 2011; Language and verbal reasoning skills in adolescents with 10 more years of cochlear implant experience. Ear & Hearing. 32:39S–48S. [PubMed: 21832889]

Geers A, Davidson LS, Uchanski RM, Nicholas JG. 2013; Interdependence of linguistic and indexical speech perception skills in school-aged children with early cochlear implantation. Ear and Hearing. 34(5):562–574. DOI: 10.1038/jid.2014.371 [PubMed: 23652814]

Gioia GA, Isquith PK, guy SC, Kenworthy L. 2000; Behavior rating inventory of executive function. Child Neuropsychol. 6:235–238. [PubMed: 11419452]

Hasher, L; Zacks, RT; Rose, KC; Sanft, H. Truly incidental encoding of frequency information. The American Journal of Psychology. 1987.

Hicks CB, Tharpe AM. 2002; Listening effort and fatigue in school-age children with and without hearing loss. Journal of Speech Language and Hearing Research. 45(3):573.doi: 10.1044/1092-4388(2002/046)

Husain FT, Carpenter-Thompson JR, Schmidt SA, Peelle JE, McGettigan C, Holloway R. 2014; The effect of mild-to-moderate hearing loss on auditory and emotion processing networks. Frontiers in Systems Neuroscience. 8:10.doi: 10.3389/fnsys.2014.00010 [PubMed: 24550791]

Jerger S. 2007; Current state of knowledge: Perceptual processing by children with hearing impairment. Ear & Hearing. 28(6):754–765. [PubMed: 17982363]

Johnson EK, Westrek E, Nazzi T, Cutler A. 2011; Infant ability to tell voices apart rests on language experience. Developmental Science. 14(5):1002–1011. DOI: 10.1111/j.1467-7687.2011.01052.x [PubMed: 21884316]

Jusczyk PW, Cutler A, Redanz NJ. 1993; Infants' Preference for the Predominant Stress Patterns of English Words. Child Development. 64:675–687. [PubMed: 8339688]

Juslin PN, Laukka P. 2001; Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. Emotion. 1:381. [PubMed: 12901399]

Kempe V, Schaeffler S, Thoresen JC. 2010; Prosodic disambiguation in child-directed speech. Journal of Memory and Language. 62(2):204–225. DOI: 10.1016/j.jml.2009.11.006

Ketelaar L, Rieffe C, Wiefferink CH, Frijns JHM. 2012; Does hearing lead to understanding? Theory of mind in toddlers and preschoolers with cochlear implants. Journal of Pediatric Psychology. 37(9):1041–1050. DOI: 10.1093/jpepsy/jss086 [PubMed: 22847880]

Ketelaar L, Rieffe C, Wiefferink CH, Frijns JHM. 2013; Social competence and empathy in young children with cochlear implants and with normal hearing. The Laryngoscope. 123(2):518–23. DOI: 10.1002/lary.23544 [PubMed: 22996702]

Klein KE, Walker EA, Kirby B, McCreery RW. 2017; Vocabulary facilitates speech perception in children with hearing aids. Journal of Speech, Language, and Hearing Research. 60(8)doi: 10.1044/2017_JSLHR-H-16-0086

Lambrecht L, Kreifelts B, Wildgruber D. 2014; Gender differences in emotion recognition: Impact of sensory modality and emotional category. Cognition & emotion. 28:452–469. [PubMed: 24151963]

Laugen NJ, Jacobsen KH, Rieffe C, Wichstrøm L. 2017; Emotion understanding in preschool children with mild-to-severe hearing loss. The Journal of Deaf Studies and Deaf Education. 22(2):155–163. DOI: 10.1093/deafed/enw069 [PubMed: 27881481]

Lewis DE, Valente DL, Spalding JL. 2015; Effect of minimal/mild hearing loss on children's speech understanding in a simulated classroom. Ear and Hearing. 36(1):136–144. DOI: 10.1097/AUD.0000000000000092 [PubMed: 25170780]

Li T, Fu Q-J. 2011; Voice gender discrimination provides a measure of more than pitch-related perception in cochlear implant users. International Journal of Audiology. 50(8):498–502. DOI: 10.1038/jid.2014.371 [PubMed: 21696330]

Lima CF, Alves T, Scott SK, et al. 2014; In the ear of the beholder: how age shapes emotion processing in nonverbal vocalizations. Emotion. 14:145. [PubMed: 24219391]

Ling, D. Speech and the hearing-impaired child: Theory and practice. Washington DC: The Alexander Graham Bell Association for the Deaf, Inc; 1976.

Linnankoski I, Leinonen L, Vihla M, et al. 2005; Conveyance of emotional connotations by a single word in English. Speech Communication. 45:27–39.

Ludlow A, Heaton P, Rosset D, Hills P, Deruelle C. 2010; Emotion recognition in children with profound and severe deafness: Do they have a deficit in perceptual processing? Journal of Clinical and Experimental Neuropsychology. 32(9):923–928. DOI: 10.1080/13803391003596447 [PubMed: 20349386]

Luo X, Fu Q-J, Galvin J. 2007; Vocal Emotion Recognition by Normal-Hearing Listeners and Cochlear Implant Users. Trends in Amplification. 11(4):301–315. [PubMed: 18003871]

McCreery RW, Spratford M, Kirby B, Brennan M. 2017; Individual differences in language and working memory affect children's speech recognition in noise. International journal of audiology. 56(5):306–315. [PubMed: 27981855]

Mildner V, Koska T. 2014; Recognition and production of emotions in children with cochlear implants. Clinical Linguistics & Phonetics. 28(7–8):543–554. [PubMed: 25000377]

Moeller MP. 2000; Early intervention and language development in children who are deaf and hard of hearing. Pediatrics. 106(3):E43. [PubMed: 10969127]

Moeller MP. 2007; Current state of knowledge: Psychosocial development in children with hearing impairment. Ear and Hearing. 28(6):729–39. DOI: 10.1097/AUD.0b013e318157f033 [PubMed: 17982361]

Most T, Weisel A, Zaychik A. 1993; Auditory, visual and auditory-visual identification of emotions by hearing and hearing-impaired adolescents. British Journal of Audiology. 27:247–253. [PubMed: 8312847]

Most T, Aviner C. 2009; Auditory, visual, and auditory-visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. Journal of Deaf Studies and Deaf Education. 14(4):449–464. DOI: 10.1093/deafed/enp007 [PubMed: 19398533]

Most T, Michaelis H. 2012; Auditory, visual, and auditory-visual percpetions of emotions by young children with hearing loss versus children with normal hearing. Journal of Speech, Language, and Hearing Research. 55:1148–1163. DOI: 10.1044/1092-4388(2011/11-0060)1148

Murray IR, Arnott JL. 1993; Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. J Acoust Soc Am. 93(2):1097–1108. [PubMed: 8445120]

Nelson D, Hirsh-Pasek K, Jusczyk P, Cassidy K. 1989; How the prosodic cues in motherese might assist language learning. Journal of Child Language. 16(1):55–68. DOI: 10.1017/S030500090001343X [PubMed: 2925815]

Orbelo DM, Grim MA, Talbott RE, et al. 2005; Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. Journal of Geriatric Psychiatry and Neurology. 18:25–32. [PubMed: 15681625]

Oster AM, Risberg A. 1986; The identification of the mood of a speaker by hearing-impaired listeners. SLT-Quarterly Progress Status Report. 4:79–90.

Pals C, Sarampalis A, van Rijn H, Ba kent D. 2015; Validation of a simple response-time measure of listening effort. The Journal of the Acoustical Society of America. 138(3):EL187–EL192. DOI: 10.1121/1.4929614 [PubMed: 26428811]

Paulmann S, Pell MD, Kotz SA. 2008; How aging affects the recognition of emotional speech. Brain and language. 104:262–269. [PubMed: 17428529]

Peterson CC, Siegal M. 1995; Deafness, conversation and theory of mind. J Child Psychol Psych. 36:459–474.

Peterson CC, Siegal M. 1998; Changing focus on the representational mind. Deaf, autistic and normal children's concepts of false photos, false drawings and false beliefs. Br J Devel Psychol. 16:301–320.

Pinheiro, J; Bates, D; DebRoy, S; Sarkar, D. R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-121. 2015. http://CRAN.Rproject.org/package=nlme

Pisoni, DB. Some thoughts on "normalization" in speech perception. In: Johnson, KJ, Mullenix, JW, editors. Talker Variability in Speech Processing. London, San Diego, CA: Academic Press; 1997. 9–32.

Pittman A, Vincent K, Carter L. 2009; Immediate and long-term effects of hearing loss on the speech perception of children. The Journal of the Acoustical Society of America. 126(3):1477–1485. DOI: 10.1121/1.3177265 [PubMed: 19739760]

Reber AS. 1967; Implicit Learning of Artificial Grammars. Journal of Verbal Learning and Verbal Behavior. 6:855–863. DOI: 10.1016/S0022-5371(67)80149-X

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2015. http://www.R-project.org/

Rieffe C, Terwogt MM. 2000; Deaf children's understanding of emotions: Desires take precedence. J Child Psychol Psych. 41:601–608.

Rodero EA. 2011; Intonation and emotion: influence of pitch levels and contour type on creating emotions. Journal of Voice. 25:25– 34.

Russel PA, Hosie JA, Gray CD, Scott C, Hunter N, Banks JS, Macaulay MC. 1998; The development of theory of mind in deaf children. J Child Psychol Psych. 39:309–910.

Saffran JR, Newport EL, Aslin RN, Tunick RA, Barrueco S, Saffran JR, … Barrueco S. 1997; Incidental language learning: Listening (and learning) out of the corner of your ear. Psychological Science. 8(2):101–105.

Sauter DA, Eisner F, Calder AJ, et al. 2010; Perceptual cues in nonverbal vocal expressions of emotion. The Quarterly Journal of Experimental Psychology. 63:2251–2272. [PubMed: 20437296]

Scherer KR, Banse R, Wallbott HG, et al. 1991; Vocal cues in emotion encoding and decoding. Motivation and emotion. 15:123–148.

Schick B, de Villiers PA, de Villiers J, Hoffmeister P. 2007; Language and theory of mind: A study of deaf children. Child Dev. 78:376–396. [PubMed: 17381779]

Schorr EA, Roth FP, Fox NA. 2009; Quality of life for children with cochlear implants: Perceived benefits and problems and the perception of single words and emotional sounds. Journal of Speech, Language, and Hearing Research. 52(1):141–152. DOI: 10.1044/1092-4388(2008/07-0213)

Singh L, Morgan JL, Best CT. 2002; 'Infants' listening preferences: Baby talk or happy talk? Infancy. 3:365–394.

Song JY, Demuth K, Morgan J. 2010; Effects of the acoustic properties of infant-directed speech on infant word recognition. The Journal of the Acoustical Society of America. 128(1):389–400. DOI: 10.1121/1.3419786 [PubMed: 20649233]

Steeds L, Rowe K, Dowker A. 1997; Deaf children's understanding of beliefs and desires. Journal of Deaf Studies and Deaf Education. 2(3):185–95. DOI: 10.1093/oxfordjournals.deafed.a014324 [PubMed: 15579847]

Studebaker GA. 1985; A "rationalized" arcsine transform. J Sp Lang Hear Res. 28(3):455–462.

Tinnemore AR, Zion DJ, Kulkarni AM, Chatterjee M. Children's Recognition of Emotional Prosody in Spectrally Degraded Speech Is Predicted by Their Age and Cognitive Status. Ear Hear. 2018 Jan.doi: 10.1097/aud.0000000000000546

Tomblin JB, Harrison M, Ambrose SE, Walker EA, Oleson JJ, moeller MP. 2015; Language outcomes in young children with mild to severe hearing loss. Ear & Hearing. 36(Supplement 1):76S–91S. [PubMed: 26731161]

Tukey, JW. Exploratory data analysis. Reading, Mass: Addison-Wesley Pub. Co; 1977.

Tun PA, Mccoy S, Wingfield A. 2009; Aging, hearing acuity, and the attentional costs of effortful listening. Psychology and Aging. 24(3):761–766. DOI: 10.1037/a0014802 [PubMed: 19739934]

Wang Y, Seidl A, Cristia A. 2015; Acoustic-Phonetic Differences between Infant- and Adult-directed Speech: The Role of Stress and Utterance Position. Journal of Child Language. 42(4):821–842. [PubMed: 25158975]

Wauters LN, Knoors H. 2008; Social integration of deaf children in inclusive settings. Journal of Deaf Studies and Deaf Education. 13(1):21–36. DOI: 10.1093/deafed/enm028 [PubMed: 17573356]

Wechsler, D. Psychological Corporation. WASI -II: Wechsler abbreviated scale of intelligence -- second edition. San Antonio, Tex: Psychological Corporation; 2011.

Wickham, H. ggplot2: elegant graphics for data analysis. Springer; New York: 2009.

Wiefferink CH, Rieffe C, Ketelaar L, De Raeve L, Frijns JHM. 2013; Emotion understanding in deaf children with a cochlear implant. Journal of Deaf Studies and Deaf Education. 18(2):175–186. DOI: 10.1093/deafed/ens042 [PubMed: 23232770]
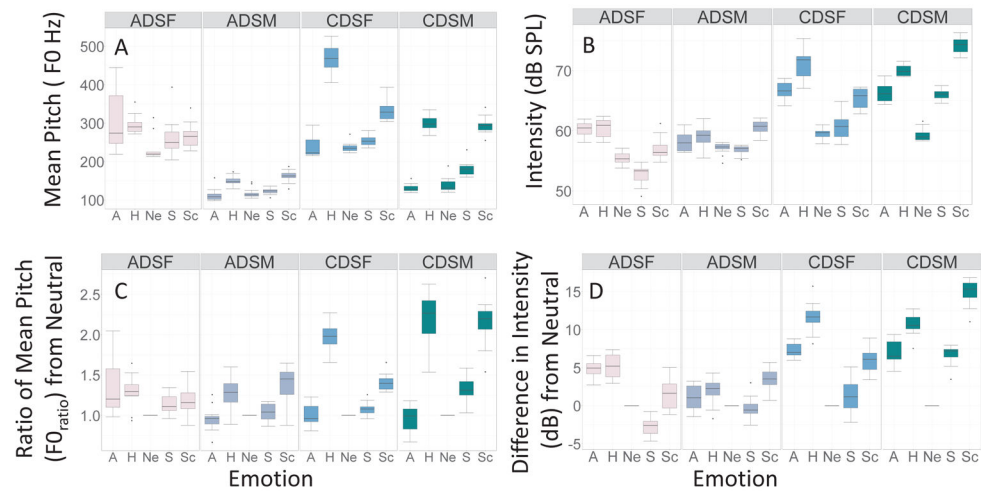
Figure 1.
A) Boxplots of mean pitch (F0 Hz) computed across all sentences spoken with each individual emotion (abcissa) by the two talkers and the two material types (shown in the four panels). B) As in Fig. 1A, but showing boxplots of mean intensity (dB SPL). C) As in Figs. 1A and 1B, but showing the ratio of mean pitch computed for each emotion to the mean pitch computed for the neutral emotion ($F0_{ratio}$). D) As in Figs. 1A, 1B and 1C, but showing the difference in intensity (dB) between individual emotions and the neutral emotion. A-angry, H-happy, Ne-neutral, S-sad, Sc-scared. ADSF-female talker, adult-directed material; ADSF/CDSF, adult/child-directed speech, female talkers; ADSM/CDSM, adult/child-directed speech, male talkers.
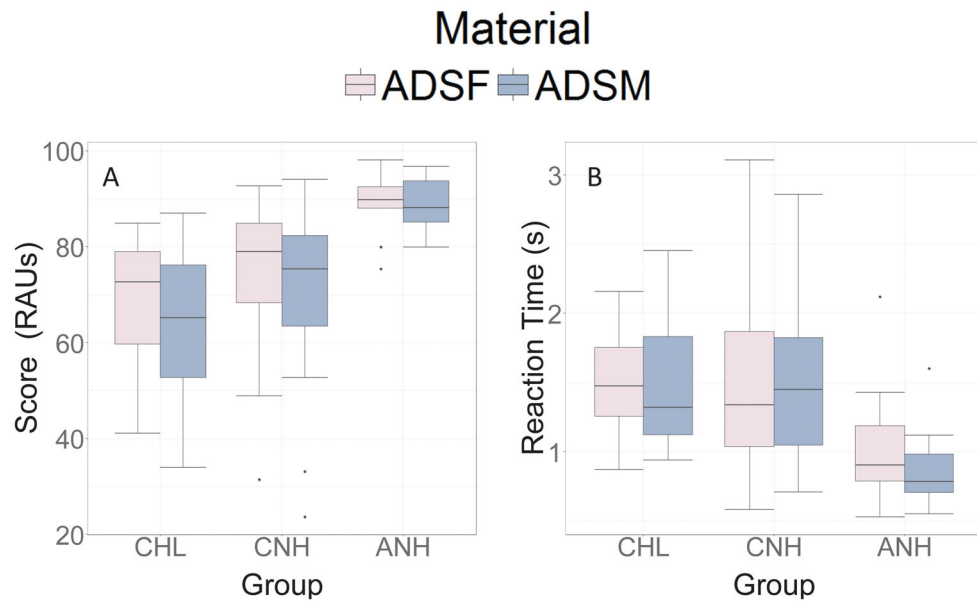
Figure 2.
Boxplots of RAU scores (A) and reaction times (B) obtained in each participant group for adult-directed materials. (CHL-children with hearing loss, CNH, children with normal hearing, ANH-adults with normal hearing; ADSF-female talker, adult-directed materials, ADSM-male talker, adult-directed materials).
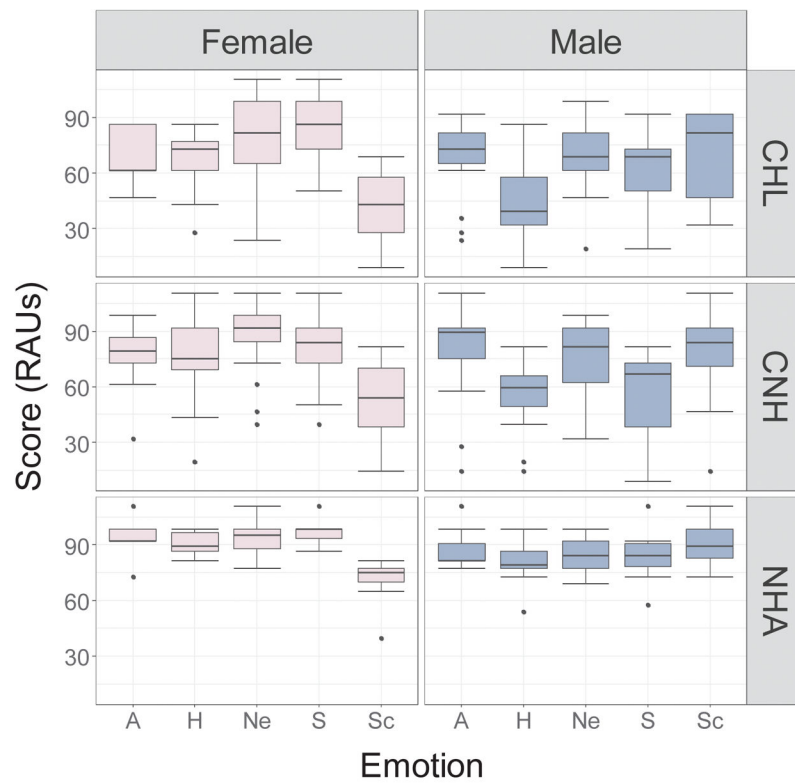
Figure 3.
Boxplots of RAU scores obtained in each participant group (rows) for adult-directed materials spoken by the female and male talker (left and right columns), plotted against the 5 emotions (abcissa). (A-angry, H-happy, Ne-neutral, S-sad, Sc-scared.; CHL-children with hearing loss, CNH-children with normal hearing, ANH-adults with normal hearing)
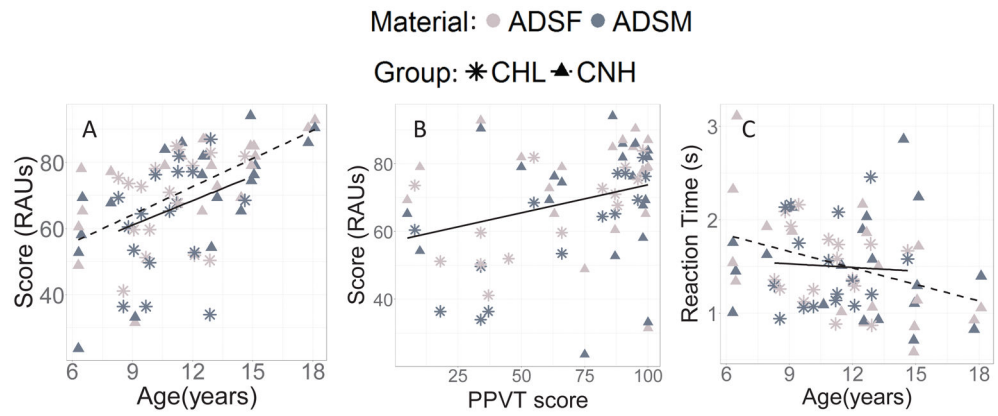
Figure 4.
Relationship between RAU scores and age (A), vocabulary score (B), and the relationship between age and reaction time (C) with adult-directed materials for children with hearing loss (CHL) and children with normal hearing (CNH). Lines indicate linear regression fits. Solid line – CHL, dashed line – CNH.
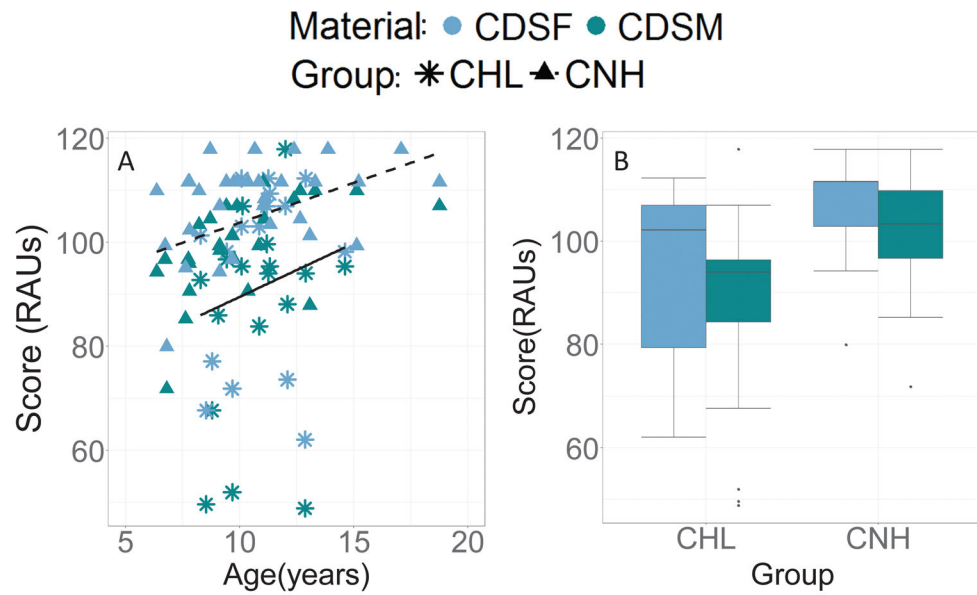
Figure 5.

Figure 5A Scatterplots of RAU scores obtained with child-directed materials in children with hearing loss (CHL) from this study and children with normal hearing (CNH) in the Chatterjee et al. (2015), plotted against the child's age. Lines indicate linear regression fits. Solid line – CHL, dashed line – CNH. **5B** Boxplots of RAU scores obtained in each group. (CDSF-female talker, CDSM-male talker).
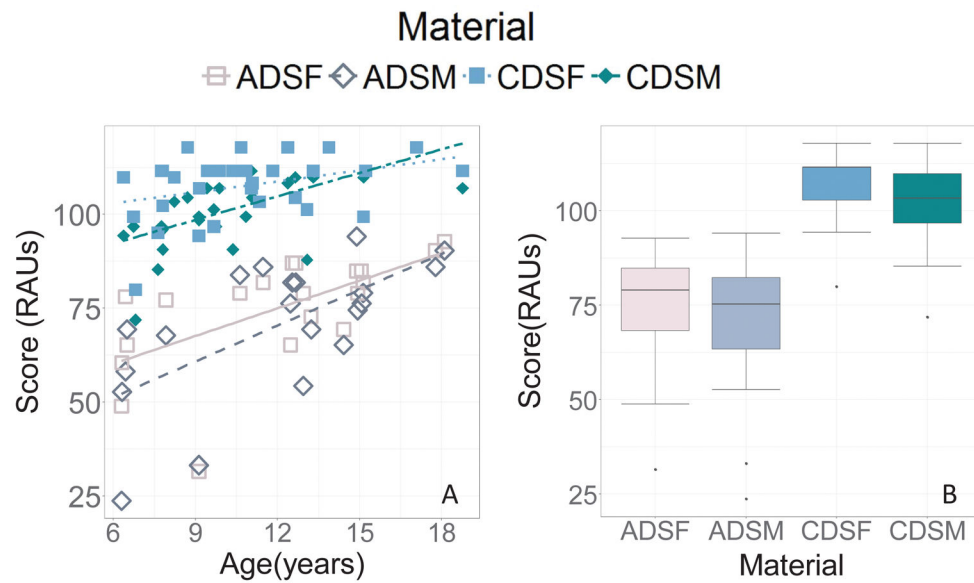
Figure 6.

Figure 6A. RAU scores plotted against age for the children with normal hearing from this study listening to adult-directed materials and the group of children with normal hearing from Chatterjee et al. (2015) listening to child-directed materials. **6B**. Boxplots of RAU scores obtained in the each group. (ADSF/CDSF, adult/child-directed speech, female talkers; ADSM/CDSM, adult/child-directed speech, male talkers). Lines indicate linear regression fits. Solid line – ADSF, dashed line – ADSM, dotted line – CDSF, two dashed line - CDSM.
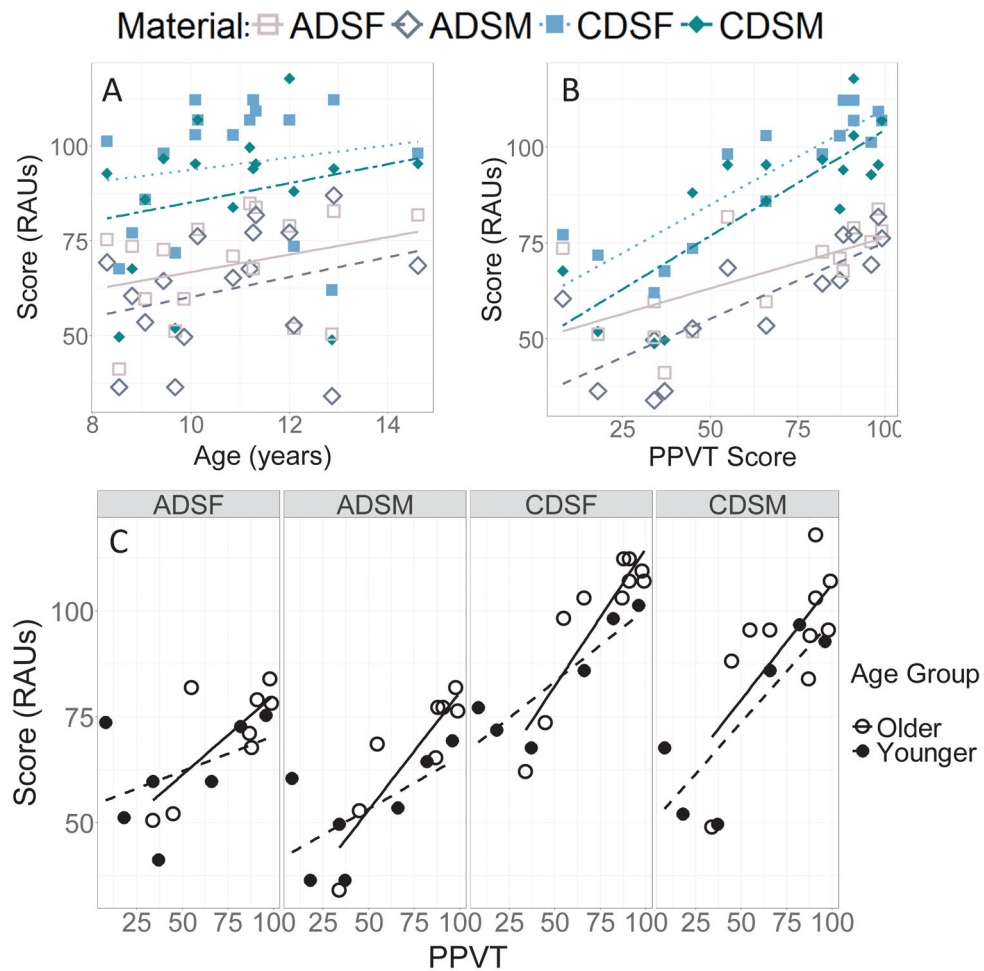
Figure 7.
Scatterplots of RAU scores against age (**7A**) and vocabulary score (**7B**) for children with hearing loss with adult-directed and child-directed materials (ADSF/CDSF-female talker, ADSM/CDSM-male talker). Fig 7C shows the relationship between RAUs and vocabulary scores for younger and older children listening to the four sets of materials (ADSF/CDSF, adult/child-directed speech, female talkers; ADSM/CDSM, adult/child-directed speech, male talkers). Lines indicate linear regression fits. Solid line – ADSF, dashed line – ADSM, dotted line – CDSF, two dashed line - CDSM.
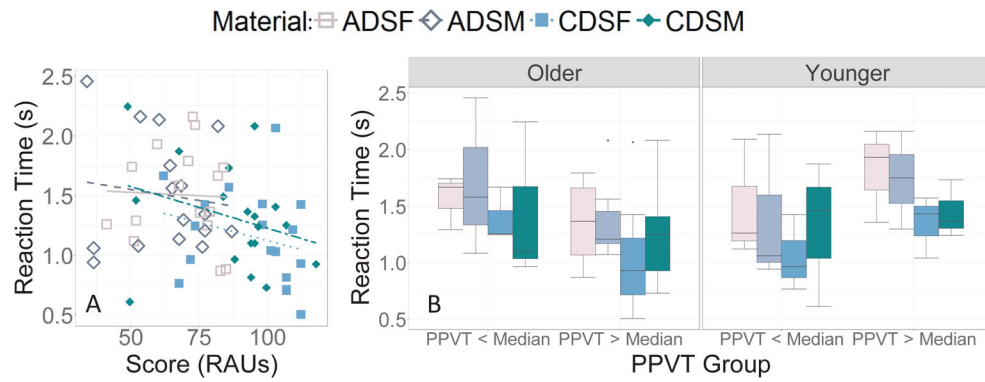
Figure 8.

Reaction times plotted against RAU scores in the children with hearing loss (**8A**). Fig 8B shows boxplots of reaction times obtained in children older than or younger than 10 years of age (left and right panels), and categorized by their vocabulary (scores greater than or less than the median PPVT score, shown in the abcissa) for children with hearing loss listening to the four sets of materials (ADSF/CDSF, adult/child-directed speech, female talkers; ADSM/CDSM, adult/child-directed speech, male talkers). Lines indicate linear regression fits. Solid line – ADSF, dashed line – ADSM, dotted line – CDSF, two dashed line - CDSM.
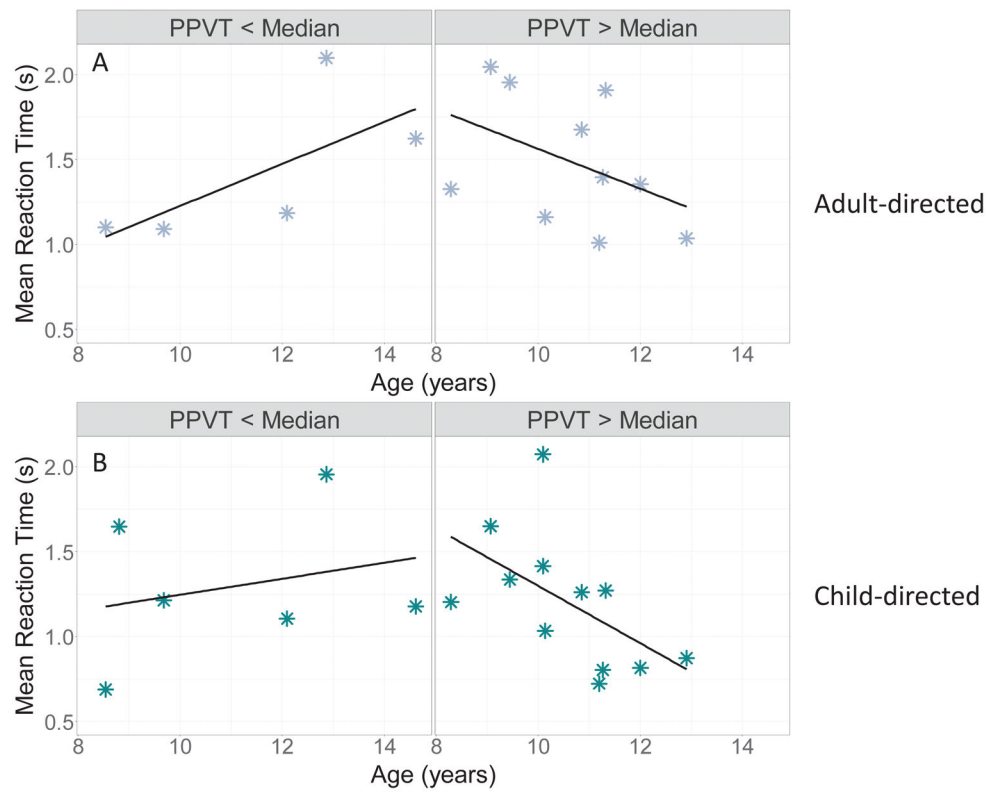
Figure 9.
Mean reaction times plotted against age for children with hearing loss categorized by vocabulary score (less than the median or greater than the median) for adult-directed materials (**9A**) and child-directed materials (**9B**). Lines indicate linear regression fits.
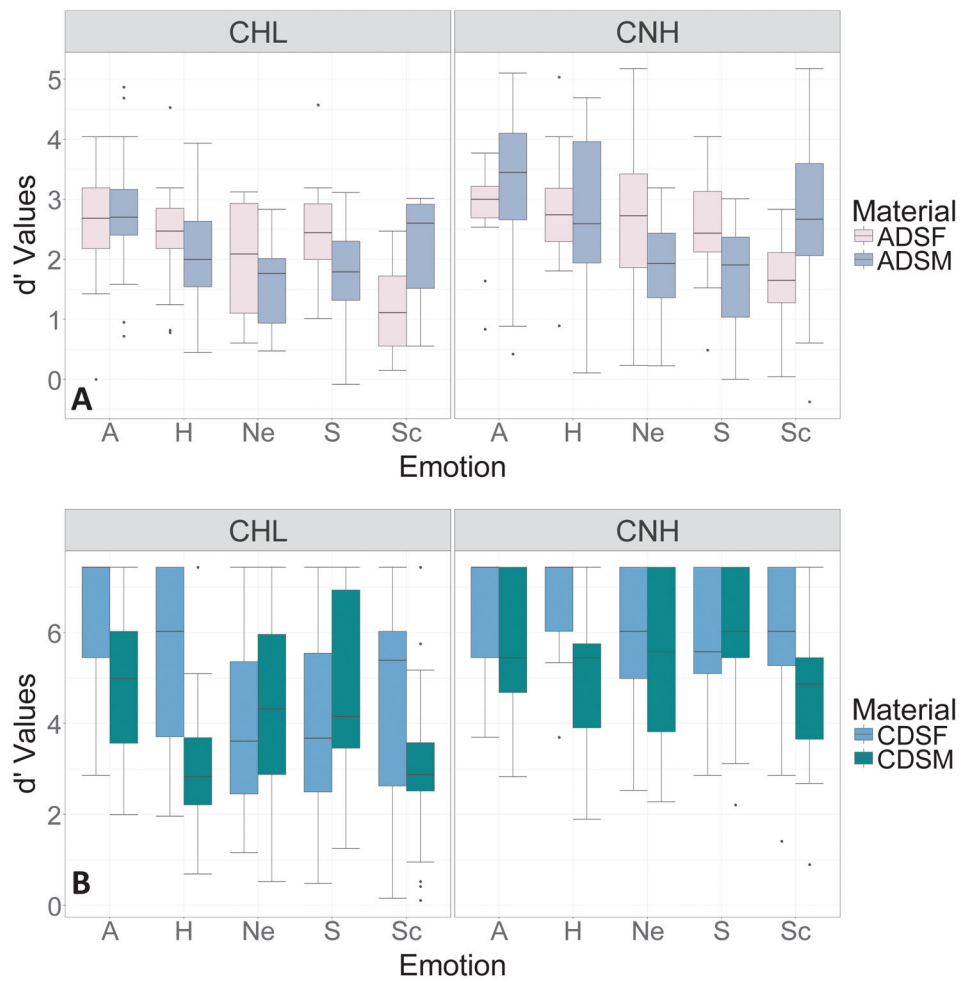
Figure 10.
Boxplots of dʹ plotted against each emotion for children with hearing loss (CHL) and children with normal hearing (CNH) with adult-directed materials (**10A**) and child-directed materials (10B). (A-angry, H-happy, Ne-neutral, Sc-scared; ADSF/CDSF-adult/child-directed speech, female talkers; ADSM/CDSM-adult/child-directed speech, male talkers).

Table 1

Relevant information about participants with hearing loss. PTA represents the average of thresholds at four frequencies (0.5, 1, 2, and 4 kHz). HI_12 had normal thresholds at low frequencies, but mild to severe hearing loss between 4 and 8 kHz.

Subject	Gender	Age (years)	Degree of hearing loss	Right Ear PTA (dB)	Left Ear PTA (dB)	Use of HA	HA used during testing
HI_01	M	10.09	mild	31.25	23.75	yes	yes
HI_02	M	10.09	mild	28.75	25.00	yes	yes
HI_03	M	8.81	moderate	30.00	37.50	yes	yes
HI_04	M	10.85	mild	28.75	18.75	yes	yes
HI_06	F	12.00	mild	26.25	27.50	no	no
HI_07	M	9.07	moderate	38.75	47.50	yes	yes
HI_08	M	10.14	mild	13.75	31.25	yes	yes
HI_09	M	14.61	moderate	46.25	46.25	no	no
HI_10	F	12.90	moderate	45.00	45.00	yes	yes
HI_11	F	11.20	moderate	53.75	42.50	yes	yes
HI_12	F	9.68	moderate	16.25	15.00	no	no
HI_13	M	8.29	mild	40.00	35.00	yes	yes
HI_14	M	11.32	moderate	48.75	48.75	yes	yes
HI_15	M	11.27	moderate	45.00	45.00	yes	yes - only wore one
HI_16	M	9.86	mild	21.25	38.75	yes	yes
HI_17	M	8.54	mild	38.75	38.75	yes	yes
HI_18	F	12.10	mild	27.50	27.50	yes	yes
HI_19	F	12.87	mild	30.00	27.50	yes	yes
HI_20	F	9.45	moderate	43.75	42.50	yes	yes