



Published in final edited form as:

*Ear Hear.* 2022 ; 43(2): 323–334. doi:10.1097/AUD.0000000000001095.

## Age-related changes in voice emotion recognition by post-lingually deafened listeners with cochlear implants

Shauntelle A. Cannon<sup>1,2</sup>, Monita Chatterjee<sup>1</sup>

<sup>1</sup>Auditory Prostheses & Perception Laboratory, Center for Hearing Research, Boys Town National Research Hospital, 555 N 30<sup>th</sup> St, Omaha, NE 68131

<sup>2</sup>Department of Speech and Hearing Science, The Ohio State University, 110 Pressey Hall, 1070 Carmack Rd, Columbus, OH 43210

### Abstract

**Objectives:** Identification of emotional prosody in speech declines with age in normally hearing (NH) adults. Cochlear implant (CI) users have deficits in the perception of prosody, but the effects of age on vocal emotion recognition by adult post-lingually deaf CI users are not known. The objective of the present study was to examine age-related changes in CI users' and NH listeners' emotion recognition.

**Design:** Participants included 18 CI users (29.6–74.5 years) and 43 NH adults (25.8–74.8 years). Participants listened to emotion-neutral sentences spoken by a male and female talker in five emotions (happy, sad, scared, angry, neutral). NH adults heard them in four conditions: unprocessed (full spectrum) speech, 16-channel, 8-channel, and 4-channel noise-band vocoded speech. The adult CI users only listened to unprocessed (full spectrum) speech. Sensitivity ( $d'$ ) to emotions and Reaction Times were obtained using a single-interval, five-alternative, forced-choice paradigm.

**Results:** For NH participants, results indicated age-related declines in Accuracy and  $d'$ , and age-related increases in Reaction Time in all conditions. Results indicated an overall deficit, as well as age-related declines in overall  $d'$  for CI users, but Reaction Times were elevated compared to NH listeners and did not show age-related changes. Analysis of Accuracy scores (hit rates) were generally consistent with  $d'$  data.

**Conclusions:** Both CI users and NH listeners showed age-related deficits in emotion identification. The CI users' overall deficit in emotion perception, and their slower response times, suggest impaired social communication which may in turn impact overall well-being, particularly so for older CI users, as lower vocal emotion recognition scores have been associated with poorer subjective quality of life in CI patients.

### Keywords

Aging; cochlear implants; vocal emotion identification; spectral degradation

## INTRODUCTION

### Emotional Prosody, Aging, and Cochlear Implants

Age-related declines in the ability to identify emotional prosody have been documented in the literature. Such declines have been observed both in individuals with normal hearing (e.g. Dupis & Pichora-Fuller, 2015; Mitchell & Kingston, 2014; Mill et al, 2009; Ryan et al., 2009; Ruffman et al, 2008; Mitchell, 2007; Orbelo et al, 2005; Orbelo et al, 2003; Brosgole & Weismann, 1995) and those with hearing loss (Christensen et al., 2019; Picou, 2016, Husain, 2014; Orbelo et al., 2005; Oster & Risberg, 1986). Although age-related hearing loss may impact listeners' access to the acoustic cues critical for emotional prosody, studies have shown that hearing thresholds do not predict the deficits seen in vocal emotion recognition by older adults with clinically normal hearing and with hearing loss (Christensen et al., 2019; Dupuis & Pichora-Fuller, 2015; Orbelo, 2005).

When compared to normally hearing individuals, cochlear implant (CI) users show significant deficits in their ability to recognize emotional prosody (Chatterjee et al., 2015; Gilbers et al., 2015; Hopyan-Misakyan et al., 2009; Luo et al, 2007). Poorer performance on emotion recognition tasks is consistent with CI users' limited access to the spectro-temporal cues that are crucial for the perception of voice pitch, a primary acoustic cue to emotional prosody (e.g., Banse & Scherer, 1996; Chatterjee & Peng, 2008; Deroche et al., 2014; Deroche et al., 2016; Oxenham, 2008; Chatterjee et al., 2015). However, the extent to which aging influences vocal emotion perception by adult CI users remains unknown.

Post-lingually deafened adult CI users show large inter-subject variability in emotion recognition outcomes (e.g., Chatterjee et al., 2015; Luo et al., 2007). The sources of this variability may lie in a number of factors other than their age, including their duration and etiology of hearing loss, age at implantation, cognitive status, etc. This can pose difficulties in interpretation of findings with CI patients. To isolate the effects of age-related changes on emotion recognition in degraded speech by listeners with an intact auditory periphery, we tested a group of normally hearing listeners spanning a similar age range as the CI users with CI-simulated versions of our stimuli. To a large extent, these simulations preserve the degradations in the speech information that CI users experience. Such simulation studies have yielded a rich body of information relevant to CI speech perception, speech intonation recognition, and speech emotion recognition (Chatterjee et al., 2015; Peng et al., 2012; Chatterjee & Peng, 2008; Luo et al., 2007; Friesen et al., 2001; Shannon et al., 1995). Some studies have shown age-related declines in speech perception outcomes in CI users as well as in normally hearing listeners attending to CI-simulated speech (Schvartz-Leyzac & Chatterjee, 2015; Roberts et al., 2013; Schvartz & Chatterjee, 2012; Schvartz et al., 2008; Chatelin et al., 2004), but how aging in normally hearing listeners impacts vocal emotion recognition in CI-simulated speech, remains unknown.

Acoustic cues for vocal emotion have been documented in the literature (Chatterjee et al., 2015; Banse & Scherer, 1996; Murray & Arnott, 1993). While emotions are conveyed through multiple cues in speech prosody, the dominant cues appear to be voice pitch and its modulations, vocal timbre, loudness, and speaking rate. Of these cues, voice pitch is most poorly represented in CIs and in CI-simulations, thus limiting performance in a number of

tasks involving the perception of music, lexical tones and speech prosody (e.g., Gaudrain & Baskent, 2018; Peng et al., 2017; Deroche et al., 2016; Chatterjee et al., 2015; Gilbers et al., 2015; Deroche et al., 2014; Peng et al., 2012; Chatterjee & Peng, 2008; Cullington et al., 2008; Galvin et al., 2007; Luo et al., 2007; Luo & Fu 2006; Fu et al., 2005; Green et al., 2004; Xu et al., 2002). Although intensity and duration cues are still represented in CIs, these cues do not fully compensate for the loss of voice pitch and related cues to emotion (Chatterjee et al., 2015; Luo et al., 2007). With increasing age, the ability to cope with degraded speech declines (Schvartz et al., 2008). In addition, sensitivity to voice pitch and its changes, particularly temporal-envelope-based cues to voice pitch (such as those represented in CIs and CI-simulations), declines in older listeners (Shen & Souza 2016; Schvartz-Leyzac & Chatterjee, 2012). It remains unknown whether CI users also show age-related declines in vocal emotion recognition, or whether similar declines may be additionally observed in aging normally hearing listeners attending to CI-simulated speech.

The objective of the present study was to explore the age-related changes in vocal emotion recognition and Reaction Time by post-lingually deafened CI users and normally hearing adults listening to speech with varying levels of spectral degradation. Previous work on vocal emotion recognition has generally focused on measures of Accuracy. Christensen et al (2019) supplemented Accuracy with measures of Reaction Time and measures of sensitivity ( $d'$ ). They reported similar patterns of vocal emotion recognition with Accuracy, Reaction Time and measures of sensitivity ( $d'$ ), but some differences were also noted. While measures of Accuracy provide a sense of overall performance, they are not bias-free. Therefore, the criterion-free  $d'$  measure of sensitivity was included in the present study. Accuracy measures were retained to allow direct comparisons to the extant literature and are presented in Supplemental Materials (see supplemental text, supplemental Tables S1–S2, & Supplemental Figures S1–S3). In addition to  $d'$ , measures of Reaction Time were included to provide an indication of processing speed and cognitive load (McMurray et al., 2017; Pals et al., 2015; Farris-Trimble et al., 2014). Studies investigating stimulus-response compatibility and its effects on P300 latencies have shown that certain aspects of stimulus processing begin to slow with age, but the time it takes to respond to a stimulus is particularly susceptible to age (Bashore et al, 1989). We anticipated that Reaction Times would be affected (lengthened) by the more degraded conditions, however, we further hypothesized that as different processes may be involved in aging, Reaction Times may also be differently impacted by age, resulting in an interaction between age and degree of spectral degradation. While some previous studies have found mixed results regarding differences between men and women in facial, gestural or spoken emotion recognition Accuracy (Hall, 1978; Thompson & Voyer, 2014), our previous study using the same stimuli as the present study showed that differences in vocal emotion recognition no longer exist once differences in hearing thresholds are accounted for (Christensen et al., 2019). Therefore, the listener's gender was not considered in the present analyses. Studies examining latencies of Reaction Time between correct and incorrect trials suggest that different cognitive processes are involved when selecting the correct answer vs making an error in discrimination tasks (e.g., Pike 1968; Estes & Wessel, 1966). Therefore, we included an additional analysis of the Reaction Time data obtained from correct and incorrect trials.

Normally hearing participants listened to stimuli that were unprocessed (full spectrum) as well as with three levels of decreasing spectral resolution. With the unprocessed (full spectrum) speech, we expected to confirm previous findings indicating age-related declines in vocal emotion recognition and age-related increases in Reaction Time in normally hearing listeners. With the spectrally degraded speech, we hypothesized that normally hearing listeners would show age-related declines in vocal emotion recognition and age-related increases in Reaction Times across degraded conditions. We also expected performance to worsen with increasingly degraded speech. Similar to our hypothesis that CI users would show smaller effects of talker than normally hearing listeners, smaller effects of talker were also expected for normally hearing listeners attending to noise-vocoded speech than to unprocessed (full spectrum) speech. Finally, we hypothesized that CI users show age-related declines in voice emotion recognition, age-related increases in Reaction Time, and that their performance with unprocessed (full spectrum) stimuli is comparable to normally hearing listeners' performance with speech at greater levels of spectral degradation.

## METHODS

### Participants

Forty-three normally hearing adults (30 Female, 13 Male; age range: 25.8–74.8 years, mean age: 47.5 years, SD 16.61 year) participated in this study, which was approved by the Boys Town National Research Hospital Institutional Review Board under protocol 11–24XP. Informed consent was obtained from all participants prior to testing. All participants were native speakers of English. Pure-tone air-conduction thresholds were obtained for all participants on the day of testing, unless participants had audiometric thresholds available from within six months of the test date. All participants were found to have thresholds of 25 dB HL or better between 250 and 4000 Hz in the better hearing ear. This criterion was modified for the older normally hearing group so that for participants over the age of 60 years, audiometric thresholds worse than 25 dB HL at 4, 6, and 8 kHz were deemed acceptable when the 4-frequency PTA (0.5, 1, 2, 4 kHz) was better than or equal to 25 dB HL. Figure 1 shows mean audiometric thresholds for the younger and older normally hearing participants plotted against frequency.

Additionally, 18 post-lingually deafened adult CI users participated in this study (9 Female, 9 Male; age range: 29.6–74.5 years, mean age: 58.7 years, SD: 13.78 years). There were no restrictions on device type or listening mode (unilateral, bilateral, bimodal). Based on self-report, patients' duration of deafness was defined as the difference in years between the age at implantation and the age at which they no longer received benefit from their hearing aids. Further information about the CI participants is available in Table 1.

CI users were tested using their first implanted ear and the configuration of their everyday MAP. One bilaterally implanted participant was tested using their later implanted device due to self-reported better speech recognition and preference for that ear. After otoscopic inspection, an earplug was placed in the contralateral ear to ensure the participant was listening with the CI only. If a CI user reported perception of sound in the contralateral ear, even with the earplug in use, sound attenuating ear muffs (Peltor 3M Optime 105) were placed in addition.

## Stimuli

Stimuli consisted of the adult-directed emotional speech recordings used in our previous studies (Cannon & Chatterjee, 2019; Christensen et al., 2019). Twelve semantically neutral sentences from the Hearing In Noise Test (HINT) corpus were recorded by one male and one female talker in one of five emotional variations (*happy, angry, sad, scared, and neutral*) for a total of 60 sentences. The sentences selected from the HINT corpus were identical to those used by Chatterjee et al. (2015). These recordings were then processed to create noise-band vocoded versions (NBV; 16-channel, 8-channel, and 4-channel) within the AngelSim™ software (Emily Shannon Fu Foundation, [www.tigerspeech.com](http://www.tigerspeech.com)) using the Greenwood frequency-place map (Greenwood, 2001) and following the method described by Shannon et al. (1995). The analysis filter settings in the simulation process used bandpass filters with 24 dB/octave slopes and the temporal envelope was lowpass filtered with a cutoff at 160 Hz. The choice of 16-channel, 8-channel and 4-channel conditions was motivated by the finding that CI users (both post-lingually deaf and prelingually deaf) show performance levels in emotion recognition that range from normally hearing listeners' performance with 4-channel to 16-channel NBV speech in the same task (Chatterjee et al., 2015).

## Procedure

Participants were seated inside a single-walled sound attenuated booth one meter away from a single loudspeaker (Grason Stadler, Inc). Stimuli were routed to the loudspeaker from the PC via an Edirol-25EX soundcard and a Crown D75 power amplifier. The normally hearing listeners completed testing with all four conditions of spectral degradation (unprocessed (full spectrum) speech, 16-, 8-, 4-channel NBV speech). The CI users participated in an identical test, but only with the unprocessed (full spectrum) speech. The stimuli were not normalized in terms of intensity across the utterances in an effort to preserve the natural intensity cues for each emotional presentation. Instead, they were presented at a mean level of 65 dB SPL calibrated with a 1 kHz pure tone that was generated with a root mean square (r.m.s.) level equal to the mean r.m.s. level computed across the utterances within the test set for each condition. The experiment was controlled by a custom Matlab program, and participants made task responses by using a mouse to click a button on the computer screen. For all participants, the initial talker (i.e., the male or the female talker) for the first condition was selected in random order. Thereafter, stimuli based on the two talkers were presented in counterbalanced order through the rest of the test conditions, and test conditions were randomized for each participant. Tests were administered in blocks of stimuli (four blocks total, blocked by talker and condition). The participants were provided with passive training prior to each block. The training included two sentences (different from the ones used in the test block) spoken in the five emotions by the same talker and in the same condition for that block, resulting in 10 passive training utterances in total. The purpose of the passive training was to familiarize the participant with the nuances in the speech and emotional productions of each talker in addition to exposing them to the different levels of spectral degradation. During the passive training, the participant heard each of the 10 utterances in turn, followed by an interval in which the correct emotion associated with that recording was highlighted on the screen. This was repeated twice. After the training, the actual test was administered in a single-interval, five-alternative forced-choice paradigm. Each utterance was played in turn, and the participant's task was to choose (by selecting an option on the screen) which of

the five emotions the talker was trying to convey. The options were displayed on the screen as cartoon images with the label of the specific emotion shown underneath. Participants were instructed to focus on Accuracy, and speed was not mentioned in the instructions. No feedback was provided, and repetitions of the stimuli were not allowed. Each block (each condition and talker) was repeated twice (not necessarily successively, depending on the randomized order). Participants responded to 60 trials per talker within each block (5 emotions, 12 sentences) for a total of 240 trials per condition (60 trials x 2 talkers x 2 repetitions). Participants were allowed to take breaks between blocks. The experiment took between two and two and a half hours to complete for the normally hearing participants. The testing time for participants with CIs, who only completed the tests with the unprocessed (full spectrum) stimuli, was an hour or less.

## Analyses

Data analyses were completed in R version 3.6.2 (R Core Team, 2019), primarily using packages *lme4* (Bates et al, 2015) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) for linear mixed-effects (LME) modeling. Figures were rendered using the package *ggplot2* (Wickham, 2009). Dependent variables considered in the analyses were  $d'$  values (computed across emotions from the confusion matrices) and Reaction Time (averaged across emotions, in seconds). The  $d'$  values were calculated as  $z(\text{hit rate}) - z(\text{false alarm rate})$  (Macmillan and Creelman, 2005). Hit rates of 1.00 or 0.00 were corrected to 0.9999 or 0.0001 (MacMillan and Kaplan, 1985).

Linear mixed effects (LME) models were built beginning with Age as the primary fixed effect, and other predictors (Talker, Degree of Spectral Degradation, Hearing Status, PTA) were introduced sequentially, alongside interactions. The categorical variables were dummy coded, with the Normal Hearing group as the reference for the CI group, the Female Talker as the reference for the Male Talker, and the Full Spectrum condition as the reference for the other degraded conditions. Comparisons of the models were completed based on chi-square analysis of the 2-log-likelihood ratio using the *anova* function in R. After exclusion of predictors and interactions that did not significantly improve model fit, random slopes were included, only if their inclusion was supported by decreases in the information criteria (Akaike's Information Criterion AIC and the Bayesian Information Criterion, BIC) and in the degrees of freedom. When the fully complex model including subject-based random slopes for the fixed effects and their interactions was underpowered and did not converge, a less complex model was selected. The final model was arrived at by removing random slope effects until convergence was achieved. Model residuals, their histograms and qqnorm plots were visually inspected to ensure that their distribution did not deviate from normality. Main effects and interactions were further investigated using pairwise t-tests as appropriate.

Outlier analyses were completed on Reaction Times using Tukey fences (Tukey, 1979). Outlier analysis was conducted separately for normally hearing and CI participants. Within the normally hearing data, analyses were conducted on spectral conditions separately. Proportion of data removed is as follows: Full spectrum, 2.35%; 16 channel, 8.33%; 8 channel, 2.35%; 4 channel, 2.38%; CI, no outliers. Additionally, outlier analyses were separately completed for the Reaction Time data obtained in correct and incorrect trials.

Proportion of data removed from correct trials by normally hearing participants is as follows: Full Spectrum, 3.49%; 16 channel, 2.33%, 8 channel, no outliers; 4 channel, 2.33%. Proportion of data removed from incorrect trials by normally hearing participants is as follows: Full Spectrum, 3.49%; 16 channel, no outliers; 8 channel, 2.33%; 4 channel, no outliers. Proportion of data removed from CI participants is as follows: correct trials, 2.78%; incorrect trials, no outliers.

Post-hoc analyses included simple linear regressions and multiple pairwise t-tests, which were conducted with the Holm correction for multiple comparisons (Holm, 1979).

For CI listeners, duration of deafness prior to implantation was also considered in preliminary analyses but showed no predictive effects.

## RESULTS

The results showed significant effects of Age, Degree of Spectral Degradation, Hearing Status, and Talker. The general patterns can be observed in Table 2, which shows the mean confusion matrices (coded in shades of grey) computed across all participants (left column) as well as for younger (<60 years) and older (> 60 years) participants separately (middle and right columns respectively) for normally hearing participants (top 4 rows) and for participants with CIs (bottom row). For the Full Spectrum condition, participants showed greater Accuracy and fewer confusions (high scores arranged along the diagonals). As spectral degradation increased or in individuals with CIs, the confusion matrices showed a more diffuse pattern, with larger numbers of confusions (higher scores away from the diagonals). Younger participants had fewer confusions than older participants. Although listener age is shown as a categorical variable in Table 2 and in Fig. 1 for illustrative purposes, age was considered a continuous variable in our statistical analyses of the data.

### Group difference between adults with normal hearing or CIs listening to unprocessed (full spectrum) stimuli

The  $d'$ s and Reaction Times obtained in normally hearing participants and participants with CIs are plotted for each talker in Figure 2 against participants' age. Correlation analysis showed that, for normally hearing participants, Reaction Times shortened with increasing  $d'$  ( $r = -0.617$ ,  $p < 0.001$ ). (A similar correlation analysis was performed between Accuracy scores and Reaction Times; note that the significance level survives the Bonferroni correction for multiple comparisons). For CI users Reaction Times were not correlated with  $d'$  values ( $p = 0.827$ ). In CI listeners, increased age may also be related to increased duration of deafness prior to implantation. To investigate this possibility, exploratory analyses were conducted on duration of deafness as a predictor of performance. Results showed no effects of the duration of deafness (defined in Methods) on  $d'$ s or Reaction Times in the CI group.

Results of linear mixed effects (LME) analyses on  $d'$  and Reaction Times are shown in Table 3, which lists only the model results for the significant effects for purposes of clarity (For results with Accuracy scores, see supplemental text, supplemental Table S1 & Supplemental Figure S1).

A LME model constructed with  $d'$  as the dependent variable, fixed effects of Age, Hearing Status (normally hearing or CI), Talker (Female or Male) and subject-based random intercepts showed a main effect of Age and Hearing Status. No main effect of Talker and no interactions were observed.

Results of parallel analysis on Reaction Times showed significant effects of Age and Hearing Status, and an Age x Hearing Status interaction. No main effect of Talker and no other interactions were observed. Post hoc analyses were completed using simple linear regression models on the normally hearing and CI data separately to investigate the Age x Hearing Status interaction. Results indicated a significant effect of Age on Reaction Times in normally hearing participants, but no effect of Age on Reaction Times in participants with CIs. These findings are consistent with the patterns observed in Fig. 2 (lower panel), where the normally hearing listeners' Reaction Times lengthen with age, while the participants with CIs show elevated Reaction Times across all ages.

### Effects of Spectral Degradation in Normally Hearing Listeners

Overall  $d'$ s and Reaction Times obtained by normally hearing participants can be viewed in Figure 3. Results of the LME analyses on these data are shown in Table 4; as in Table 3, only the significant effects are reported for the sake of clarity (For results with Accuracy scores, see supplemental text, supplemental Table S2 & Supplemental Figures S2–S3). Neither PTA (.5, 1, 2 kHz), High Frequency PTA (4, 6, 8 kHz), nor Low Frequency PTA (.25, .5, 1, 2 kHz) were found to be predictive for  $d'$  or Reaction Time, so these variables were excluded from the final models.

A LME model with  $d'$  as the dependent variable, fixed effects of Age, Talker, Degree of Spectral Degradation, and subject-based random intercepts showed significant main effects of Age, Talker, and Degree of Spectral Degradation, and two-way interactions between Age and Degree of Spectral Degradation, and Talker and Degree of Spectral Degradation (Table 4).

The two-way interaction between Age and Degree of Spectral Degradation in the  $d'$  data (Table 4) suggested that each of the degraded conditions resulted in a flatter function relating  $d'$  with Age. To obtain a deeper understanding of the effects, this was confirmed by using simple linear regression models with Age as a predictor for each spectral condition separately. The estimated coefficients varied with Degree of Spectral Degradation (beta=-0.026, s.e.=0.005 for the Full Spectrum condition; beta=-0.016, s.e.=0.004 for the 16-channel condition; beta=-0.012, s.e.=0.003 for the 8-channel condition, and beta=-0.010, s.e.=0.003 for the 4-channel condition). The lower coefficients (flatter functions) obtained with noise-vocoded speech may in part be due to floor effects, as evident from visual inspection of Figure 4A.

The two-way interaction between Talker and Degree of Spectral Degradation in the  $d'$  data (Table 4) suggested that performance with the Male Talker's stimuli decreased more from Full Spectrum to the 16- and 8-channel conditions than did performance with the Female Talker's stimuli. This was followed up by parsing the data by spectral degradation and completing pairwise t-tests with Holm's correction by Talker. As expected, results indicated



a significant difference in  $d'$  between the Female and Male talkers (Male Talker < Female Talker) for all conditions except in the Full Spectrum condition (Figure 4B).

Parallel LME analyses of Reaction Time data (Table 4) indicated a significant main effect of Age (longer Reaction Times in older participants), and a two-way interaction between Talker and Degree of Spectral Degradation. No main effects of Talker or Degree of Spectral Degradation, and no other interactions were observed.

The two-way interaction between Talker and Degree of Spectral Degradation in the Reaction Time data (Table 4) suggested that the Reaction Times increased more for the Male Talker's materials than the Female Talker's materials when the spectral degradation changed from Full Spectrum to 16 channel. This was followed up by parsing the data by Talker and completing pairwise t-tests with Holm's correction based on spectral degradation. Results indicated a difference in patterns between Talkers. For the Female talker, Reaction Times were significantly different only between the Full-Spectrum and 8-Channel condition (longer for 8-Channel than for Full Spectrum,  $p < 0.05$ ); however, for the Male Talker Reaction Times were significantly different between the Full Spectrum and 8 channel condition as well as between the 16 Channel and 8 Channel conditions (Figure 5).

**Reaction Times Separated by Correct and Incorrect Trials**—The Reaction Time data were separated into the Reaction Times for correct trials and incorrect trials. These were separately analyzed for the Full Spectrum condition to investigate the effects of group differences on Age, Talker, and Hearing Status. Results are shown in Table 5. Results of LME analysis for both correct and incorrect trials in the Full Spectrum condition showed main effects of Hearing Status and Age, and a two-way interaction between Hearing Status and Age. There were no effects of Talker and no other interactions were observed. The two-way interaction between Hearing Status and Age for both correct and incorrect trials occurs because CI users' Reaction Times were longer overall regardless of age, whereas normally hearing participants' Reaction Times for both the correct and incorrect trials became slower with increasing age (Figure 6). As seen by the estimated coefficients (Table 5) and significance levels, the effect was stronger for the incorrect trials than for the correct trials.

Similar analyses were conducted on the normally hearing participants' Reaction Times to investigate the effects of Degree of Spectral Degradation. For the correct trials, results showed a main effect of Age, and two-way interactions between Age and Degree of Spectral Degradation and between Talker and Degree of Spectral Degradation. No main effects of Talker or Degree of Spectral Degradation and no other interactions were observed. Based on the estimated coefficients (Table 5), the interaction between Age and Degree of Spectral Degradation was due to a greater lengthening of the Reaction Time with Age in the 8-channel condition than in the Full Spectrum condition. The interaction between Talker and Degree of Spectral Degradation was due to the Reaction Time lengthening more for the Male Talker in going from the Full Spectrum to the 8 Channel condition.

For the Incorrect trials, results showed only a main effect of Age, and no main effect of Talker or Degree of Spectral Degradation and no interactions.

Taken together these analyses show an overall elevation in CI users' Reaction Times for both correct and incorrect trials relative to normally hearing listeners. Relative to CI users, normally hearing listeners showed a greater effect of longer Reaction Time with Age, for both correct and incorrect trials. This resulted in an interaction, with younger CI users' Reaction Times being longer than their normally hearing counterparts, while older CI users' Reaction times were more similar to older normally hearing listeners' Reaction Times (Figure 5).

## DISCUSSION

The purpose of the present study was to investigate the effects of age on voice emotion recognition performance and Reaction Time in post-lingually deaf CI users and normally hearing adults listening to increasingly degraded speech. The findings confirmed results of previously published studies showing age-related declines in spoken emotion recognition by normally hearing adults and extended them to novel results obtained in listeners with CIs and normally hearing adults listening to CI-simulations. These results are the first to show age-related declines in adult CI users' ability to decipher vocal emotions. The results also showed elevated response times across the adult life span for CI users in general, compared to normally hearing listeners. The CI participants did not show the increase in Reaction Time with age that was observed in the normally hearing group. Analyses indicated that the interaction between Hearing Status and Age in this pattern of results was stronger for the incorrect trials than for the correct trials (larger negative estimated coefficient, Table 5), although both sets of data look similar (Figure 6). Consistent with previous reports (Christensen et al., 2019; Dupuis et al., 2015; Orbelo, 2005), aging effects in the normally hearing group were not predicted by their audiometric thresholds (high- or low-frequency PTAs).

### Group differences and Similarities for Full Spectrum Speech

In the Full Spectrum condition, results confirmed that normally hearing participants'  $d'$ s on the emotion recognition task decreased with increasing age, while Reaction Times lengthened with increasing age. As expected, the performance of CI users was poorer than normally hearing participants. An unexpected result was that CI users showed overall elevated Reaction Times relative to normally hearing listeners across age and showed no age-related change in Reaction Times. No interaction was observed between age and hearing status for  $d'$  measures, suggesting that the decline in performance with age was similar between groups. This result might be taken to indicate that the mechanisms of age-related decline are similar between the groups, but the interaction between age and hearing status in the Reaction Time data argue for a group-based difference in mechanisms, at least in terms of cognitive load. While the  $d'$  scores supported our hypothesis that CI users' emotion recognition would decline with age, the Reaction Time data did not. The elevated Reaction Times in CI users across the age range suggest that the vocal emotion recognition task is more cognitively demanding for CI users than for normally hearing listeners, to the extent that age-related benefits in younger CI patients are not observable.

## Effects of Spectral Degradation on Normally Hearing Listeners' Emotion Recognition: Comparison with CI Listeners

Considering the overall data, we found that  $d'$  and Reaction Times revealed significant effects of Age (decreased  $d'$ , longer Reaction Times, with increasing age), Talker (better  $d'$ , shorter Reaction Times for the female talker), and Degree of Spectral Degradation (decreased  $d'$ ). The analyses, followed up by simple linear regression, indicated that the two-way interaction between Age and Degree of Spectral Degradation for  $d'$  measures was due to smaller estimated coefficients (less change with Age) for the degraded conditions than the full Spectrum condition. Consistent with previous studies of emotion perception and of speech recognition in general, CI patients'  $d'$  scores in the Full Spectrum condition fell largely in between the normally-hearing listeners' performance in the 8-channel and 16-channel conditions (visual inspection of Figures 2, 4; Chatterjee et al, 2015; Friesen et al., 2001; Croghan et al., 2017). However, normally hearing participants' Reaction Times in conditions with higher spectral degradation were not consistent with CI users' Reaction Times. Even in the most degraded conditions, the normally hearing listeners were still quicker at responding than the CI group (who listened to Full Spectrum stimuli). This lengthening of Reaction Times was not predicted by age. For the normally hearing participants in the unprocessed (full spectrum) condition, the lengthening of Reaction Time was driven by age, and this increased response time for older normally hearing listeners was seen for both incorrect and correct trials.

Lengthened Reaction Times have been reported in CI users relative to normally hearing counterparts in eye-tracking tasks involving word identification (Farris-Trimble et al., 2014; McMurray et al., 2017). McMurray et al. (2017) propose that CI users adopt a “wait-and-see” approach in which they wait to accrue additional information before committing to a specific response. Winn & Moore (2018) report that pupil diameters in CI patients remain dilated well after a trial is completed, suggesting that CI listeners spend some time after the event reconsidering what they heard. The stronger effect for incorrect trials (larger negative coefficient of the Age x Hearing Status interaction term in Table 5) in CI users in our study suggest a greater uncertainty about what was heard, in which they might be considering and reconsidering the auditory percept before committing to a response. Another possibility includes effects of prolonged hearing loss, which may alter central auditory pathways and processing efficiency in ways that increase processing time in CI users. Although duration of deafness prior to implantation did not predict outcomes in this study, hearing loss in itself may cause changes in the central auditory system specific to emotion processing (Husain et al., 2014).

Analyses of Accuracy scores (see supplemental text, supplemental Tables S1–S2, & supplemental Figures S1–S3) show broadly consistent patterns relative to the analyses presented here on the  $d'$  values. However, there are also some differences, indicating that a different picture may be obtained when an index derived from the full confusion matrix is considered in the analyses.

In contrast to normally hearing listeners in the spectrally degraded condition, the CI participants did not show talker-variability effects. A consideration of the interactions observed between Talker and Degree of Spectral Degradation observed in the normally

hearing listeners may place the data obtained in the CI listeners within a useful framework. Thus, normally hearing listeners may be assumed to be a more homogeneous group than the CI group, even though the middle-aged and older normally hearing listeners may be different from younger listeners in both peripheral and central processing. The CI group, on the other hand, is likely to be highly heterogeneous in their peripheral neural coding, central pathways, etc., based on differences in etiology and audiological history. Individual CI patients may have peripheral channel interaction levels that correspond to a variety of spectral degradation conditions: some may have extremely broad channels equivalent to the 4-channel condition, while others may have much less channel interaction, similar to the 16-channel condition (Chatterjee et al., 2015). Thus, different CI patients may have different sensitivity to Talker based variations, obscuring the effects of these factors in a group analysis. Taken together, the results obtained with the normally hearing listeners suggest that the degree to which acoustic cues to emotions are altered by spectral degradation varies from talker to talker. However, a three-way interaction with Age was not observed, suggesting that older and younger listeners did not differ substantially in their ability to cope with the combined and interactive effects of Talker and Degree of Spectral Degradation.

### Limitations

First, it must be noted that it is difficult to recruit young, post-lingually deafened CI users. In this study there were only three under the age of 35, a significant limitation in a study of aging effects. Indeed, a *t*-test showed a significant difference between the ages of the participants in the normally hearing and CI groups ( $p=0.018$ ). Further, analyses separating the data into the older ( $> 60$  years of age) and younger ( $< 60$  years of age) groups showed significant age effects in both normally hearing ( $d'$  and Reaction Time measures) and CI ( $d'$  scores) listener groups, a reassuring finding given the relatively small sample size. Finally, if the older age of the majority of CI participants had been a driving factor, we would have seen an interaction between Hearing Status and Age. We did not see such an interaction. The interaction was observed for Reaction Time, but it was driven by overall elevated Reaction Times for the CI group independent of age. Thus, we conclude that overall age differences between groups were not a major factor in our findings.

A second limitation was that cognitive screening measures were not completed, and therefore the potential effect of cognition cannot be ruled out as a factor contributing to the age-related effects seen. Third, the speech stimuli used here only consisted of one male and one female talker, and therefore we cannot generalize the talker effects present to all male or female talkers. Additionally, no measures of psychoacoustic sensitivity to voice pitch cues or other cues to emotion were made and thus specific mechanisms underlying the observed effects cannot be elucidated. Finally, it is possible that the effect of spectral degradation on normally hearing participants would have been ameliorated by providing them with extensive training with noise-vocoded speech. Such efforts, although time-consuming, would have the advantage of allowing a more reasonable comparison between normally hearing and CI groups. However, it needs to be acknowledged that CI simulations do not replicate the effects of auditory deprivation and damage, or the effects of electrical stimuli, which evoke different responses in the auditory system than natural acoustic stimuli. Thus, CI simulations can at best provide a reference framework to aid

interpretation of results obtained in actual CI users and should not be considered as a true representation of the listening experience of CI patients.

## Conclusions

The results of this study confirmed that age-related changes in voice emotion recognition are present in normally hearing listeners with unprocessed (full spectrum) speech stimuli. Additionally, similar effects of age are seen for  $d'$  and Reaction Time for normally hearing adults listening to spectrally degraded, CI-simulated speech stimuli. Furthermore, age-related changes are present for  $d'$  in CI users. Reaction Times for CI users are delayed compared to normally hearing listeners with unprocessed (full spectrum) speech materials and appear to be even longer than when normally hearing participants were presented with the most spectrally degraded materials. These findings are important for real-time conversation dynamics as voice emotion recognition has been found to predict subjective quality of life for adult CI users (Luo et al., 2018). Impaired, slowed ability to decipher spoken emotions would not only result in less effective communication overall, it would also impede the normal rhythm and flow of natural conversation. These reductions in the quality and quantity of social communication may have cumulative effects, resulting in social withdrawal, smaller social circles, and reduced overall quality of life, not only for CI patients, but also for older normally hearing individuals.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

The authors would like to thank Aditya Kulkarni, Julie Christensen, Jenni Sis, Rizwan Siddiqui, Adam Bosen, Mohsen Hozan, and Harley Wheeler for their help with this work. The authors would also like to thank the participants for their time and assistance in the study. Portions of this work were presented at CI Crash 2017 (Madison, WI), ARO 2018 (San Diego, CA), AAS 2018 (Scottsdale, AZ), ARO 2019 (Baltimore, MD), and ASA 2019 (Louisville, KY). This research was funded by the National Institutes of Health (NIH) grants R01 DC014233 and P20 GM10923. S. C. was supported by NIH grant R01 DC014233 04S1. The authors have no conflicts of interest to disclose.

## REFERENCES

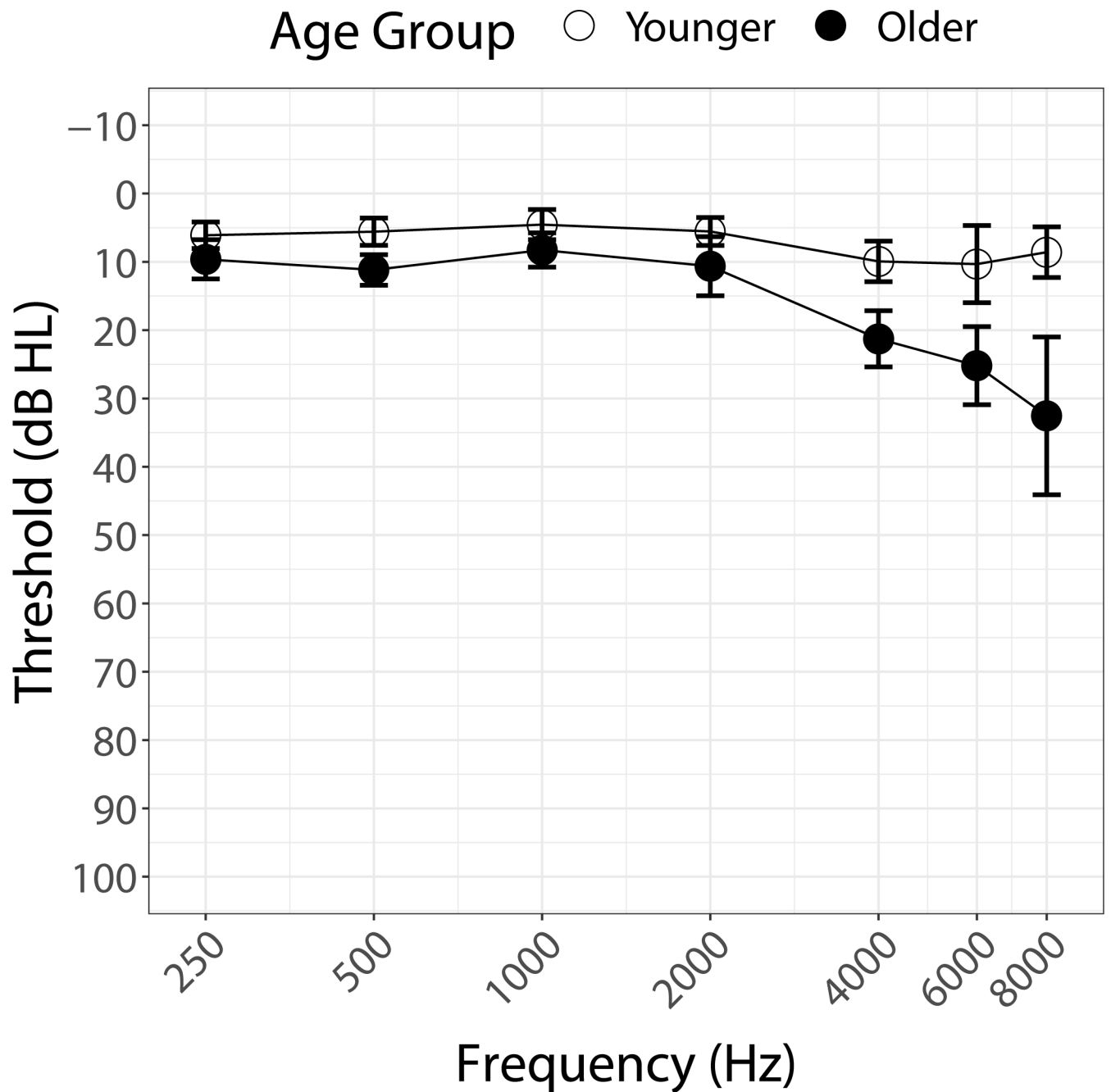
- Banse R, & Scherer KR (1996). Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol*, 70, 614–636. DOI: 10.1037//0022-3514.70.3.614 [PubMed: 8851745]
- Bashore TR, Osman A, & Heffley III EF (1989). Mental slowing in elderly persons: A cognitive psychophysiological analysis. *Psychology and Aging*, 4(2), 235–244. [PubMed: 2789753]
- Bates D, Maechler M, Bolker Ben., & Walker S (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01.
- Brosigole L & Weisman J (1995) Mood Recognition Across the Ages, *International Journal of Neuroscience*, 82(3–4), 169–189, DOI: 10.3109/00207459508999800
- Cannon SA, & Chatterjee M (2019). Voice Emotion Recognition by Children With Mild-to-Moderate Hearing Loss. *Ear and Hearing*, 40(3), 477–492. doi.org/10.1097/aud.0000000000000637 [PubMed: 30074504]
- Chatelin V, Kim EJ, Driscoll C, Larky J, Polite C, Price L, & Lalwani AK (2004). Cochlear implant outcomes in the elderly. *Otology & Neurotology*, 25(3), 298–301. doi.org/10.1002/lary.23676 [PubMed: 15129109]

- Chatterjee M, Peng S-C (2008). Processing F0 with cochlear implants: modulation frequency discrimination and speech intonation recognition. *Hear. Res.* 235, 143–156. DOI: 10.1016/j.heares.2007.11.004 [PubMed: 18093766]
- Chatterjee M, Zion DJ, Deroche ML, Burianek BA, Limb CJ, Goren AP, ... Christensen JA (2015). Voice emotion recognition by cochlear-implanted children and their normally hearing peers. *Hearing Research*, 322, 151–162. DOI: 10.1016/j.heares.2014.10.003 [PubMed: 25448167]
- Christensen JA, Sis J, Kulkarni AM, & Chatterjee M (2019). Effects of Age and Hearing Loss on the Recognition of Emotions in Speech. *Ear and Hearing*, 40(5):1069–1083 doi: 10.1097/AUD.0000000000000694 [PubMed: 30614835]
- Cullington HE & Zeng FG (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear implant, and implant simulation subjects. *J Acoust Soc Am*, 123(1), 450–461. doi: 10.1121/1.2805617. [PubMed: 18177173]
- Deroche ML, Lu HP, Limb CJ, Lin YS, & Chatterjee M (2014). Deficits in the pitch sensitivity of cochlear-implanted children speaking English or Mandarin. *Frontiers in neuroscience*, 8, 282. doi:10.3389/fnins.2014.00282. [PubMed: 25249932]
- Deroche ML, Kulkarni AM, Christensen JA, et al. (2016). Deficits in the sensitivity to pitch sweeps by school-aged children wearing cochlear implants. *Front Neurosci*, 10, 73. DOI: 10.3389/fnins.2016.00073 [PubMed: 26973451]
- Dupuis K & Pichora-Fuller MK (2015). Aging Affects Identification of Vocal Emotions in Semantically Neutral Sentences. *J Speech Lang Hear Res*, 58(3), 1061–1076. doi: 10.1037/a0018777 [PubMed: 25810032]
- Estes WK, & Wessel DL (1966). Reaction Time in relation to display size and correctness of response in forced-choice visual signal detection. *Perception & Psychophysics*, 1(5), 369–373.
- Farris-Trimble A, McMurray B, Cigrand N, & Tomblin JB (2014). The process of spoken word recognition in the face of signal degradation. *Journal of experimental psychology. Human perception and performance*, 40(1), 308–327. doi:10.1037/a0034353 [PubMed: 24041330]
- Friesen LM, Shannon RV, Baskent D, Wang X (2001). Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am*, 110, 1150e1163. DOI: 10.1121/1.1381538 [PubMed: 11519582]
- Fu QJ, Chinchilla S, Nogaki G, Galvin JJ (2005) Voice gender identification by cochlear implant users: The role of spectral and temporal resolution. *J. Acoust. Soc. Am*, 118, 1711–1718. [PubMed: 16240829]
- Fu Q-J (2010). TigerCIS: Cochlear Implant and Hearing Loss Simulation (Version 1.05.03; now titled AngelSim) [Computer program]. Available online at: [www.tigerspeech.com/tsttigercis.html](http://www.tigerspeech.com/tsttigercis.html)
- Galvin JJ, Fu QJ, & Nogaki G (2007). Melodic contour identification by cochlear implant listeners. *Ear Hear*, 28(3), 302–219. doi: 10.1097/01.aud.0000261689.35445.20 [PubMed: 17485980]
- Gaudrain E, & Ba kent D (2018). Discrimination of Voice Pitch and Vocal-Tract Length in Cochlear Implant Users. *Ear and hearing*, 39(2), 226–237. doi:10.1097/AUD.0000000000000480 [PubMed: 28799983]
- Gilbers S, Fuller C, Gilbers D, Broersma M, Goudbeek M, Free R, & Ba kent D (2015). Normal-Hearing Listeners' and Cochlear Implant Users' Perception of Pitch Cues in Emotional Speech. *I-Perception*. 6(5), 1–19. 10.1177/0301006615599139 [PubMed: 26034566]
- Green KM, Bhatt YM, Saeed SR, & Ramsden RT (2004). Complications following adult cochlear implantation: Experience in Manchester. *J Laryngol Otol*, 118(6), 417–420. DOI: 10.1258/002221504323219518 [PubMed: 15285857]
- Greenwood DD (1990). A cochlear frequency-position function for several species—29 years later. *J Acoust Soc Am*, 87(6), 2592–605. DOI: 10.1121/1.399052 [PubMed: 2373794]
- Wickham H. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009
- Hall JA (1978). Gender effects in decoding non-verbal cues. *Psychol Bull*, 85, 845–857.
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand. J. Stat* 6 (2), 65–70.
- Hopyan-Misakyan TM, Gordon KA, Dennis M, Papsin BC, 2009. Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child. Neuropsychol*, 15, 136–146. DOI: 10.1080/09297040802403682 [PubMed: 18828045]

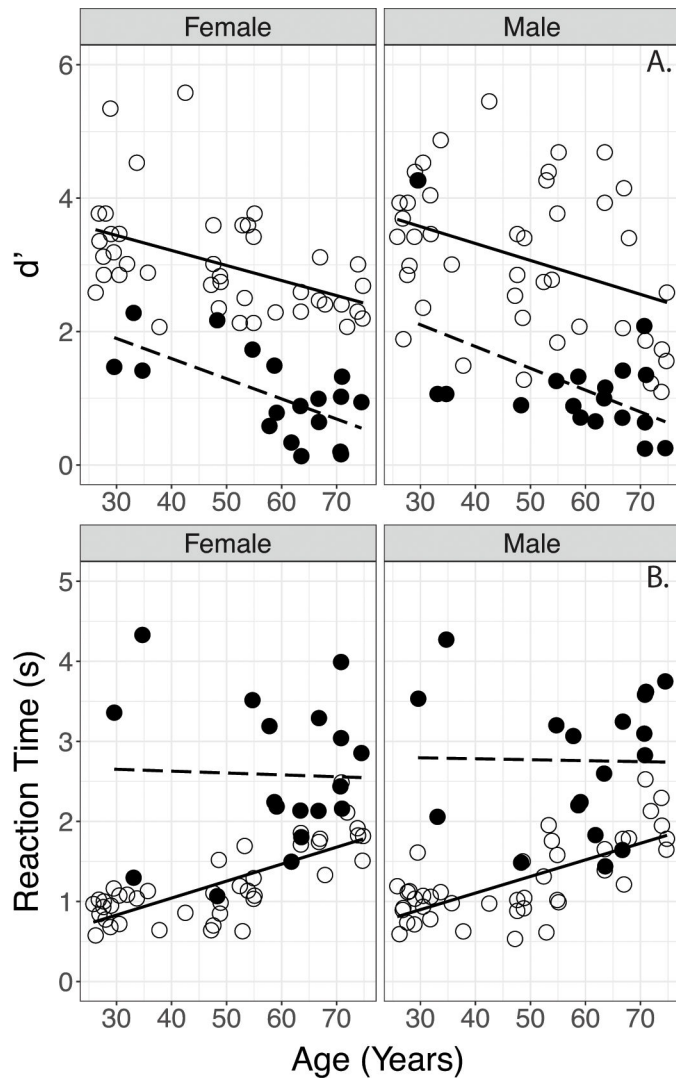
- Husain FT, Carpenter-Thompson JR, & Schmidt SA (2014). The effect of mild-to-moderate hearing loss on auditory and emotion processing networks. *Frontiers in Systems Neuroscience*, 8, 10. doi:10.3389/fnsys.2014.00010 [PubMed: 24550791]
- Kuznetsova A, Brockhoff PB, Christensen RHB (2017). “lmerTest Package: Tests in Linear Mixed Effects Models.” *Journal of Statistical Software*, 82(13), 1–26. doi: 10.18637/jss.v082.i13
- Luo X, Kern A, & Pulling KR (2018). Vocal emotion recognition performance predicts the quality of life in adult cochlear implant users. *J. Acoust Soc Am*, 144(5), EL429–EL435. [PubMed: 30522282]
- Luo X (2016). Talker variability effects on vocal emotion recognition in acoustic and simulated electric hearing. *J. Acoust. Soc. Am*, 140(6), EL497–EL503. [PubMed: 28040001]
- Luo X & Fu Q-J. (2006). Contribution of low-frequency acoustic information to Chinese speech recognition in cochlear implant simulations. *J. Acoust. Soc. Am*, 120(4), 2260. doi: 10.1121/1.2336990 [PubMed: 17069321]
- Luo X, Fu Q-J, & Galvin JJ (2007). Vocal Emotion Recognition by Normal-Hearing Listeners and Cochlear Implant Users. *Trends in Amplification*, 11(4), 301–315. DOI: 10.1177/1084713807305301 [PubMed: 18003871]
- Macmillan NA, & Kaplan HL (1985). Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. *Psychological bulletin*, 98(1), 185. [PubMed: 4034817]
- Macmillan NA, & Creelman CD (2005). *Signal Detection theory: A user’s guide* (2nd ed.). Mahwah, NJ: Erlbaum
- McMurray Bob & Farris-Trimble Ashley & Rigler Hannah. (2017). Waiting for lexical access: Cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition*, 169, 147–164. DOI: 10.1016/j.cognition.2017.08.013. [PubMed: 28917133]
- Mill A, Allik J, Realo A, et al. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion*, 9(5), 619–630. doi: 10.1037/a0016562. [PubMed: 19803584]
- Mitchell R (2007). Age-related decline in the ability to decode emotional prosody: Primary or secondary phenomenon? *Cognition & Emotion*, 21, 1435–1454. 10.1080/02699930601133994.
- Mitchell RLC, & Kingston RA (2014). Age-Related Decline in Emotional Prosody Discrimination: acoustic correlates. *Experimental Psychology*, 61(3), 215–223. doi:10.1027/1618-3169/a000241 [PubMed: 24217140]
- 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, 1097–1108. DOI: 10.1121/1.405558 [PubMed: 8445120]
- Orbelo DM, Testa JA, & Ross ED (2003). Age-related impairments in comprehending affective prosody: Comparison to brain damaged subjects. *Journal of Geriatric Psychiatry and Neurology*, 16, 44–52. DOI: 10.1177/0891988702250565 [PubMed: 12641373]
- 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. *J Geriatr Psychiatry Neurol*, 18, 25–32. DOI: 10.1177/0891988704272214 [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.
- Oxenham AJ (2008). Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends Amplif*, 12(4), 316–331. [PubMed: 18974203]
- Pals C, Sarampalis A, van Rijn H, et al. (2015). Validation of a simple response-time measure of listening effort. *J Acoust Soc Am*, 138, 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. doi: 10.1016/j.bandl.2007.03.002 [PubMed: 17428529]
- Peng S-C, Chatterjee M, & Lu N (2012). Acoustic Cue Integration in Speech Intonation Recognition with Cochlear Implants. *Trends in Amplification*, 16(2), 67–82. doi: 10.1177/1084713812451159 [PubMed: 22790392]
- Peng SC, Lu HP, Lu N, Lin YS, Deroche M, & Chatterjee M (2017). Processing of Acoustic Cues in Lexical-Tone Identification by Pediatric Cochlear-Implant Recipients. *JSLHR*, 60(5), 1223–1235. doi:10.1044/2016\_JSLHR-S-16-0048 [PubMed: 28388709]

- Picou EM (2016). How hearing loss and age affect emotional responses to nonspeech sounds. *JSLHR*, 59, 1233–1246. DOI: 10.1044/2016\_JSLHR-H-15-0231 [PubMed: 27768178]
- Pike AR (1968). Latency and relative frequency of response in psychophysical discrimination. *British J Math Stat Psychol*, 21(2), 161–182.
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Roberts DS, Lin HW, Herrmann BS and Lee DJ (2013). Differential cochlear implant outcomes in older adults. *The Laryngoscope*, 123(8), 1952–1956. doi:10.1002/lary.23676 [PubMed: 23737286]
- Ruffman T, Henry JD, Livingstone V, & Phillips LH (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*, 32, 863–881. [PubMed: 18276008]
- Ryan M, Murray J, & Ruffman T (2009). Aging and the Perception of Emotion: Processing Vocal Expressions Alone and With Faces, *Experimental Aging Research*, 36(1), 1–22, DOI: 10.1080/03610730903418372
- Schwartz KC, Chatterjee M, & Gordon-Salant S (2008). Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. *J. Acoust. Soc. Am*, 124(6), 3972–3988. DOI: 10.1121/1.2997434 [PubMed: 19206821]
- Schwartz KC & Chatterjee M (2012). Gender Identification in Younger and Older Adults. *Ear and Hearing*, 33(3), 411–420. DOI: 10.1097/AUD.0b013e31823d78dc [PubMed: 22237163]
- Schwartz-Leyzac KC & Chatterjee M (2015). Fundamental-frequency discrimination using noise-band-vocoded harmonic complexes in older listeners with normal hearing. *J. Acoust. Soc. Am*, 138(3), 1687–1695. DOI: 10.1121/1.4929938 [PubMed: 26428806]
- Shannon RV, Zeng FG, Kamath V, Wygonski J, & Ekelid M (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303–304. DOI: 10.1126/science.270.5234.303 [PubMed: 7569981]
- Shen J, Wright R, & Souza PE (2016). On Older Listeners' Ability to Perceive Dynamic Pitch. *JSLHR*, 59(3), 572–582. doi:10.1044/2015\_JSLHR-H-15-0228 [PubMed: 27177161]
- Thompson AE, & Voyer D (2014). Sex differences in the ability to recognize non-verbal displays of emotion: A meta-analysis. *Cogn Emot*, 28, 1164–1195. [PubMed: 24400860]
- Tinnemore A, Zion DJ, Kulkarni AM, & Chatterjee M (2018). Children's Recognition of Emotional Prosody in Spectrally-Degraded Speech is Predicted by Their Age and Cognitive Status. *Ear & Hearing*, 39, 874–880. DOI: 10.1097/AUD.0000000000000546 [PubMed: 29337761]
- Tukey JW (1977). *Exploratory Data Analysis*. Reading, MA: Addison- Wesley Pub. Co.
- Xu L, Tsai Y, & Pfingst BE (2002). Features of stimulation affecting tonal-speech perception: implications for cochlear prostheses. *J. Acoust. Soc. Am*, 112(1), 247–258. doi:10.1121/1.1487843 [PubMed: 12141350]

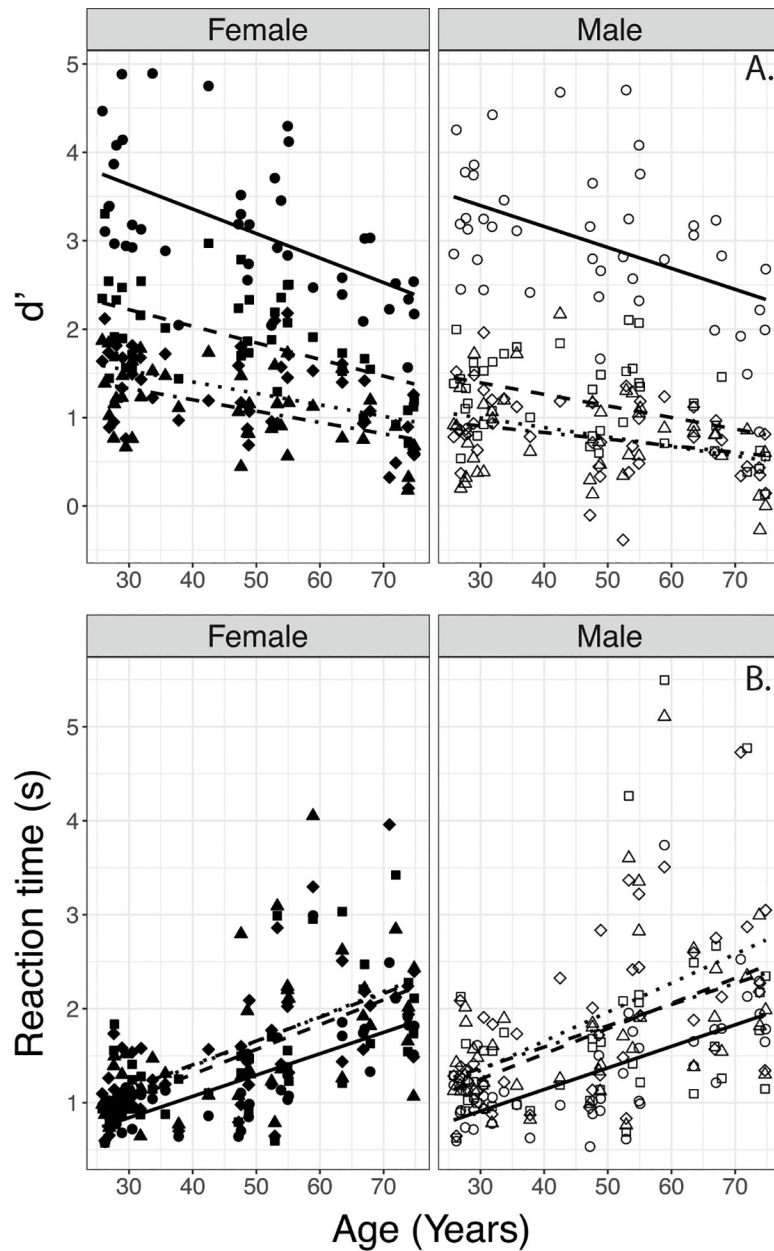




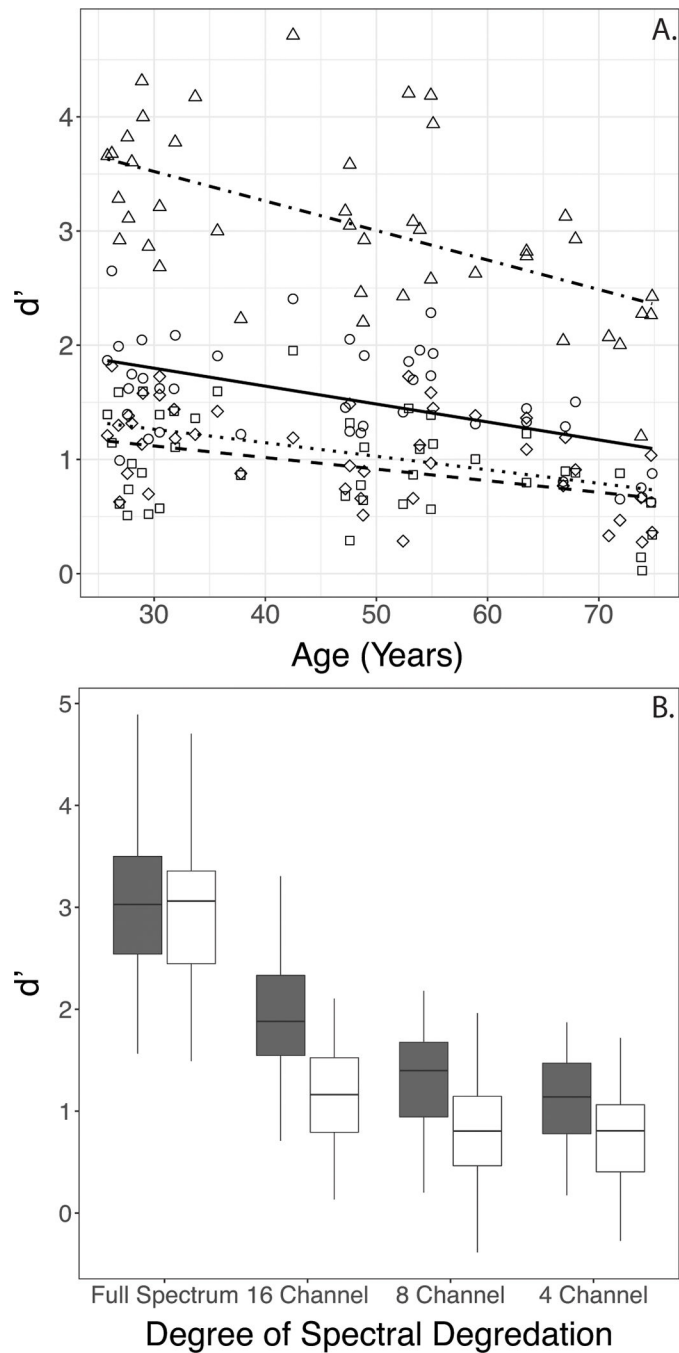
**Figure 1.** Pure-tone air-conduction thresholds (dB HL) averaged across participants from 250–8000 Hz by younger (< 60 years, open circles) and older (> 60 years, filled circles) normally hearing participants. Standard deviations are shown as error bars for each frequency.



**Figure 2.** Comparison of  $d'$  (top) and reaction times (bottom) obtained in normally hearing and CI participants. Data are separated by Female (left panel) and Male Talker (right panel). Lines indicate linear regressions. (Open circles/solid line: normally hearing; filled circles/dashed line: CI participants).



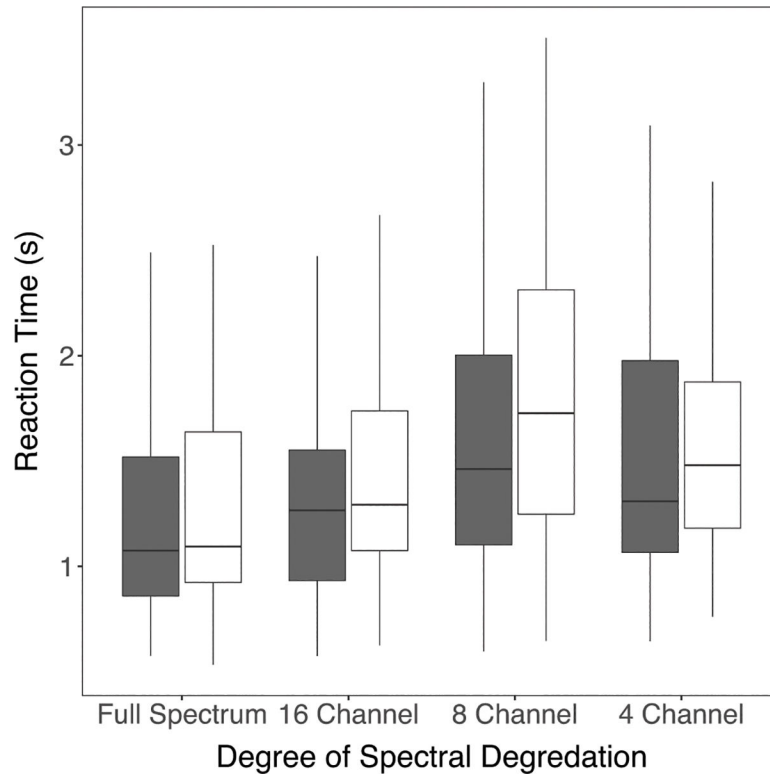
**Figure 3.** Effects Age on  $d'$  (top) and reaction times (bottom) with the four degrees of spectral degradation (shapes) by normally hearing participants. Data are separated by Female (left panel) and Male Talker (right panel). Lines indicate linear regressions. (Circles/solid line - Full Spectrum, squares/dashed line - 16 Channel, diamonds/dotted line - 8 Channel, triangles/dot dashed line - 4 channel)



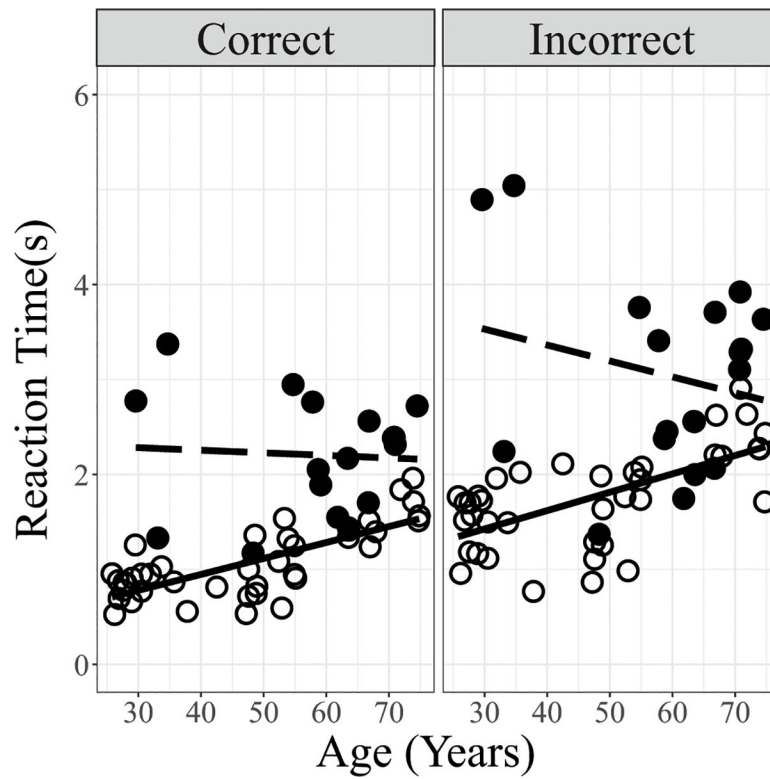
**Figure 4.**

4A. Two-way interaction between Age and Degree of Spectral Degradation (shapes) in  $d'$  data. Lines indicate linear regressions. (Circles/solid line - Full Spectrum, squares/dashed line - 16 Channel, diamonds/dotted line - 8 Channel, triangles/dot dashed line - 4 channel)

4B. Boxplots showing two-way interaction between Talker and Degree of Spectral Degradation in  $d'$  data (filled = Female; open = Male).



**Figure 5.** Boxplots showing two-way interaction between Talker and Degree of Spectral Degradation in reaction time data (filled = Female; open = Male).



**Figure 6.** Comparison of reaction time data (averaged across the two talkers) obtained in normally hearing and CI participants plotted against Age. Data are separated into Correct (left panel) and Incorrect trials (right panel). Lines indicate linear regressions. (Open circles/solid line: normally hearing; filled circles/dashed line: CI participants).

**Table 1.**

Detailed information about CI participants.

Participant	Sex	Age at testing (years)	Duration of Test Ear Device Use (~years)	CI Manufacturer	Test Ear	Everyday Listening Mode (Unilateral, Bilateral, Bimodal)
CI_01	M	70.7	15	Advanced Bionics	Left	Unilateral
CI_02	F	61.8	4	Advanced Bionics	Left	Unilateral
CI_03	M	70.8	2	Advanced Bionics	Left	Bimodal
CI_04	F	34.7	10	Advanced Bionics	Right	Bimodal
CI_05	F	57.8	7	Cochlear Corporation	Right	Bilateral
CI_06	F	63.6	13	Cochlear Corporation	Right	Bilateral
CI_07	M	54.7	12	Cochlear Corporation	Right	Bilateral
CI_08	M	63.4	5	Cochlear Corporation	Left	Bimodal
CI_09	M	29.6	4	Cochlear Corporation	Right	Bilateral
CI_10	M	66.8	4	R-Cochlear Corporation, L-Advanced Bionics	Right	Bilateral
CI_11	F	74.5	5	Cochlear Corporation	Right	Bimodal
CI_12	M	58.7	17	Cochlear Corporation	Right	Bilateral
CI_13	F	66.7	4	Cochlear Corporation	Right <sup>*</sup>	Bilateral
CI_14	F	48.3	6	Cochlear Corporation	Left	Bilateral
CI_15	F	59.1	20	Cochlear Corporation	Left	Bilateral
CI_16	F	33.1	11	Cochlear Corporation	Left	Unilateral
CI_17	M	70.8	6	Cochlear Corporation	Right	Bimodal
CI_18	M	71.0	2	Cochlear Corporation	Right	Bimodal

\* Participant used later implanted device due to self-reported better speech recognition and preference of that ear.

**Table 2.**

Confusion matrix showing mean accuracy computed across all participants (left column), across younger (middle column) and across older (right column) participants. Data obtained by normally hearing participants are shown in the top four rows and data from CI users are shown in the bottom row. Accuracy is best along the diagonal.

Condition	Overall												Younger						Older																
	Emotion			Perceived			Ang			Sad			Neu			Hap			Emotion			Perceived			Ang			Sad			Neu				
	Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang					
<b>Full Spectrum</b>	Happy	76.81	8.92	12.45	0.78	1.03	Happy	79.89	8.86	10.12	0.33	0.79	Happy	67.99	9.09	19.13	2.08	1.70	Scared	5.69	71.86	7.21	11.47	3.77	Scared	10.98	53.41	15.34	15.91	4.36					
	Neutral	2.11	0.98	89.56	6.42	0.93	Neutral	2.38	0.79	91.53	4.56	0.73	Neutral	1.33	1.52	83.90	11.74	1.52	Sad	0.83	6.47	11.27	81.13	0.29	Sad	0.95	4.17	19.32	74.81	0.76					
	Sad	4.46	3.53	9.75	5.29	76.96	Angry	3.97	3.64	7.54	4.03	80.82	Angry	5.87	3.22	16.10	8.90	65.91	Angry	4.46	3.53	9.75	5.29	76.96	Angry	5.87	3.22	16.10	8.90	65.91					
	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang
	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang
<b>16 Channel</b>	Happy	54.17	19.27	17.71	2.34	6.51	Happy	66.07	17.33	12.24	0.93	3.44	Happy	45.63	18.13	23.13	3.75	9.38	Scared	20.44	34.64	21.74	14.45	8.72	Scared	21.25	23.13	29.58	16.46	9.58					
	Neutral	7.55	6.25	68.75	14.97	2.47	Neutral	6.02	5.36	72.49	12.37	3.77	Neutral	9.58	6.46	64.38	16.67	2.92	Sad	4.95	8.98	32.16	50.52	3.39	Sad	6.46	6.46	39.79	43.75	3.54					
	Sad	11.33	13.02	21.48	12.89	41.28	Angry	6.68	12.96	17.46	11.77	51.12	Angry	13.96	10.00	26.25	15.21	34.58	Angry	11.33	13.02	21.48	12.89	41.28	Angry	13.96	10.00	26.25	15.21	34.58					
	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang
	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang
<b>8 Channel</b>	Happy	43.22	21.51	18.02	5.47	11.77	Happy	46.74	21.88	16.47	3.71	11.20	Happy	32.95	20.45	22.54	10.61	13.45	Scared	18.80	27.81	24.56	13.08	15.75	Scared	15.91	20.64	34.47	10.80	18.18					
	Neutral	14.05	6.44	63.03	12.98	3.49	Neutral	14.65	6.45	63.02	13.02	2.86	Neutral	12.31	6.44	63.07	12.88	5.30	Sad	6.98	12.55	26.55	49.95	3.97	Sad	9.09	13.26	37.31	35.80	4.55					
	Sad	12.60	13.08	18.41	12.16	43.75	Angry	12.96	12.89	16.73	11.46	45.96	Angry	11.55	13.64	23.30	14.20	37.31	Angry	12.60	13.08	18.41	12.16	43.75	Angry	11.55	13.64	23.30	14.20	37.31					
	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang
	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang	Perceived	Portrayed	Hap	Sca	Neu	Sad	Ang
<b>4 Channel</b>	Happy	33.88	27.18	18.35	4.81	15.77	Happy	37.37	28.52	15.17	4.30	14.65	Happy	22.71	22.92	28.54	6.46	19.38	Scared	18.35	30.06	22.72	13.10	15.77	Scared	16.67	23.33	30.21	12.92	16.88					
	Neutral	18.35	30.06	22.72	13.10	15.77	Neutral	18.88	32.16	20.38	13.15	15.43	Neutral	16.67	23.33	30.21	12.92	16.88	Angry	18.35	30.06	22.72	13.10	15.77	Angry	16.67	23.33	30.21	12.92	16.88					
	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang	Emotion	Portrayed	Hap	Sca	Neu	Sad	Ang



Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Emotion	Perceived			Emotion			Perceived			Emotion			Perceived			Emotion			Perceived				
	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang
	Neutral	13.84	8.13	60.27	14.14	3.62	Neutral	14.00	8.20	60.03	14.00	3.78	Neutral	13.33	7.92	61.04	14.58	3.13	Neutral	10.00	11.67	35.00	38.75
Sad	9.03	11.66	26.59	48.07	4.66	Sad	8.72	11.65	23.96	50.98	4.69	Sad	10.00	11.67	35.00	38.75	4.58	Sad	10.00	11.67	35.00	38.75	4.58
Angry	12.00	14.58	16.47	15.03	41.91	Angry	12.24	15.17	14.58	14.32	43.68	Angry	11.25	12.71	22.50	17.29	36.25	Angry	11.25	12.71	22.50	17.29	36.25
Emotion	Perceived			Emotion			Perceived			Emotion			Perceived			Emotion			Perceived				
Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang	Portrayed	Hap	Sca	Neu	Sad	Ang
Happy	43.75	21.30	17.48	3.94	13.54	Happy	47.40	22.66	12.76	4.17	13.02	Happy	40.83	20.21	21.25	3.75	13.96	Happy	40.83	20.21	21.25	3.75	13.96
Scared	17.01	32.41	23.15	13.19	14.24	Scared	14.58	38.28	20.31	11.46	15.36	Scared	18.96	27.71	25.42	14.58	13.33	Scared	18.96	27.71	25.42	14.58	13.33
Neutral	10.30	5.21	61.81	14.93	7.75	Neutral	7.55	5.21	64.58	12.50	10.16	Neutral	12.50	5.21	59.58	16.88	5.83	Neutral	12.50	5.21	59.58	16.88	5.83
Sad	4.75	8.68	26.74	53.47	6.37	Sad	2.86	11.72	24.48	55.47	5.47	Sad	6.25	6.25	28.54	51.88	7.08	Sad	6.25	6.25	28.54	51.88	7.08
Angry	11.92	10.30	19.33	14.47	43.98	Angry	8.59	10.42	18.75	11.98	50.26	Angry	14.58	10.21	19.79	16.46	38.96	Angry	14.58	10.21	19.79	16.46	38.96

Results of linear mixed effects modeling on  $d'$  and reaction times comparing group differences when listening with unprocessed stimuli. In the interest of clarity, only significant effects are shown.

**Table 3.**

	$d'$				<u>Reaction Time</u>					
	$\beta$	SE	df	t	p	$\beta$	SE	df	t	p
Age	-0.028	0.005	61	-4.90	<0.001	0.021	0.005	70	4.21	<0.001
Male Talker (ref: Female)	-	-	-	-	-	-	-	-	-	-
Hearing Status: CI (ref: Normally Hearing)	-1.71	0.202	61	-8.50	<0.001	2.49	0.601	61	4.14	<0.001
Age x Hearing Status: CI (ref: Normally Hearing)	-	-	-	-	-	-0.023	0.010	60	-2.21	0.03

Results of Linear Mixed Effects model results on  $d'$  and reaction times for normally hearing adults listening to unprocessed and spectrally degraded speech. In the interest of clarity, only significant effects are shown.

**Table 4.**

	$d'$				<u>Reaction Time</u>					
	$\beta$	SE	df	t	p	$\beta$	SE	df	t	p
Age	-0.023	0.004	94	-6.55	<0.001	0.022	0.003	46	7.47	<0.001
Male Talker (ref: Female Talker)	-0.378	0.144	296	-2.63	<0.01	-	-	-	-	-
Degree of Spectral Degradation (ref: Full Spectrum)										
16 Channel	-1.68	0.188	296	-9.00	<0.001	-	-	-	-	-
8 Channel	-2.51	0.184	296	-13.59	<0.001	-	-	-	-	-
4 Channel	-2.76	0.186	296	-14.85	<0.001	-	-	-	-	-
<b>Overall Performance</b>										
Age x 16 Channel (ref: Full Spectrum)	0.008	0.004	297	2.52	0.01	-	-	-	-	-
Age x 8 channel (ref: Full Spectrum)	0.014	0.003	295	4.01	<0.001	-	-	-	-	-
Age x 4 Channel (ref: Full Spectrum)	0.015	0.004	296	4.22	<0.001	-	-	-	-	-
Male Talker x 16 Channel (ref: Female Talker, Full Spectrum)	-0.567	0.115	295	-4.92	<0.001	0.22	0.073	161	3.06	<0.01
Male Talker x 8 Channel (ref: Female Talker, Full Spectrum)	-0.329	0.114	295	-2.89	<0.01	-	-	-	-	-

**Table 5.**

Results of linear mixed effects modeling on correct and incorrect trials of reaction times. Group differences between listeners with NH and CIs with unprocessed (Full Spectrum) stimuli are shown on top. Results from NH listeners in the different conditions of spectral degradation are shown in the lower part. In the interest of clarity, only significant effects are shown.

	Correct Trials				Incorrect Trials			
	$\beta$	SE	t	p	$\beta$	SE	t	p
<b>Group Differences</b>								
Age	0.017	0.004	4.16	<0.001	0.022	0.006	3.45	<0.001
Male Talker (ref: Female)	-	-	-	-	-	-	-	-
Hearing Status: CI (ref: Normally Hearing)	1.97	0.46	4.24	<0.001	3.19	0.743	4.29	<0.001
Age x Hearing Status: CI	-0.018	0.008	-2.29	<0.05	-0.037	0.013	-2.94	<0.01
Age	0.018	0.004	5.44	<0.001	0.027	0.006	4.12	<0.001
Male Talker (ref: Female)	-	-	-	-	-	-	-	-
Degree of Spectral Degradation (ref: Full Spectrum)								
16 Channel	-	-	-	-	-	-	-	-
8 Channel	-	-	-	-	-	-	-	-
4 Channel	-	-	-	-	-	-	-	-
Age x 8 channel (ref: Full Spectrum)	0.006	0.003	2.36	<0.05	-	-	-	-
Male Talker x 8 Channel (ref: Female Talker, Full Spectrum)	0.259	0.082	3.17	<0.01	-	-	-	-