Research Article

How Vocal Emotions Produced by Children With Cochlear Implants Are Perceived by Their Hearing Peers

Sara A. Damm,[a] Jenni L. Sis,[a,b] Aditya M. Kulkarni,[a] and Monita Chatterjee[a]

Purpose: Cochlear implants (CIs) transmit a degraded version of the acoustic input to the listener. This impacts the perception of harmonic pitch, resulting in deficits in the perception of voice features critical to speech prosody. Such deficits may relate to changes in how children with CIs (CCIs) learn to produce vocal emotions. The purpose of this study was to investigate happy and sad emotional speech productions by school-age CCIs, compared to productions by children with normal hearing (NH), postlingually deaf adults with CIs, and adults with NH.
Method: All individuals recorded the same emotion-neutral sentences in a happy manner and a sad manner. These recordings were then used as stimuli in an emotion recognition task performed by child and adult listeners with NH. Their performance was taken as a measure of how well the 4 groups of talkers communicated the 2 emotions.

Results: Results showed high variability in the identifiability of emotions produced by CCIs, relative to other groups. Some CCIs produced highly identifiable emotions, while others showed deficits. The postlingually deaf adults with CIs produced highly identifiable emotions and relatively small intersubject variability. Age at implantation was found to be a significant predictor of performance by CCIs. In addition, the NH listeners' age predicted how well they could identify the emotions produced by CCIs. Thus, older NH child listeners were better able to identify the CCIs' intended emotions than younger NH child listeners. In contrast to the deficits in their emotion productions, CCIs produced highly intelligible words in the sentences carrying the emotions.
Conclusions: These results confirm previous findings showing deficits in CCIs' productions of prosodic cues and indicate that early auditory experience plays an important role in vocal emotion productions by individuals with CIs.

C ochlear implants (CIs) allow prelingually deaf children to have access to speech communication, develop receptive and expressive spoken language skills, attend mainstream schools, and experience a broader range of social interactions than was possible before pediatric cochlear implantation with modern-day devices. Although CIs transmit sufficient speech information to support a reasonable level of oral communication in day-to-day interaction, prosodic cues transmitted by these devices are not adequate to support the full range of speech communication enjoyed by listeners with normal

hearing (NH). In particular, voice pitch and its changes are poorly represented by CI processing. Unlike the normal auditory system, which provides a rich representation of the harmonic structure associated with voice pitch, CIs primarily represent voice pitch information in the form of temporal envelope periodicity cues. A large body of work has demonstrated that representing periodicity cues in this way does not support adequate perception of voice pitch or musical intervals and speech intonation cues important for the perception of vocal emotions, question/statement contrasts, or lexical tones (Burns & Viemeister, 1981; Chatterjee & Peng, 2008; Hopyan-Misakyan, Gordon, Dennis, & Papsin, 2009; Chatterjee et al., 2015; Deroche, Kulkarni, Christensen, Limb, & Chatterjee, 2016; Deroche, Lu, Limb, Lin, & Chatterjee, 2014; Green, Faulkner, Rosen, & Macherey, 2005; Luo, Fu, & Galvin, 2007; Tinnemore, Zion, Kulkarni, & Chatterjee, 2018; Xu & Pfingst, 2003).

For young children, access to prosodic cues is crucial for multiple reasons. Infants demonstrate greater attention to prosodically enriched speech (infant-direct speech), and it is known that this enhanced attention to exaggerated

[a]Auditory Prostheses and Perception Laboratory, Center for Hearing Research, Boys Town National Research Hospital, Omaha, NE
[b]Department of Special Education and Communication Disorders, University of Nebraska–Lincoln Barkley Memorial Center

Correspondence to Monita Chatterjee: monita.chatterjee@boystown.org

Sara A. Damm is now at Omaha Public Schools, NE.

prosody may help them to acquire their native language (Fernald, 1985; Fernald & Kuhl, 1987; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989). Lack of access to salient prosodic cues early in development may result in linguistic deficits in children with CIs (CCIs). A second aspect of such deficits may manifest itself in social interactions. As children develop into socially interactive adults, the ability to identify and interpret emotional prosody is likely important in the development of social cognition skills, the degree and quality of social interactions, and overall quality of life. "Theory of mind" refers to cognizance of others' beliefs and internal state through interactions. To know how others are feeling (affective theory of mind), we must be able to glean information about their emotions and internal state from their behavior and communication. A deficit in the ability to process such information (e.g., in listening through the degradation of a CI) may result in a theory of mind deficit. However, direct causative links between emotion recognition and theory of mind are yet to be drawn in CCIs. There is evidence of links between emotion understanding in children and theory of mind, a key aspect of social cognition (Grazzani, Ornaghi, Conte, Pepe, & Caprin, 2018; Mier et al., 2010), and young CCIs have been found to show delays in advanced aspects of theory of mind development (Wiefferink, Rieffe, Ketelaar, De Raeve, & Frijns, 2013). The link between emotion recognition and theory of mind is not well established in general. As an example, individuals with high risk for psychosis were found to have impaired theory of mind and poorer facial emotion perception than healthy controls, but once age and intelligence were controlled, the group difference in facial emotion perception was no longer significant (Barbato et al., 2015). On the other hand, social communication involves emotion perception and production, and it seems reasonable to postulate that impaired social communication would be related to diminished peer networks and poorer quality of life. Indeed, self-perceived quality of life has been found to be correlated to vocal emotion recognition, but not to speech perception, by both CCIs and adults with CIs (ACIs; Luo, Kern, & Pulling, 2018; Schorr, Roth, & Fox, 2009).

Deficits in the perception of emotional prosody may translate to altered production of vocal emotions in the pediatric CI population. If children are unable to monitor the pitch of their voices adequately, the ability to produce vocal changes in the graded and nuanced manner that underlies much of our vocal emotion communication may be impaired. Additionally, if children do not have full access to the primary cues for emotion in the voices of adults, siblings, and other communication partners around them, their ability to learn to produce specific emotions in specific ways (e.g., the knowledge that when people are happy they generally speak with higher mean voice pitch, that they modulate their voice pitch more, and that these features of speech are associated with higher intensity and a faster speaking rate) may be limited. As a result, CCIs may show deficits in their productions of vocal emotions.

Only a handful of studies have investigated the production of speech intonation and emotions by CCIs. These have focused on the imitation of exemplars rather than intentional production of emotions. Peng and colleagues (Peng, Tomblin, Spencer, & Hurtig, 2007; Peng, Tomblin, & Turner, 2008) found significant deficits in the ability of CCIs to imitate question/statement contrasts, compared to their NH peers. Nakata, Trehub, and Kanda (2012) found deficits in the ability of Japanese CCIs to imitate specific emotions (disappointment, surprise); they also found deficits in these children's emotion perceptions, compared to NH children. The perception and production data were significantly correlated. Chin, Bergeson, and Phan (2012) also found deficits in the ability of CCIs to imitate prosodic information in sentences. More recently, Wang, Trehub, Volkova, and van Lieshout (2013) studied 5- to 7-year-old bilaterally implanted children and found significant deficits in their imitation of happy and sad emotions relative to an NH comparison group. Furthermore, they conducted acoustic analyses of the productions and found that the CCIs produced smaller contrasts in voice pitch modulation to differentiate the emotions than their NH peers. In a group of school-age CCIs aged 6–12 years, Van de Velde et al. (2019) found a correlation between emotion recognition and production by CCIs. In speakers of tonal languages, moderate correlations have been reported between the perception and production of lexical tones (Deroche, Lu, Lin, Chatterjee, & Peng, 2019; Xu et al., 2011; Zhou, Huang, Chen, & Xu, 2013). Lexical tones are contrasted primarily by voice pitch and its changes, with duration cues playing a secondary role. In contrast, vocal emotions are acoustically more complex, with a number of secondary cues such as intensity, speaking rate, and spectral timbre playing key roles in addition to the pitch contour. Many of these secondary cues are well represented in electric hearing with a CI. Thus, compared to lexical tones, there are more opportunities for CCIs to communicate vocal emotions if their pitch production is compromised by the degraded input.

Adult users of CIs who became postlingually deaf (ACIs) present an important comparison to pediatric, prelingually deaf CI users. Unlike CCIs who are learning/have acquired language through the degraded input of the CI, the ACIs learned to hear and speak through a normal (or reasonably well-functioning) auditory system. Thus, while both populations are limited by the same device constraints, they bring very different brains to the task of speech communication. A comparison of productions by ACIs and CCIs would provide critical information about the perception–production link, particularly how it might be modified by the role of early auditory experience with acoustic versus electric hearing.

The CCIs and ACIs differ in yet another aspect. The CCIs were implanted within a more sensitive period and may be better able to take advantage of the greater plasticity of the brain than ACIs, who were implanted in adulthood, often in middle age or in older age, when the brain is less able to cope with a degraded input. On the other

hand, CCIs are developing language, social, and general cognition with a degraded input through the acoustic modality. To some extent, this may be a disadvantage, as suggested by demonstrated deficits in verbal working memory and sequential processing of inputs by CCIs (AuBuchon, Pisoni, & Kronenberger, 2015; Conway, Pisoni, Anaya, Karpicke, & Henning, 2011; Nittrouer, Caldwell-Tarr, & Lowenstein, 2013). In contrast, the ACIs have well-established cognitive resources, patterns of language, and concepts of emotion by the time they receive their CIs and can fall back on these resources to reconstruct the intended message from the degraded input. A key element of this study is a comparison of NH adult listeners' identification of happy and sad emotions produced by CCIs and ACIs.

As discussed above, previous studies of emotion productions by CCIs have used perceptions of the productions by adult listeners with NH as the primary measure of how well the emotions were produced. However, children communicate with each other for large portions of their days—on the playground, at school, and at home, with siblings. Little is known about how well CCIs' emotion productions are understood by their peers with NH. The literature suggests the young children with NH (CNHs) generally struggle more to understand speech that has been degraded in some way—listening to CI simulations, in background noise, or in reverberant rooms (Eisenberg, Shannon, Martinez, & Wygonski, 2000; Leibold, 2017; Valente, Plevinsky, Franco, Heinrichs-Graham, & Lewis, 2012). If the CCIs' productions of prosodic cues are considered degraded versions of speech prosody, then young CNHs may also struggle to decipher CCI peers' intended emotions. Information about this aspect of CCIs' speech communication could shed light on their everyday communication challenges.

In this study, we investigated the identification of vocal emotion productions by CCIs by their peers with NH and by adults with NH (ANHs). We asked school-age CCIs who were implanted by the age of 2 years and who had no usable pre-CI auditory experience to read simple sentences in a happy manner and a sad manner, with no training, no examples, and no feedback. We recorded the productions and, in a second protocol, played them to listeners with NH (child and adult), asking them to indicate whether each utterance sounded happy or sad in a single alternative, two-interval forced-choice procedure. Similarly, productions were recorded by CNHs, ANHs, and postlingually deaf ACIs. The productions by children were heard by both adult and child listeners with NH. The productions by the adults were heard by adult listeners with NH.

We selected only children who had been implanted by the age of 2 years and who had no usable hearing at birth for two reasons. First, we hoped that this selection would reduce sources of variability in the data that might stem from factors such as the effect of prolonged sensory deprivation that were not the focus of this study. Second, we had observed in preliminary explorations that the presence of hearing at birth has a positive impact on prosodic

productions in CCIs, and we had too small of a sample size to investigate this effect.

We selected happy and sad emotions because these are highly contrastive acoustically and because young children know these concepts early on (Pons, Harris, & de Rosnay, 2004) in development. We chose simple sentences that were easy to read, and we noted that all talkers articulated the words within the sentences correctly. To verify this objectively among the CCIs, we asked a separate group of NH adult listeners to perform a speech perception task in which they heard the productions by the CCIs and repeated them back without regard to the emotions.

The CI users in our sample included unilaterally and bilaterally implanted patients. We elected to record their productions with the earlier implanted device only. This choice had two purposes: (a) to retain consistency with our other studies of perception and (b) to constrain an additional source of variability that might arise from differences in asymmetry between the ears and variability in sequential bilateral implantation. Large differences in age at implantation between the ears in bilateral patients are likely to result in both peripheral neural degeneration and alteration in neural pathways on that side, resulting in informational differences along the ascending auditory pathways on the two sides. In children, this asymmetry would likely interact with development, complicating the picture. In contrast, simultaneous bilateral implantation early in infancy is likely to result in more similar auditory encoding on the two sides. However, in all cases of bilateral implantation, differences in insertion depth between the two ears are likely to contribute to asymmetric perception and variable binaural fusion. All of these factors may impact the perception of speech prosody in ways that are not well understood as yet. Constraining participants to listen on the earlier implanted side also allowed us to investigate the effects of age at implantation in a consistent way.

We expected that CCIs would show deficits in communicating the happy and sad emotions in their productions. We also expected high intersubject variability. We were interested in predictors of this variability and hypothesized that age at implantation and duration of experience with the device might play a role. In NH children, we expected developmental effects. In addition, we also expected "listener" effects—specifically that, among child listeners, younger NH children might struggle more to identify the intended emotions of CCIs than older NH children.

To summarize, this study was different from the majority of previous studies in a few specific ways. First, the productions were not imitations of an exemplar. While they were not spontaneous, they did require the child to associate their concept of the emotion with the production on their own. In this way, the productions were likely elicited in a manner that resembles real-world communication more closely than productions elicited via imitation. Second, the listeners included both ANHs and CNHs. Third,

we included comparison groups of NH children, NH adults, and postlingually deaf adult CI users as talkers. Fourth, rather than obtain subjective ratings of the productions, we asked the NH listeners to perform an emotion identification task in an objective procedure. Fifth, we attempted to reduce the heterogeneity of the results by only recording the CI users with their earlier implanted device active. Finally, we were careful to only include CCIs who had been implanted by the age of 2 years and had no hearing at birth. The outcome measure was the accuracy with which child and adult listeners with NH could identify the emotions produced by the different groups of talkers. Statistical analyses focused on comparisons of these results with those obtained in CCIs with peers with NH, ANHs, and postlingually deaf ACIs. In addition, statistical analyses also investigated predictors of performance by the CCIs, focusing on their age at implantation and duration of experience with the device.

Throughout this article, we refer to "performance" by talkers in communicating the intended emotion. We note here that the measure of performance is based on how well the talker's emotion is identified by listeners (children and adults) with NH. This is an indirect measure, but it may be argued that it also provides us with a real-life indicator of success in emotional communication.

Method

Participants were native speakers of American English who had no diagnosis of cognitive or linguistic impairments. All NH participants were screened at a level of 20 dB HL for NH from 250 to 8000 Hz. Participants were recruited using Boys Town National Research Hospital volunteer database and tested at Boys Town National Research Hospital in Omaha, NE. Written informed consent was obtained from all adult participants prior to participation. For child participants, written informed assent was obtained alongside parental permission to participate. Participants were compensated for travel and for their participation time. In addition, child participants were offered a toy or a book of their choice as a token of appreciation. Protocols were approved by the Boys Town National Research Hospital Institutional Review Board (Protocol 11-24-XP).

Productions

Participants

Child Talkers

Two groups of CCI talkers participated (they were recorded in successive years). Group 1 CCI talkers included seven CCIs (two boys, five girls; age range: 7.0–18.14 years, $M_{age} = 11.74$, $Mdn_{age} = 11.89$, $SD = 4.27$, mean duration of CI use = 10.34 years). Group 2 CCI talkers included six CCIs (two boys, four girls; age range: 7.9–18.49 years, $M_{age} = 13.50$, $Mdn_{age} = 14.025$, $SD = 4.07$, mean duration of CI use = 13.02 years). All CCIs were prelingually deaf

and implanted by the age of 2 years. They had no usable acoustic hearing, did not use hearing aids, and were primarily oral communicators. These two groups of talkers were not selected with different criteria; the only difference between them was that they participated in their study over two different summers and their productions were listened to by different groups of child and adult listeners (described below). Nine child talkers with NH (four boys, five girls; age range: 6.56–18.1 years, $M_{age} = 12.5$, $Mdn_{age} = 12.86$, $SD = 4.37$) were also recruited.

Adult Talkers

Adult talkers included nine NH adults (three men, six women; age range: 21–45 years) and 10 postlingually deafened adult CI users (four men, six women; age range: 27–75 years).

Information About Individual CCI and ACI Talkers

Table 1 shows relevant information about the CI users who participated. All information was obtained from a case history form filled out by participants and by the parents of child participants at the time of obtaining informed consent, immediately prior to participation.

Recordings

Each talker read a list of 20 simple emotion-neutral sentences (e.g., "This is it," "She is back," "It's my turn") in a happy manner and the same 20 sentences in a sad manner, for a total of 40 sentences. Table 2 lists the sentences. Emotion-neutral sentences were chosen to avoid any semantic bias during emotion production and during the task. No training, modeling, or feedback was given to the talker, with the exception of positive encouragement throughout the recording session. Each participant recorded the 20 sentences with a happy emotion three times, followed by the same sentences spoken with a sad emotion three times. Their productions were recorded while seated in a sound booth using a microphone (AKG C 2000 B) placed 12 in. from their mouth. Adobe Audition v.3.0 or v.6.0 (sample rate: 44100, 16-bit resolution) was used to record the productions. The second production of each recorded sentence was used for the remainder of the study unless it included nonspeech artifacts, in which case one of the other two productions was used. The recordings were high-pass filtered (75-Hz cutoff frequency) prior to analyses. CI users who were bilaterally implanted were recorded wearing only their earlier implanted device, and adult CI users with any residual hearing in the other ear were recorded with that ear plugged. Recall that the CCIs had no usable acoustic hearing.

Listening Tasks

The recordings were used as stimuli in two listening tasks. In Task 1 (Emotion Recognition), NH children and adults heard the recordings and indicated whether they were associated with happy or sad emotions. In Task 2

Table 1. Information about participants with cochlear implant in this study.

CCI participant	Age of testing (years)	Age of implantation	Duration of CI use	Bilateral implant (yes/no)	Gender	Manufacturer/device	Pre-/postlingual deafness
Group 1 CCI							
CICH02	18.14	2	16.14	No	Male	Cochlear	Prelingual
CICH03	11.89	1.4	10.49	No	Female	Advanced Bionics	Prelingual
CICH13	7.72	0.83	6.89	Yes	Female	Advanced Bionics	Prelingual
CICH18	17.2	1.7	15.5	No	Female	Advanced Bionics	Prelingual
CICH19	7	0.9	6.1	No	Female	Advanced Bionics	Prelingual
CICH20	7.6	1.1	6.5	Yes	Male	Advanced Bionics	Prelingual
CICH22	12.62	1.86	10.76	Yes	Female	Advanced Bionics	Prelingual
Group 2 CCI							
CICH35	12.73	1	11.73	Yes	Male	Advanced Bionics	Prelingual
CICH36	16.27	1.5	14.77	Yes	Female	Med-El	Prelingual
CICH37	18.49	1.5	16.99	Yes	Female	Advanced Bionics	Prelingual
CICH38	7.9	1.25	6.65	Yes	Female	Cochlear	Prelingual
CICH39	16.61	1.17	15.44	Yes	Female	Advanced Bionics	Prelingual
CICH40	14.025	1.5	12.53	Yes	Male	Advanced Bionics	Prelingual

ACI Participant	Age at testing (years)	Age at implantation (years)	Duration of CI use	Bilateral implant (yes/no)	Gender	Manufacturer/device
Postlingually deaf adult CI users						
C01	37	31	6	Yes	Female	Advanced Bionics
C03	67	55	12	No	Male	Advanced Bionics
C05	68	63	5	No	Female	Advanced Bionics
C06	75	55	20	No	Female	Advanced Bionics
C07	68	67	1	No	Female	Advanced Bionics
N5	53	50	3	No	Female	Cochlear
N6	51	44	7	Yes	Male	Cochlear
N7	57	51	6	No	Female	Cochlear
N15	61	59	2	No	Male	Cochlear
N16	27	25	2	Yes	Male	Cochlear

Note. CCI = children with cochlear implant; CI = cochlear implant.

(Speech Intelligibility), a different group of NH adults heard the recordings made by talkers in Group 1 CCIs one at a time and repeated back the sentences without regard to the emotions. Care was taken to ensure that these listeners had not participated in previous vocal emotion recognition tasks in our laboratory.

Table 2. List of sentences.

1. Time to go.
2. Here we are.
3. This is it.
4. This is mine.
5. The bus is here.
6. It's my turn.
7. They are here.
8. Today is the day.
9. Time for a bath.
10. She is back.
11. It's snowing again.
12. It's Halloween.
13. Time for bed.
14. Time for lunch.
15. I see a dog.
16. I see a car.
17. I see a cat.
18. That is the book.
19. I saw a bug.
20. That is a big tree.

Participants: Task 1 (Identification of Talkers' Intended Emotions)

Independent groups of listeners with NH listened to the productions by the different groups of talkers. These are as follows:

Group 1 CCI productions: Twenty-one CNH listeners (nine boys, 12 girls; age range: 6.84–18.49 years, M_{age} = 13.33, SD = 3.45) and six NH adult listeners (three men, three women; age range: 20.54–21.90 years) listened to Group 1 CCIs' productions in Task 1.

Group 2 CCI productions: A different set of 23 CNH listeners (14 boys, nine girls; age range: 6.88–18.90 years, M_{age} = 11.38, SD = 3.34) and seven NH adult listeners (four men, three women; age range: 19.36–29.53 years) listened to Group 2 CCIs' productions in Task 1.

CNH productions: 11 NH children (five girls, seven boys; age range: 6.5–16.7 years) and five NH adult listeners (three women, two men; age range: 22.17–25.75 years) listened to CNHs' productions in Task 1.

ACI productions: Six NH adults (three men, three women; age range: 20.2–21.9 years) listened to ACIs' productions in Task 1.

ANH productions: A different group of six NH adults (two men, four women; age range: 21.12–31.24 years) listened to ANHs' productions in Task 1.

Procedure: Task 1

Listeners were seated approximately 1 m in front of a loudspeaker in a sound booth. A 1-kHz calibration tone was adjusted to 65 dB SPL at the participants' left ear. The tone was generated to have a level corresponding to the mean root-mean-square level calculated across all 40 productions by each individual talker. Thus, intensity cues for the stimuli were not altered. There was no training prior to testing and no feedback given throughout the test. Each utterance was presented twice in randomized order. The experiments were controlled using a custom-built software program. The word indicating each emotion ("happy," "sad"), along with a corresponding black and white cartoon face, was placed in a white box located vertically on the right side of the computer screen. The rest of the screen had a colorful image that changed every few trials. Participants indicated the perceived emotion by using the computer mouse to click on the appropriate box. Percent correct scores were recorded. For each repetition, each talker's 40 utterances were presented in a single block and in randomized order within the block. Each listener heard each talker's productions, but the talker order was randomized for each listener.

Participants: Task 2 (Speech Intelligibility)

A group of 11 NH adults (two men, nine women; age range: 19.33–25.01 years) who had not participated in Task 1 were recruited to participate in Task 2.

Procedure: Task 2

Listeners were seated in the same position in the sound booth. They were presented with 20 utterances in total, selected from the happy/sad productions by Group 1 CCI. Each listener heard one of a number of test lists created by including four sentences spoken by each of five talkers. The test list of 20 sentences that each listener heard was created as follows. The 20 sentences listed in Table 2 were divided into five groups of four each. Each test set consisted of productions by five of the seven CCI talkers, each contributing one group of four sentences. Different listeners heard different test sets. These were designed in such a way that each listener heard productions by five CCIs, and each CCI's productions were heard by five listeners. The utterances were randomly selected to be happy or sad versions of the sentences. These test sets of 20 sentences were presented at a mean level of 65 dB to the listener. No visual display was presented to the listener during the task. After hearing each sentence, they repeated back the sentence without

regard to the emotion. The total words correct were recorded by the experimenter. For each talker, the number of total words correct was divided by the total possible number of correct words and converted to total percent words correct. Time restrictions necessitated the use of only a subset of the CCI talkers' recordings for this task. Group 1 CCI was selected because they spanned a wider range of age at implantation and a wider range of performance in emotion production.

Statistical Analyses

Prior to analyses, outliers were identified using "Tukey fences" (Tukey, 1977): Data points that fell above or below the upper and lower fences (third quartile + 1.5 * interquartile range or first quartile − 1.5 * interquartile range) were deemed outliers and excluded from analyses. Statistical analyses were conducted using R v. 3.6.0 (R Core Team, 2016). Specifically, the *lme4* and *lmerTest* packages (Bates, Maechler, Bolker, & Walker, 2015; Kuznetsova, Brockhoff, & Christensen, 2017) were used to construct linear mixed-effects models and the *anova* function within the *car* package (Fox & Weisberg, 2011) was used to compare models. Model residuals (histograms and qq plots) were visually inspected for normality. Conditional and marginal coefficients of determination for the models were obtained using the *r.squaredGLMM* function in the *MuMIn* package (Bartoń, 2019). The graphical package *ggplot2* (Wickham, 2016) was used to render plots.

Results

Adult Listeners' Recognition of Happy/Sad Emotions Produced by the Five Talker Groups

First, the mean scores obtained by the NH adult listeners were calculated for each talker. Figure 1 shows boxplots of these mean scores plotted against the four talker groups (ANH, CNH, ACI, and CCI). The actual data are overlaid on the boxplots. To remind the reader that two different groups of NH adult listeners heard the earlier tested and later tested CCIs, respectively, the data corresponding to each group are shown in different colors. The data set were severely nonnormal because of ceiling effects. Therefore, a nonparametric test, the Kruskal–Wallis rank sum test, was conducted to assess group-level differences in the scores. The result showed a significant difference between groups ($\chi^2 = 11.13$, $df = 3$, $p = .011$). A follow-up Dunn's test (with Holm correction) comparing the groups showed the following:

1. a marginal effect of poorer scores obtained with the ACIs' productions than the ANH group's productions (z statistic = −2.238, $p = .063$);

2. a significant effect of better scores obtained with the ANH group's productions than the CCIs' productions (z statistic = 3.052, $p = .0068$); and

3. a marginal effect of poorer scores obtained with the CCIs' productions than the CNHs' productions (z statistic = −2.099, p = .072).

Comparison Between ACI and CCI Groups

A particular focus of interest in this study is the comparison between the identifiability of emotions produced by ACI and CCI groups. The high variability of the CCI group in comparison to the ACI group (see Figure 1) is of note. It appears that a number of CCIs are high performers in emotion productions, while a number fall well below all other groups combined. The high variability in the CCI group also makes across-group comparisons less meaningful. To address our specific question regarding ACIs and CCIs, we divided their scores (from Figure 1) into those above the mean (better performers) and those below the mean (poorer performers) of each group. For the ACI group, the mean score was 96.52% correct; for the CCI group, the mean score was 89.43% correct. The Kruskal–Wallis rank sum test showed a significant difference between the poorer performers in the ACI and CCI groups (χ^2 = 7.5, df = 1, p = .0062) and no significant difference between the better performers in the two groups.

Figure 1. Boxplots of the percent correct scores obtained by normal hearing adult listeners attending to the productions by the four talker groups (abscissa). A mean percent correct score was calculated for each talker by averaging scores across all listeners for that talker. The actual mean scores are overlaid on the boxplots as dots. The two sets of mutually offset dots overlaid on the boxplot for children with cochlear implants' (CCI) data correspond to scores from the two listener groups attending to the two subsets of CCIs (Group 1 CCI and Group 2 CCI) who were tested at different times. ACI = adults with cochlear implants; ANH = adults with normal hearing; CNH = children with normal hearing.
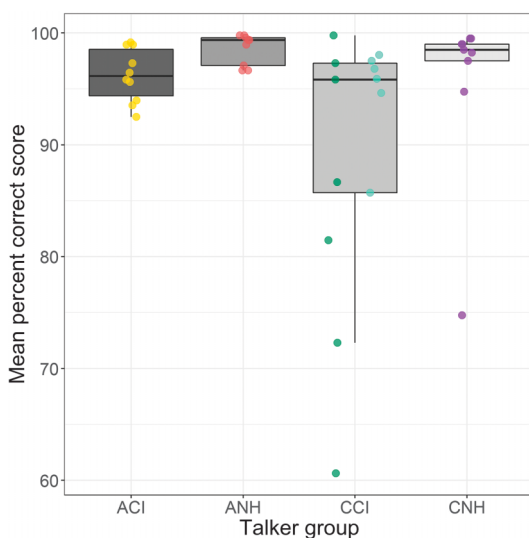
Adult Listeners' Recognition of Happy/Sad Emotions Produced by CNHs and CCIs: Predictors of Performance

Child Talkers With NH

Outlier analysis resulted in removal of low recognition scores obtained with productions by the youngest talker (6.6 years of age), comprising 8.88% of the total data set. An lme analysis conducted on the remaining data showed no significant effects of talker age. The results were generally limited by ceiling effects, but findings were similar with both original percent correct scores and rau (rationalized arcsine unit)–transformed versions.

Child Talkers With CIs

The CCI talkers' age at implantation and duration of device experience at the time of testing were the primary predictors considered in analyses of identification scores for their emotional productions. Exploratory analyses showed a significant correlation between these variables (see Table 3). Consistent with the evolution of clinical practice in cochlear implantation over the last two decades, the age at implantation was positively and significantly correlated with the year of implantation, that is, children implanted in previous years were also implanted at a later age. Additionally, we noted that the age at which the child was recorded in our laboratory (their chronological age, described as "age at testing" in Table 3) was significantly correlated with the child's age at implantation (older children were implanted earlier in time than younger children and had older ages at implantation than younger children did). We only considered the age at implantation and the duration of experience with the device as predictors in the present analyses.

Scores obtained by NH adult listeners attending to productions by individual child talkers with CIs are shown in Figures 2A and 2B (percent correct scores and rau scores, respectively), plotted against the CCIs' age at implantation. In the figures, each color represents a different talker, and the larger symbols show the mean scores computed across all listeners who attended to an individual talker. For these mean scores, the different shapes indicate the two groups of CCIs whose recordings were heard by different listeners. The line shows a simple regression through these mean scores.

The rau data were analyzed in a linear mixed-effects model including age at implantation as a fixed effect and listener-based random intercepts (nested within group).

Table 3. Correlations between children with cochlear implants' age at implantation, duration of device experience, year at implantation, and age at testing (all correlations were statistically significant).

Variable	Age at implantation	Age at testing	Duration of device experience
Age at testing	.688		
Duration of device experience	.639	.998	
Year of implantation	−.740	−.935	−.924

Figure 2. The effect of age at implantation of the child talkers with cochlear implant on the recognition of their emotions by adult listeners with normal hearing. The ordinates in A and B show percent correct scores and rau scores, respectively. Each talker is represented in a different color. The different symbols in the same color show the scores obtained by the different normal hearing listeners attending to each talker's productions. The larger symbols show the mean score computed across all listeners for each talker. For these larger symbols, the different shapes indicate the two groups of talkers (and listeners). The regression line represents a simple regression through the mean scores.
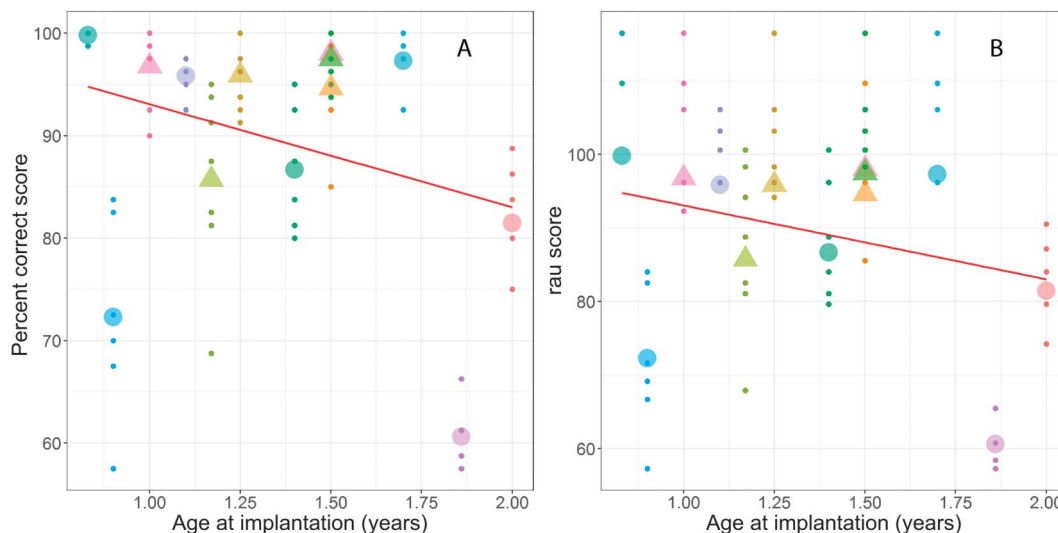
Results showed a significant effect of age at implantation, $t(82.31) = -2.544$, $p = .0128$. The model showed an estimated intercept of 113.414 rau units and an estimated slope of -12.65 rau ($SE = 4.97$) units per year of increase in the age at implantation. The addition of duration of device experience as a fixed effect did not result in a significant improvement to model fit. Model residuals showed a reasonable approximation to a normal distribution. The estimated marginal and conditional R^2 values for the model were .065 and .178, respectively, indicating a small effect size. A parallel analysis with the percent correct score (untransformed) as the dependent variable showed similar results, but the model residuals showed more obvious deviation from normality.

To summarize, these analyses showed that increasing age at implantation was linked with a significant decrease in recognition of the emotions conveyed by CCIs, at a rate estimated at -12.65 rau (decrease in performance) per year of delay in age at implantation. The results suggest no significant mitigating effect of duration of device experience.

Child Listeners' Recognition of Happy/Sad Emotions Produced by CNHs and CCIs

Outlier analysis conducted on the rau scores obtained by NH child listeners when listening to CNH talkers resulted in the removal of 5.05% of the data (all were scores corresponding to productions by the same 6.6-year-old CNH talker discussed in the previous section, heard by five of the child listeners). Note that not all of the data based on productions by this youngest talker were outliers. No

outliers were observed in the data obtained by NH child listeners attending to CCI talkers, possibly because of the broader distribution of scores obtained with CCI talkers.

Outlier analysis conducted on the percent correct scores obtained by the NH child listeners attending to CNH talkers resulted in the exclusion of a larger number of data points (10.1% of the data). No outliers were observed in the data obtained by the NH child listeners attending to CCI talkers.

Excluding outliers, mean percent correct scores obtained by each group of NH child listeners attending to the two groups of child talkers (CCI, CNH) are shown in boxplots in Figure 3. As in Figure 1, the means were computed across all listeners for each talker. The patterns are consistent with those observed with NH adult listeners. The Kruskal–Wallis rank sum test was conducted on the mean rau scores to assess group-level differences. The result showed a significant difference between groups ($\chi^2 = 57.337$, $df = 1$, $p < .0001$). The Kruskal–Wallis rank sum test conducted on the mean percent correct scores confirmed these findings ($\chi^2 = 10.926$, $df = 1$, $p = .001$).

Child Listeners' Recognition of Happy/Sad Emotions Produced by CNHs and CCIs: Predictors of Performance

Child Talkers With NH

The percent correct scores in CNHs were not normally distributed, and an attempt at linear mixed-effects modeling resulted in a skewed distribution of residuals. A linear

Figure 3. Boxplots of the mean percent correct scores obtained by normal hearing child listeners attending to the productions by child talkers with normal hearing (CNHs) and cochlear implants (CCIs). The mean scores were computed as the average score obtained across all listeners for each talker's productions. As in Figure 1, the data are overlaid on the boxplots. The two sets of data points overlaid on the CCIs' boxplot indicate the scores obtained by the two groups of listeners attending to the two groups of CCI talkers who were tested at different times.
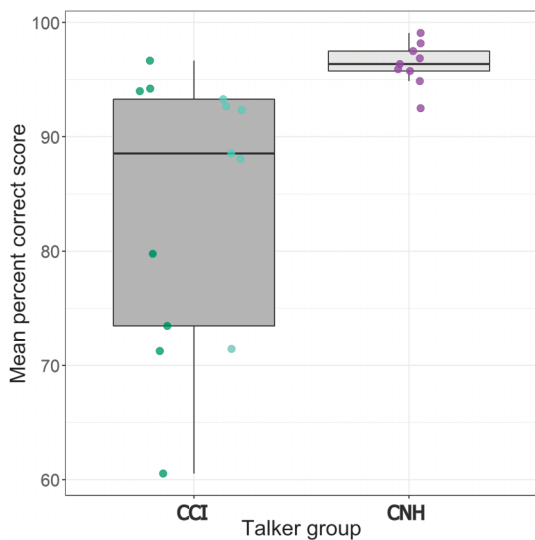
mixed-effects model conducted on the NH children's rau scores with CNH talkers' productions to examine the effect of listener and talker ages resulted in a reasonable distribution of residuals. The results showed a weak effect of talker age, $t(83.73) = 1.731$, $p = .087$ (estimated slope = 0.389, $SE = 0.225$), that did not reach significance and no effect of listener age.

Child Talkers With CIs

A linear mixed-effects analysis was conducted on the CCI talkers' emotion productions as identified by the child listeners with NH. The percent correct scores showed a reasonably normal distribution, so a linear mixed-effects model was first attempted with percent correct score as the dependent variable. The talkers' age at implantation was included as a fixed effect, with listener-based random intercepts (nested within the listener group). The inclusion of listeners' age improved the model significantly. The model residuals showed an acceptable distribution, with a slight deviation from normality. Results showed significant effects of talkers' age at implantation, $t(244.89) = -5.543$, $p < .0001$, and listeners' age, $t(42.077) = 3.457$, $p = .0013$. The model estimated an intercept of 88.7% correct and a slope of −11.95% correct per year of age at implantation. Scores improved slightly with listeners' age, at a rate of 0.979% correct per year.

The distribution of the rau-transformed scores was closer to a normal distribution, so the same model was applied with rau score as the dependent variable. Results

showed significant effects of talkers' age at implantation, $t(244.67) = -5.529$, $p < .0001$, and listeners' age, $t(41.95) = 3.264$, $p = .0022$. The model summary estimated an intercept of 94.46 rau units and a slope of −15.44 rau units ($SE = 2.793$) per year of increase in the age at implantation of the CCI talker. The model summary estimated an improvement of 1.23 rau units ($SE = 0.377$) per year of increase in the child listener's age. Including duration of device experience as a fixed effect did not improve the model fit. Model residuals confirmed an approximately normal distribution. The conditional and marginal estimated R^2 values for the model were .125 and .268, respectively, suggesting a small effect.

Figures 4A and 4B show the percent correct scores and rau scores obtained with the CCI talkers and with the listeners with NH. Upper and lower panels show the same scores plotted against the talkers' age at implantation and listeners' age, respectively. As in Figure 2, in the upper panels, different colors show data obtained with different talkers' productions. The larger symbols show the mean scores computed across all listeners attending to each talker (different shapes indicate the two groups of CCI talkers), and the line shows a simple regression through the mean scores. In the lower panels, each symbol represents a different talker, and the lighter lines show individual regressions through each talker's data. The larger symbols show the mean scores obtained by each listener across all talkers, and the thick line shows a simple regression through these mean scores.

Speech Intelligibility of Child Talkers With CIs

The productions of the child talkers with CIs (Group 1) were heard by 11 ANHs who were asked to repeat back the sentences in Task 2. The scores obtained by the NH adult listeners showed ceiling levels of intelligibility, with a mean of 96.69% correct across talkers ($SD = 0.032$).

Discussion

Summary of Findings

The CCIs in this study produced less recognizable emotions than their NH peers. The difference between CCI and CNH talkers was highly significant for child listeners with NH and only marginal for adult listeners with NH (possibly due to adult listeners' better ability to compensate for deficits in the productions). However, large variability was observed in the data, with a number of CCIs' intended emotions being highly identifiable to both NH child and adult listeners and others much less so.

The ACIs' productions were marginally less identifiable than the ANHs' productions, not significantly different from the CNHs' productions, and overall not significantly different from the CCIs' productions. The high variability across CCI participants made it difficult to draw firm conclusions from across-group comparisons. When the ACI and CCI talkers were divided into better and poorer

Figure 4. The effect of the child talkers with cochlear implants' age at implantation (upper) and the age of the child listeners with normal hearing (lower) on the latter group's ability to identify emotions produced by children with cochlear implants. The ordinates in A and B show percent correct scores and rau scores, respectively. In the upper panels, the abscissa represents the age at implantation. In the lower panels, the abscissa represents the listeners' age. As in Figure 2, in the upper panels, each talker is represented by a different color, and different data points in the same color show scores obtained for that talker by the different listeners. The larger symbols represent the mean score computed across all the talkers. For these larger symbols, the different shapes indicate the two groups of talkers (and listeners). The line shows a simple regression through the mean scores. In the lower panels, each talker is represented by a different (small) symbol, and the pale lines show simple regressions through scores for each talker. The larger symbols show the mean scores obtained by each listener across
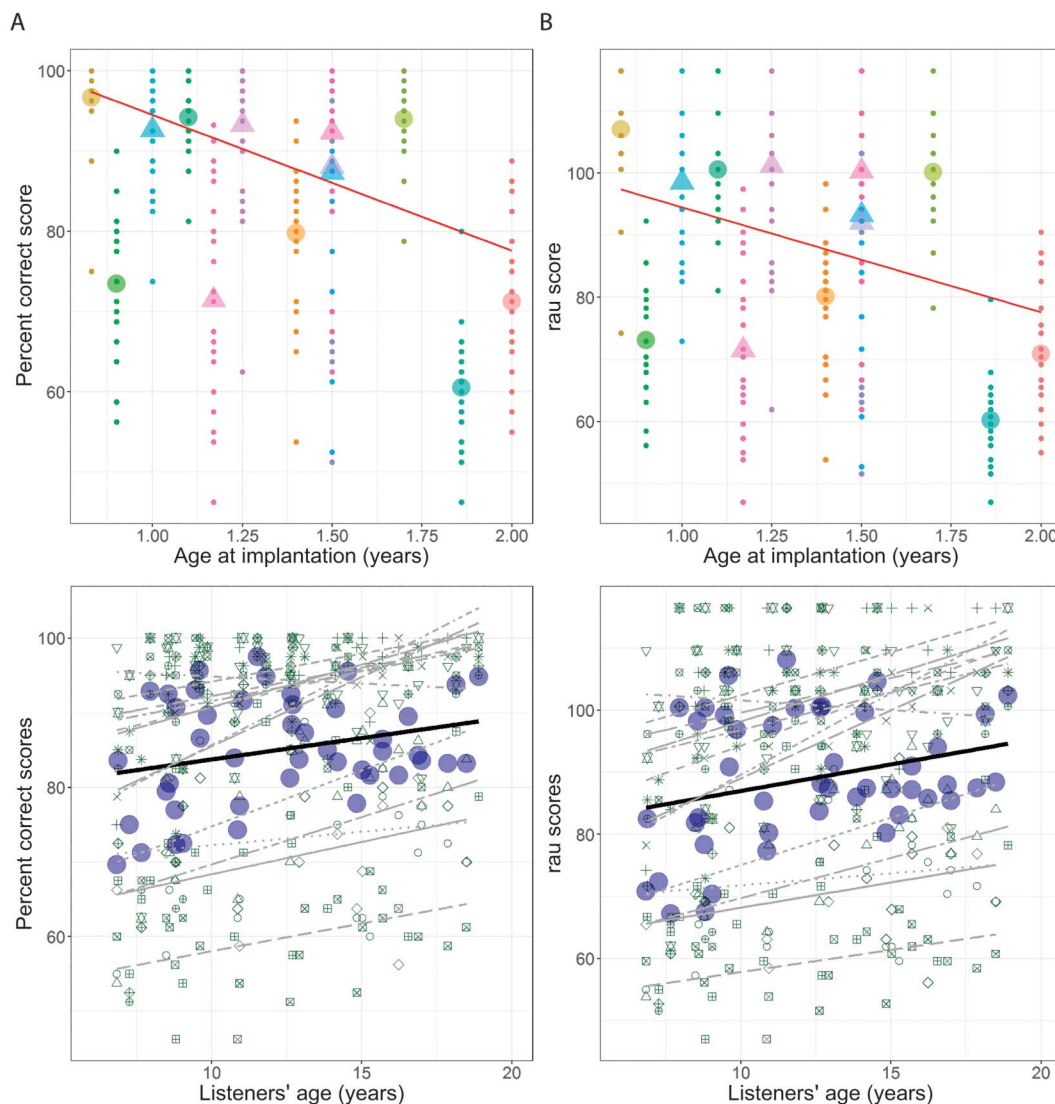
performers, we observed that the poorer performing CCIs' productions were significantly more poorly identified than the poorer performing ACIs' scores, while the better performers in the two groups were not significantly different from one another.

The contributions of CCIs' age at testing and duration of device experience to the large variability in NH listeners' identification of the CCIs' productions were investigated. The ANHs' recognition of the CCIs' productions was significantly predicted by the CCI talkers' age at

implantation. Similarly, the CNH listeners' scores were significantly predicted by the CCI talkers' age at implantation. An additional significant predictor was the CNH listeners' age, with younger CNHs showing poorer identification than older CNHs.

Finally, the results of Task 2 showed that, although the CCI talkers had significant deficits in their emotion productions, their ability to produce the words within the sentences was at ceiling, with an intelligibility level averaging 96.69% correct.

Talker-Related Factors: Early Auditory Experience

Two aspects of the present results point to the importance of early auditory experience in CI users' production of emotional prosody. The first is the difference in the patterns of postlingually deaf ACIs' productions and those of the prelingually deaf CCIs. The high variability of the CCIs' productions compared to the ACIs' productions is of note. It is apparent from visual inspection of the data points in Figures 1 and 3 that the high-performing CCIs in our sample were comparable to the ACI, ANH, and CNH groups, while the emotional productions of poorer performing CCIs were much harder to identify than the other groups. This was supported by the statistical comparison of better and poorer performers in the ACI and CCI groups, which showed no significant difference between the better performers but significantly higher scores among the poorer ACI performers than the poorer CCI performers.

The second aspect is the significant effect of the CCIs' age at implantation, despite the fact that all of the CCIs had been implanted by the age of 2 years. This finding underscores the importance of early electric experience in prelingually deaf CCIs. This is remarkable, given that electric hearing does not provide listeners with salient vocal pitch cues and the spectrotemporal degradation results in deficits in many areas involving the perception of suprasegmental speech information, as discussed in the introduction. We infer that earlier implanted children benefit in the production of prosodic cues from the greater plasticity of the brain at the time of device activation and in the months thereafter. Earlier age of implantation has been shown to benefit CCIs in various aspects of receptive and expressive language (Niparko et al., 2010). It is unknown how the motor systems that engage in the communication of emotional prosody develop in CCIs and how this interacts with the critical period. This finding is additionally remarkable in the light of our recent studies showing little to no effect of the age at implantation in children when the task involves the perception of vocal emotions or complex pitch (Chatterjee et al., 2015; Deroche et al., 2016, 2014). Emotion communication involves the use of multiple acoustic cues that covary. It is possible that production patterns that maximally utilize these different cues are the most successful in communicating emotions. Recognition of emotions requires the listener to map the multidimensional sensory input onto perceptual space. While this internal representation is likely utilized in the development of speech motor patterns, it is possible that this mapping process involves greater distortions in CCIs than in those learning to speak through high-fidelity acoustic hearing. Thus, while such a distorted map/internal representation may help CCIs to learn to correctly identify and discriminate vocal emotions, it may not be their best guide as they learn to produce the same emotions (particularly in such a way as to be understood by NH conversational partners). For example, a recent study by Deroche et al. (2019) found that CCIs' productions of lexical tones focused on contrasting the secondary cue (duration), while CNHs' productions focused on voice pitch

changes. Additionally, they found that CCIs' emphasis on the voice pitch cue in their productions was inversely related to their reliance on the duration cue in the perception domain: CCIs who relied more on the duration cue in perception, emphasized the voice pitch cue less in production.

We note here that albeit significant, the effects reported here were small (as indicated by the low R^2 values), indicating that other factors must account for the large intersubject variability. The high variability is consistent with the literature on CI speech perception outcomes in general. Possible factors underlying this variability that were not taken into account in this study include perceptual sensitivity to the acoustic cues for emotion, general cognition, the individual's affective state, language abilities, socioeconomic status, and access to speech therapy.

It is possible that, compared with other factors, the role of age at implantation is even smaller in the perception domain, which may have made it difficult to capture in previous studies. A weak link between perception and production of emotional prosody has been reported by Van de Velde et al. (2019). This suggests that the factors predicting variability in perception and production of vocal emotions are either different, or similar but differently weighted. Further studies investigating the perception–production link in CI users are clearly warranted.

The Listener's Age Matters

One result of this study is the finding that younger NH children show greater deficits in recognizing CCI talkers' emotion productions than do older NH children. The greater ability of ANHs to cope with distorted inputs is also evident in their better identification of the CCIs' productions than child listeners with NH (evident from visual inspection of Figures 1 and 3). As discussed in the introduction, it is known that young children struggle more to identify degraded speech in general, whether the degradation is due to background noise or a CI simulation and whether the task is speech recognition or emotion recognition (Chatterjee et al., 2015; Eisenberg et al., 2000; Leibold, 2017; Tinnemore et al., 2018). To the extent that CCI talkers' productions of emotional prosody represent a form of degraded speech, the present finding is consistent with the broader literature. The implication is quite significant, suggesting that peer-to-peer communication is at risk in CCIs as they communicate with younger friends of siblings or peers at younger ages. Older child listeners and ANHs show improved performance in deciphering the intended emotion of CCI talkers, confirming that the ability to compensate and reconstruct the intended message improves with development.

Relation to Previous Studies

The present results both confirm and extend previous findings of deficits in prelingually deaf CCIs' productions of emotional prosody. The finding of deficits in the identifiability of emotions produced by CCIs relative to other

groups provides general confirmation of previous findings. The fact that these findings were obtained without asking the talkers to imitate a specific talker indicates that the emotion production deficits by pediatric CI users can be extended to more real-life scenarios than established in previous research. The present work also extends previous findings by showing that poorer performing prelingually deaf CCIs produce significantly less recognizable emotions than poorer performers among their postlingually deaf adult CI user counterparts, while better performing individuals in the two groups are not significantly different. This may suggest that early auditory experience with salient acoustic cues plays an important role in the production of speech prosody, perhaps more so in those individuals who may be at some disadvantage in how well they perceive sounds through electric hearing or how well their cognitive and linguistic systems support their adjustment to the CI. Thus, we speculate that poorer performers among the ACIs may have had impairments in a number of areas that led to their deficit, but that their superior performance relative to poorer performing CCIs was due to the fact that they learned to communicate with good acoustic hearing.

The present results suggest a key role for early experience with sounds, even presented through electric hearing. This finding of an advantage in earlier implanted children is broadly consistent with previous studies showing advantages in earlier implanted children in both expressive and receptive language outcomes. An important novel aspect of this study is the finding that the listener's age is a significant predictor of how well the emotion communications of CCIs are recognized and the implication that peer-to-peer communication by the pediatric CI population may be at higher risk than communication with adults.

Limitations of This Study

The primary limitations of this study were the relatively small sample size and high variability among the CCIs. A larger sample size would allow for the inclusion of factors such as socioeconomic status, general language and cognition skills, perceptual skills, and so forth in statistical models. The ceiling effects encountered in the data are another limitation and may be countered in future studies by the inclusion of emotions with more nuanced differences. Additionally, although the productions were somewhat more relevant to real-world scenarios than those elicited by imitations, the laboratory recording environment and the paradigm were still somewhat artificial. The development of methods that might elicit more natural emotive productions would greatly benefit research in this area in the future.

Clinical Implications: The Importance of Rehabilitative Efforts and Residual Low-Frequency Hearing

The two emotions selected for this study were chosen for their highly contrastive acoustic features and the relative ease of their recognition by younger children. Given the significant deficits observed in CCIs' productions of these emotions, it is likely that the deficits are even greater for the communication of more subtle differences in vocal emotions. Thus, the present results may be taken to underscore the importance of emphasizing prosodic communication in CCIs.

The excellent production of words in the sentences by the CCIs speaks not only to the success of CIs as a neural prosthesis but also to the success of audiological and speech therapy aspects of rehabilitation. The deficits in vocal emotion productions, however, suggest an urgent need for a comprehensive and focused effort targeting suprasegmental communication by CI users, particularly pediatric CI recipients who do not have the benefit of early acoustic experience. The present results, as well as the perception-related findings in the literature, also suggest that CCIs who have some residual acoustic hearing in the low frequencies may benefit from efforts aimed at hearing preservation or the use of a hearing aid in addition to the CI. Even if the low-frequency hearing is only preserved in the early developmental years, it may help CCIs achieve improved production of prosodic cues.

References

AuBuchon, A. M., Pisoni, D. B., & Kronenberger, W. G. (2015). Short-term and working memory impairments in early-implanted, long-term cochlear implant users are independent of audibility and speech production. *Ear and Hearing, 36*(6), 733–737.

Barbato, M., Liu, L., Cadenhead, K. S., Cannon, T. D., Cornblatt, B. A., McGlashan, T. H., . . . Addington, J. (2015). Theory of mind, emotion recognition and social perception in individuals at clinical high risk for psychosis: Findings from the NAPLS-2 cohort. *Schizophrenia Research: Cognition, 2*(3), 133–139.

Bartoń, K. (2019). *MuMIn: Multi-Model Inference. R Package* (Version 1.43.6). Retrieved from https://CRAN.R-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Burns, E. M., & Viemeister, N. F. (1981). Played-again SAM: Further observations on the pitch of amplitude-modulated noise. *The Journal of the Acoustical Society of America, 70,* 1655–1660.

Chatterjee, M., & Peng, S. C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hearing Research, 235*(1–2), 143–156.

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., . . . Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing Research, 322,* 151–162.

Chin, S. B., Bergeson, T. R., & Phan, J. (2012). Speech intelligibility and prosody production in children with cochlear implants. *Journal of Communication Disorders, 45*(5), 355–366.

Conway, C. M., Pisoni, D. B., Anaya, E. M., Karpicke, J., & Henning, S. C. (2011). Implicit sequence learning in deaf children with cochlear implants. *Developmental Science, 14*(1), 69–82.

Deroche, M. L. D., Kulkarni, A. M., Christensen, J. A., Limb, C. J., & Chatterjee, M. (2016). Deficits in the sensitivity to pitch sweeps by school-aged children wearing cochlear implants. *Frontiers in Neuroscience, 10,* 73. https://doi.org/10.3389/fnins.2016.00073

Deroche, M. L. D., Lu, H.-P., Limb, C. J., Lin, Y.-S., & Chatterjee, M. (2014). Deficits in the pitch sensitivity of cochlear-implanted children speaking English or Mandarin. *Frontiers in Neuroscience, 8,* 282. https://doi.org/10.3389/fnins.2014.00282

Deroche, M. L. D., Lu, H.-P., Lin, Y.-S., Chatterjee, M., & Peng, S.-C. (2019). Processing of acoustic information in lexical tone production and perception by pediatric cochlear implant recipients. *Frontiers in Neuroscience, 13,* 639. https://doi.org/10.3389/fnins.2019.00639

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., & Wygonski, J. (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America, 107,* 2704–2710.

Fernald, A. (1985). Four-month olds prefer to listen to motherese. *Infant Behavior & Development, 8,* 181–195.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior & Development, 10,* 279–293.

Fox, J., & Weisberg, S. (2011). *An {R} companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage. Retrieved from http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Grazzani, I., Ornaghi, V., Conte, E., Pepe, A., & Caprin, C. (2018). The relation between emotion understanding and theory of mind in children aged 3 to 8: The key role of language. *Frontiers in Psychology, 9*(724). https://doi.org/10.3389/fpsyg.2018.00724

Green, T., Faulkner, A., Rosen, S., & Macherey, O. (2005). Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification. *The Journal of the Acoustical Society of America, 118*(1), 375–385.

Hopyan-Misakyan, T. M., Gordon, K. A., Dennis, M., & Papsin, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychology, 15*(2), 136–146.

Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language, 16*(1), 55–68.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Leibold, L. J. (2017). Speech perception in complex acoustic environments: Developmental effects. *Journal of Speech, Language, and Hearing Research, 60*(10), 3001–3008.

Luo, X., Fu, Q. J., & Galvin, J. J., III. (2007). Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends in Amplification, 11*(4), 301–315.

Luo, X., Kern, A., & Pulling, K. R. (2018). Vocal emotion recognition performance predicts the quality of life in adult cochlear implant users. *The Journal of the Acoustical Society of America, 144,* EL429–EL435.

Mier, D., Lis, S., Neuthe, K., Sauer, C., Esslinger, C., Gallhofer, B., & Kirsch, P. (2010). The involvement of emotion recognition in affective theory of mind. *Psychophysiology, 47*(6), 1028–1039.

Nakata, T., Trehub, S. E., & Kanda, Y. (2012). Effect of cochlear implants on children's perception and production of speech prosody. *The Journal of the Acoustical Society of America, 131*(2), 1307–1314.

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N. Y., Quittner, A. L., . . . CDaCI Investigative Team. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association, 303*(15), 1498–1506.

Nittrouer, S., Caldwell-Tarr, A., & Lowenstein, J. H. (2013). Working memory in children with cochlear implants: Problems are in storage, not processing. *International Journal of Pediatric Otorhinolaryngology, 77*(11), 1886–1898.

Peng, S. C., Tomblin, J. B., Spencer, L. J., & Hurtig, R. R. (2007). Imitative production of rising speech intonation in pediatric cochlear implant recipients. *Journal of Speech, Language, and Hearing Research, 50*(5), 1210–1227.

Peng, S. C., Tomblin, J. B., & Turner, C. W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear and Hearing, 29*(3), 336–351.

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology, 1,* 127–152. https://doi.org/10.1080/17405620344000022

R Core Team. (2016). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Schorr, E. A., Roth, F. P., & Fox, N. A. (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.

Tinnemore, A. R., Zion, D. J., Kulkarni, A. M., & Chatterjee, M. (2018). Children's recognition of emotional prosody in spectrally-degraded speech is predicted by their age and cognitive status. *Ear and Hearing, 39*(5), 874–880.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Valente, D. L., Plevinsky, H. M., Franco, J. M., Heinrichs-Graham, E. C., & Lewis, D. E. (2012). Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children. *The Journal of the Acoustical Society of America, 131,* 232–246.

Van de Velde, D. J., Schiller, N. O., Levelt, C. C., Van Heuven, V. J., Beers, M., Briaire, J. J., & Frijns, J. H. (2019). Prosody perception and production by children with cochlear implants. *Journal of Child Language, 46*(1), 111–141.

Wang, D. J., Trehub, S. E., Volkova, A., & van Lieshout, P. (2013). Child implant users' imitation of happy- and sad-sounding speech. *Frontiers in Psychology, 4*(351). https://doi.org/10.3389/fpsyg.2013.00351

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* New York, NY: Springer-Verlag.

Wiefferink, C. H., Rieffe, C., Ketelaar, L., De Raeve, L., & Frijns, J. H. (2013). Emotion understanding in deaf children with a cochlear implant. *Journal of Deaf Studies and Deaf Education, 18*(2), 175–186.

Xu, L., Chen, X., Lu, H., Zhou, N., Wang, S., Liu, Q., . . . Han, D. (2011). Tone perception and production in pediatric cochlear implant users. *Acta Oto-Laryngologica, 131,* 395–398.

Xu, L., & Pfingst, B. E. (2003). Relative importance of temporal envelope and fine structure in lexical-tone perception. *The Journal of the Acoustical Society of America, 114,* 3024–3027.

Zhou, N., Huang, J., Chen, X., & Xu, L. (2013). Relationship between tone perception and production in prelingually-deafened children with cochlear implants. *Otology & Neurotology, 34*(3), 499–506.