Children with bilateral cochlear implants identify emotion in speech and music

Anna Volkova[1], Sandra E Trehub[1*], E Glenn Schellenberg[1], Blake C Papsin[2], Karen A Gordon[2]

[1]Department of Psychology, University of Toronto, Toronto, Canada, [2]Department of Otolaryngology, University of Toronto, Toronto, Canada

Objectives: This study examined the ability of prelingually deaf children with bilateral implants to identify emotion (i.e. happiness or sadness) in speech and music.
Methods: Participants in Experiment 1 were 14 prelingually deaf children from 5–7 years of age who had bilateral implants and 18 normally hearing children from 4–6 years of age. They judged whether linguistically neutral utterances produced by a man and woman sounded happy or sad. Participants in Experiment 2 were 14 bilateral implant users from 4–6 years of age and the same normally hearing children as in Experiment 1. They judged whether synthesized piano excerpts sounded happy or sad.
Results: Child implant users' accuracy of identifying happiness and sadness in speech was well above chance levels but significantly below the accuracy achieved by children with normal hearing. Similarly, their accuracy of identifying happiness and sadness in music was well above chance levels but significantly below that of children with normal hearing, who performed at ceiling. For the 12 implant users who participated in both experiments, performance on the speech task correlated significantly with performance on the music task and implant experience was correlated with performance on both tasks.
Discussion: Child implant users' accurate identification of emotion in speech exceeded performance in previous studies, which may be attributable to fewer response alternatives and the use of child-directed speech. Moreover, child implant users' successful identification of emotion in music indicates that the relevant cues are accessible at a relatively young age.

Keywords: Cochlear implants, Children, Emotion, Perception, Speech, Music

Introduction

Cochlear implants (CIs) have made auditory-verbal communication accessible to many children with profound hearing loss (e.g. Geers *et al.*, 2003). The benefits for congenitally deaf children seem greatest when they receive their implants early, say by 1–3 years of age (Geers, 2004; Nicholas and Geers, 2006; Waltzman and Cohen, 1998). Such early access to language facilitates social and emotional development. For example, deaf children with good language skills (oral or sign language) are reported to have fewer psychosocial difficulties than their peers with more limited skills (Dammeyer, 2010). For child CI users with hearing parents, exposure to a spoken language has been linked to greater social well-being (Percy-Smith *et al.*, 2008), perhaps because few hearing parents are sufficiently proficient in sign language to enable optimal parent–child communication.

A critical aspect of interpersonal communication in general and of parent–child communication in particular is the exchange of affective messages, both verbal and non-verbal. Aspects of visual communication of affect by facial expression, posture, and movement are universal, as are aspects of auditory communication by vocal and non-vocal means (Bachorowski, 1999; Balkwill and Thompson, 1999; Scherer, 2003; Zentner *et al.*, 2008). The ease with which children with CIs identify the emotional valence of vocal, non-verbal sounds has implications for their emotional well-being, as assessed by self-report (Schorr *et al.*, 2009). There is relatively little research, however, on the ability of CI users to identify the intended emotions in speech and music.

Portrayals of emotion in speech and music are guided both by universal aspects of emotional expression and by socio-cultural conventions (Bachorowski, 1999; Balkwill and Thompson, 1999; Scherer *et al.*, 2001). Well before children acquire the rudiments of language, they are exposed to some of these modes of expression. For example, mothers convey emotion in their speech to pre-verbal infants by means of exaggerated prosody (Fernald, 1991; Papoušek, 1992) and heightened affect (Trainor

*Correspondence to: Sandra E Trehub, Department of Psychology, University of Toronto, Mississauga, ON, Canada L5L 1C6.
Email: sandra.trehub@utoronto.ca

DOI 10.1179/1754762812Y.0000000004

et al., 2000), which confer a musical flavor to such speech (Fernald, 1989). From the early months, infants prefer infant-directed speech to adult-directed speech, the latter being more neutral in emotional tone (Cooper and Aslin, 1990; Fernald, 1985; Singh *et al.*, 2002). Mothers across cultures also sing expressively to their infants (Trehub *et al.*, 1993), at times for playful purposes and at times for soothing purposes (Trehub and Trainor, 1998). Infants prefer infant-directed singing to informal but non-infant-directed singing (Trainor, 1996), even in the newborn period (Masataka, 1999), which suggests that such expressive singing may have intrinsic appeal.

Congenitally deaf infants do not have access to these affective auditory stimuli in the early months of life because even with early diagnosis, they are unlikely to receive implants much before 12 months of age (Holt and Svirsky, 2008). Owing to delays in receiving expressive vocal input, one would expect delays in child CI user's ability to interpret and transmit vocal affect. Even in the post-implant period, the degraded pitch and spectral cues provided by implants (Geurts and Wouters, 2001; Loizou, 1998) interfere with the processing of vocal emotion. Variations in fundamental frequency (F0), or pitch, contribute to the differentiation of emotions in speech and music, although amplitude, tempo (rate), and rhythm are also important (Bachorowski, 1999, Laukka *et al.*, 2005; Scherer, 1986, 2003). Amplitude and timing variations are accessible to CI users (Loizou, 1998), but for normal-hearing (NH) individuals, these cues are typically used in conjunction with pitch variations to differentiate vocal emotions (Bachorowski, 1999; Scherer, 2003). The available literature indicates that CI users have considerable difficulty identifying emotion in speech (Hopyan-Misakyan *et al.*, 2009; Luo *et al.*, 2007; Most and Aviner, 2009). For example, prelingually deaf children from 7–13 years of age with unilateral implants are as accurate at identifying facial expressions (happy, sad, angry, fearful) as their hearing peers, but CI users perform at chance levels on affective speech prosody that is perceptible to NH children (Hopyan-Misakyan *et al.*, 2009).

Variations in pitch, amplitude, and timing characterize music as well as speech. Music and speech also share some cues to specific emotions (Juslin and Laukka, 2003). The adverse consequences of degraded pitch and spectral information for music processing have been well documented (see Kraus *et al.*, 2009; McDermott, 2004, for reviews). This research has focused largely on CI users' difficulty differentiating melodies (Galvin *et al.*, 2007; Vongpaisal *et al.*, 2006, 2009) and identifying familiar music on the basis of pitch cues alone (Cooper *et al.*, 2008; Cullington and Zeng, 2011; Hsiao, 2008; Kong *et al.*, 2004; Nimmons, 2007; Stordahl, 2002). Not surprisingly, this difficulty makes music unpalatable to many postlingually deafened CI users (Gfeller *et al.*, 2000; Lassaletta *et al.*, 2007; Leal *et al.*, 2003), but congenitally or prelingually deaf children with CIs typically enjoy music listening and music making (Gfeller *et al.*, 2000; Nakata *et al.*, 2006; van Besouw *et al.*, 2011; Vongpaisal *et al.*, 2006). To date, however, there is only one published study of emotion identification in music by child CI users (Hopyan *et al.*, 2011). Hopyan *et al.* (2011) found that prelingually deaf children from 7–13 years of age with unilateral implants distinguished happy from sad classical music excerpts but their accuracy was significantly below that of same-age NH children.

The goal of this investigation was to ascertain whether bilateral CI users from 4–7 years of age could identify happiness and sadness in speech and music. Previous research with CI users with 7–13 years of age indicated that they were unable to differentiate speech prosody expressing happiness, sadness, anger, and fear (Hopyan-Misakyan *et al.*, 2009), but they could distinguish happy from sad intentions in instrumental (piano) music (Hopyan *et al.*, 2011). Children's judgments of emotion in speech were examined in Experiment 1, and their judgments of emotion in music were tested in Experiment 2.

Experiment 1

In typical studies of emotion identification in speech among NH listeners, participants hear utterances with neutral content spoken in a manner that portrays specific emotions. They respond by selecting one of several emotions from a closed set (see Scherer, 2003, for a review). Luo *et al.* (2007) used such a task to present adult CI users with men's and women's utterances that conveyed happiness, sadness, anger, anxiety, and emotional neutrality. Amplitude cues were preserved in some cases and normalized in others. Even with amplitude cues preserved, CI users identified less than half of the emotions, performing substantially below the levels attained by NH adults. Most and Aviner (2009) compared children and adolescents with CIs (10–17 years of age) with same-age NH individuals and with deaf hearing aid users on their identification of happiness, sadness, fear, anger, surprise, and disgust in auditory, visual, and audio-visual contexts. In the auditory context, the performance of both deaf groups was significantly worse than that of hearing individuals. Although hearing individuals performed better in audio-visual than in visual contexts, the addition of auditory cues provided no advantage for CI and hearing aid users. In other words, auditory cues resolved the ambiguity of some facial expressions for hearing listeners but not for deaf listeners. CI users' difficulty differentiating the six vocal emotions is in line with 7- to 13-year olds' difficulty differentiating happy, sad, angry, and fearful vocal emotions (Hopyan-Misakyan *et al.*, 2009). In short, the available research seems to indicate that the pitch and spectral cues provided by CIs are insufficient for reliable differentiation of vocal emotions.

Unquestionably, NH individuals outperform CI users on the identification of vocal emotion, but their performance is far from perfect. In a number of studies, NH adults have identified 60–70% of the target emotions in tasks involving multiple emotions, with some emotions identified more readily than others (Luo *et al.*, 2007; Scherer, 2003). Moreover, NH individuals perform better with familiar talkers or with talkers who have well-modulated voices such as those exhibiting expressive variations in pitch, loudness, and timing (see Bachorowski, 1999, for a review). Luo *et al.* (2007) found that CI users often confused happiness with anger or fear, which have overlapping acoustic cues, and sadness with neutrality, which also have overlapping cues. Although NH listeners were much more accurate than CI listeners, both had similar confusion patterns. Furthermore, the removal of amplitude cues impaired the performance of NH listeners as well as CI users. Most and Aviner (2009) found similar confusion patterns in NH listeners and their counterparts with hearing loss. Happiness was identified the most accurately, followed by anger and disgust. It is clear, then, that vocal emotions are more difficult to discern than emotional facial expressions, even for listeners who have access to the full range of pitch and spectral cues.

Acoustic cues may be insufficient for accurate identification of discrete emotional categories in adult speech (e.g. Bachorowski, 1999). Such cues may be more useful for indicating the talker's arousal level, whether high or low, than for differentiating positive from negative valence (Bachorowski, 1999; Bänziger and Scherer, 2005). Indeed, NH listeners tend to confuse emotions associated with similar arousal levels. For example, they confuse happiness with anger, which share high arousal levels and the acoustic cues of high pitch, rapid speaking rate, and pitch and amplitude variability. They also confuse neutrality and sadness, which share low arousal levels and the acoustic cues of low pitch, slow speaking rate, as well as reduced pitch and amplitude variability. Perhaps CI users would be more successful at emotion identification if the emotions in question had contrastive arousal levels (high or low), such as happiness and sadness, and the cognitive demands of the task were reduced by having two response choices rather than four or five. They might also be more successful with child-directed speech, which is more affectively expressive than adult-directed speech. Aside from the fact that child-directed speech is more familiar and more engaging to young children, emotional intentions (e.g. approval, comfort, and prohibition) are more transparent in child-directed than in adult-directed speech because of its exaggerated prosody (Bryant and Barrett, 2007; Fernald, 1989).

Maternal prosody to infants and toddlers with CIs is tuned to their listening experience and linguistic capabilities as opposed to their chronological age (Bergeson, 2011; Kondaurova and Bergeson, 2011). There is no comparable research on maternal prosody to older children with CIs, but the available research indicates that other aspects of maternal speech to deaf children are responsive to the children's linguistic proficiency (DesJardin and Eisenberg, 2007; Nienhuys *et al.*, 1984). The result is that maternal speech to CI users tends to be similar to speech to younger NH children.

In this experiment, young bilateral CI users were required to identify child-directed speech samples as happy or sad. CI users from 5–7 years of age and a control group of NH children listened to natural, expressive utterances produced by a man and a woman. The man's speech had potential processing advantages for CI users because of its lower pitch range (Chatterjee and Peng, 2007; Vongpaisal *et al.*, 2010), but it was also possible that the woman would express emotion more distinctively (Luo *et al.*, 2007). Owing to the use of a child-directed speaking style, emotion categories that contrasted in arousal, and only two response alternatives, young CI users in the present experiment were expected to differentiate emotions more effectively than older CI users in previous studies (Luo *et al.*, 2007, Most and Aviner, 2009, Hopyan-Misakyan *et al.*, 2009). Nevertheless, because of child CI users' diminished access to pitch and spectral cues, they were expected to perform more poorly than their NH peers.

Method

Participants

The participants included 14 bilateral CI users (five girls and nine boys, $M = 5.8$ years, SD = 0.6, range: 5.1–7.0) from middle-class families who were recruited from a large metropolitan area (see Table 1). Four children had progressive hearing loss from birth and 10 children were congenitally or prelingually deaf. All participants used Nucleus 24 Contour implants with Freedom, SPrint, or CP810 processors (see Table 1) programmed with the Advanced Combinational Encoder processing strategy, and each had a minimum of 2.5 years of implant experience ($M =$ 4.3 years, SD = 0.8, range: 2.8−5.3). With the exception of children with progressive hearing loss, their first implant was activated between 9 and 27 months of age. When tested with their implants, absolute thresholds for tones within the speech range were

Table 1 Demographic variables of children with CIs

Participant	Gender	Age at test (years) Exp 1; 2	Age at CI activation Ear 1; 2	Processors (L/R when different)	Etiology
CI-1*	M	6.4; 6.4	3.4; 3.4	Freedom	Genetic
CI-2	M	5.2; 5.5	0.8; 1.7	CP810	Genetic
CI-3	M	5.4; 5.4	1.1; 1.1	Freedom	Genetic
CI-4	F	6.3; 6.3	1.0; 3.6	Freedom/SPrint	Genetic
CI-5*	F	7.0; 6.9	2.5; 4.0	Freedom	Unknown
CI-6	M	5.3; 5.3	1.0; 4.6	Freedom	Genetic
CI-7	M	5.1; 5.8	0.9; 1.8	Freedom	Genetic
CI-8	M	6.1; 6.3	0.8; 1.5	Freedom	Genetic
CI-9*	M	6.4; 6.4	3.1; 6.3	Freedom	Mondini dysplasia
CI-10	F	5.1; -	1.1; 1.1	Freedom	Genetic
CI-11	F	6.3; 6.9	1.0; 3.5	Freedom/SPrint	Unknown
CI-12	M	6.4; -	1.3; 2.3	Freedom	Genetic
CI-13	F	5.1; 5.1	1.1; 3.4	Freedom	Unknown
CI-14*	M	5.5; 5.5	2.7; 2.7	Freedom	Unknown
CI-15	F	5.8	1.7; 1.7	Freedom/CP810	Unknown
CI-16	F	4.1	1.1; 1.1	Freedom	Unknown

*Progressive hearing loss from birth.

within normal limits (10–30 dB HL). All children with CIs participated in auditory-verbal therapy for at least 2 years after implantation. They also communicated exclusively by auditory-verbal means and were in age-appropriate school classes with their NH peers. A comparison sample of NH children from middle-class families consisted of 18 preschoolers (12 girls and 6 boys, $M = 5.4$ years, SD = 0.5, range: 4.8–6.2) recruited from the local community whose mean age was slightly younger than the mean for the CI group, $t(30) = 2.05$, $P = 0.049$. No NH child had a personal or family history of hearing problems, and all children were free of colds on the day of testing.

Apparatus and stimuli

A man and a woman were instructed to produce happy and sad versions of the following three sentences: *A chair has four legs*, *Flowers grow in the garden*, and *The lamp is on the table*. The 'actors' were asked to speak naturally but in an expressive manner, as if talking to children. Both had considerable experience interacting with young children. Fundamental frequency (F0), amplitude, and duration of all utterances from the two talkers were calculated by means of PRAAT software (Boersma and Weenink, 2005). Table 2 provides information about mean F0, F0 range, amplitude, and duration for happy and sad utterances produced by the man and woman. F0 contours of the stimuli produced by the two talkers are illustrated in Fig. 1. Stimuli largely conformed to cultural conventions for the vocal portrayal of happiness and sadness, as described by Johnstone and Scherer (2000). Specifically, the man's and woman's sad utterances had a smaller F0 range and were lower in overall amplitude than their happy utterances. Although the man's sad utterances were longer in duration than his happy utterances, as expected, the woman's happy utterances were longer in duration than her sad utterances (Table 2). As a result, duration was not a reliable cue to the target emotion in the present set of stimuli.

The stimuli were recorded in a 3 m × 2.5 m double-walled, sound-attenuating chamber (Industrial Acoustics Corporation) with a microphone (Sony F-V30T, Tokyo, Japan) connected directly to a Windows XP computer workstation. High-quality digital sound files (44.1 kHz, 16-bit, mono) were created using a digital audio editor (Sound Forge 6.0). Four-colored digital photographs (headshots) of

Table 2 Happy- and sad-sounding speech stimuli

Emotion/Gender	Sentence	Mean F0 (Hz)	F0 range (Hz)	Mean Amplitude (dB)	Duration (s)
Happy/female	Chair	313.8 (SD = 83.2)	184.6–520.3	67.7 (SD = 9.5)	1.7
Happy/female	Flowers	289.2 (SD = 86.1)	169.8–519.2	67.0 (SD = 10.9)	1.9
Happy/female	Lamp	285.2 (SD = 99.2)	156.4–486.5	66.6 (SD = 11.5)	1.6
Happy/male	Chair	132.4 (SD = 28.5)	90.9–199.2	65.2 (SD = 11.6)	1.9
Happy/male	Flowers	156.2 (SD = 41.6)	95.9–231.9	67.3 (SD = 11.0)	1.6
Happy/male	Lamp	134.9 (SD = 27.6)	76.3–183.3	66.3 (SD = 10.7)	1.4
Sad/female	Chair	190.2 (SD = 32.4)	154.8–269.4	63.9 (SD = 10.8)	1.5
Sad/female	Flowers	176.2 (SD = 19.6)	149.3–232.9	63.9 (SD = 10.1)	1.5
Sad/female	Lamp	186.3 (SD = 18.3)	162.2–219.6	64.5 (SD = 11.2)	1.4
Sad/male	Chair	94.2 (SD = 6.8)	79.1–115.2	63.6 (SD = 10.1)	1.9
Sad/male	Flowers	96.7 (SD = 10.1)	81.9–127.3	64.1 (SD = 11.2)	1.8
Sad/male	Lamp	102.1 (SD = 9.7)	82.9–125.9	64.6 (SD = 11.2)	1.6

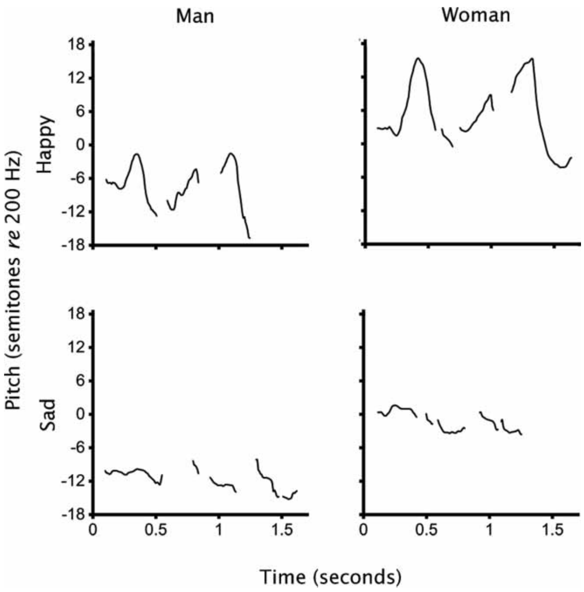

Figure 1 Pitch contours (displayed in semitones referenced to 200 Hz) of happy and sad versions of *The lamp is on the table* by male and female talker (Experiment 1).

the speakers were taken against a white background. One showed the man with a happy facial expression and another with a sad expression. The two other photographs showed happy and sad facial expressions featuring the woman (see Fig. 2).

Testing took place in a double-walled sound-attenuating booth, either at the university laboratory (Industrial Acoustics Corporation, 3 m × 2.5 m) or at a local children's hospital (4.3 m × 2.7 m). A computer workstation and amplifier (Harman/Kardon HK3380, Harman International Industries, Stanford, CT) located outside the university booth were connected with a 17-inch touch-screen monitor (Elo LCD TouchSystems, Tyco Electronics Corporation, Berwyn, PA) and two high-quality loudspeakers (Audiological GSI, Eden Prairie, MN) inside the booth. At the hospital venue, a GSI-61 two-channel clinical audiometer (Grason-Stadler Instruments, Eden Prairie, MN) replaced the amplifier. The loudspeakers were placed at 45° azimuth to the participant with the touch-screen monitor at the midpoint. An interactive computer program (customized for Windows XP) presented stimuli and recorded response selections when the participant touched the screen.

Figure 2 Response alternatives (photos of happy and sad facial expressions) for the female talker (Experiment 1).

The experimenter could record participants' responses using a portable keyboard connected to the workstation when young children preferred to point to their selection rather than touching the screen. All stimuli were played at a comfortable listening level of approximately 65 dB SPL.

Procedure

Participants were tested individually. A parent was present in the booth for a number of participants, especially the younger ones who wanted their parent nearby. The child was seated facing the computer monitor, with the experimenter off to one side and the parent out of the child's line of sight. Parents were instructed to refrain from interacting with their child, and all complied fully. The experimenter explained to the children that they were going to hear a man or a lady talk and that they should decide whether the talker sounded happy or sad. Participants were explicitly instructed to attend to *how* the talker sounded, not *what* he or she said. There were four practice trials to familiarize children with the task. Such practice trials were considered necessary because of young children's tendency to focus on the content of spoken messages rather than on the speaking style (Morton and Trehub, 2001). During practice trials, the experimenter answered the children's questions but provided no feedback about the accuracy of their responses. Although children were told that they could listen to any utterance more than once before responding (i.e. as often as they liked), no child asked for repetitions. When children were very hesitant after hearing an utterance, which occurred periodically for CI users, the experimenter provided a repetition, which usually elicited a prompt response.

Auditory stimuli were presented in two blocks, one for both talkers, with a short break between blocks. In both groups of children, the order of blocks (man–woman or woman–man) was balanced. Stimuli within each block consisted of three sentences in both emotions, with each sentence repeated three times and order randomized for a total of 18 trials in both blocks. As noted, the participant had the option of hearing a stimulus repeated before making a decision. At the beginning of each trial, colored photographic images of a man or woman (depending on the trial block) with a happy and sad facial expression were presented on the monitor. Before beginning each block of trials, the experimenter verified that the child could identify the facial expression in each image by asking 'Is the man/lady in this picture happy or sad?' The spatial arrangement of images (left/right) was identical for all participants across all trials. After listening to each stimulus, children responded by touching one of the images. No feedback was provided other than periodic encouragement and praise to maintain the children's enthusiasm and cooperation.

Results and Discussion

Preliminary analyses revealed that performance was not distributed normally for either the man talker, $P = 0.027$, or the woman talker, $P = 0.020$, or for performance averaged across talkers, $P = 0.034$ (Kolmogorov–Smirnoff tests). Although overall performance in the CI group was more variable than in the NH group, this difference fell short of statistical significance, $F(13, 17) = 3.37$, $P = 0.076$ (Levene's test). Nevertheless, subsequent analyses used nonparametric tests.

We first examined the number of children who performed significantly better than chance. With a binomial test (normal approximation, correcting for continuity, $P < 0.05$, one-tailed), overall performance required 24 correct responses on the 36 trials to exceed chance levels. For child CI users, 12 of 14 surpassed chance, and for NH children, 17 of 18 did so. In short, performance was remarkably good in both groups and consistent across individuals.

The performance of child CI users and NH children is shown in Fig. 3. A direct comparison of groups on overall performance confirmed that the NH children performed significantly better than the child CI users, $P = 0.047$ (Mann–Whitney U test). Poorer performance for child CI users than for NH children was evident for the woman talker ($P = 0.029$), but not for the man ($P = 0.107$) (Mann–Whitney U tests). Although NH children showed significant improvement from the first to the second block of trials ($P = 0.013$), comparable improvement was not evident among the CI children ($P = 0.135$) (Wilcoxon signed rank tests).

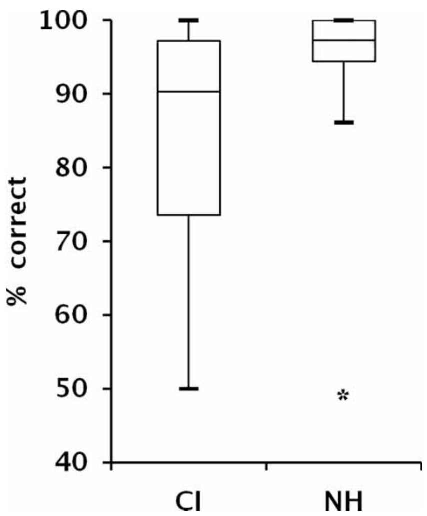

Figure 3 Box plots showing performance of child CI users and NH children on happy and sad speech (Experiment 1). Lower and upper borders of each box indicate first and third quartiles, respectively; line within each box indicates the median; and whiskers depict the highest and lowest scores, except for outliers. The single outlier in the NH group is indicated by an asterisk.

Performance of the child CI users collapsed across talkers was associated positively with duration of implant use (i.e. years since first implant activation), $r = 0.60$, $N = 14$, $P = 0.012$ (one-tailed), which is an impressive finding in view of the small sample. Fig. 4 depicts individual scores of children with CIs. Two of them exhibited error-free performance across talkers. An additional two had error-free performance for the woman talker only, and two others had error-free performance for the man talker only. Interestingly, only six NH children achieved error-free performance on both trial blocks, although their performance was generally more consistent than that of children with CIs.

CI users mistook sad-sounding speech as happy (76% of all errors) more often than they mistook happy-sounding speech as sad. This was especially noticeable for the poorest performing CI users. This pattern of confusions was also reported by Luo *et al.* (2007) and Most and Aviner (2009), which may reflect a response bias for positive emotions rather than greater ease of identifying happy-sounding speech. In any case, young CI users performed successfully on a two-alternative forced-choice task that required them to differentiate happy from sad speech, unlike older CI users who failed to identify happy and sad utterances in the context of a four-alternative forced-choice task (Hopyan-Misakyan *et al.*, 2009).

As noted, natural speaking rate and amplitude variations were preserved in the present stimuli and were available, in principle, as cues. In practice, however, differences in overall amplitude and amplitude variation were consistent across talkers, but differences in speaking rate across emotion categories were inconsistent across speakers (see Table 2). As a result, speaking

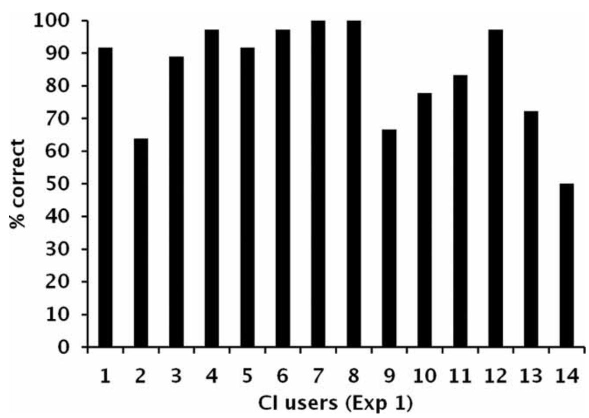

Figure 4 Performance of individual child CI users on happy and sad speech (Experiment 1).

rate, which usually distinguishes happy- from sad--sounding speech, was an unreliable cue for the female talker, which may help to explain group differences for the woman talker but not for the man. Differences in F0 range between happy and sad stimuli were large and consistent across talkers (Table 2). Differences in F0 contour also provided reliable cues to happy and sad speech but they may have been inaccessible to CI users (Geurts and Wouters, 2001; Loizou, 1998). When contour contrasts are substantial, however, as in statements versus questions, CI users perform above chance levels although well below the levels attained by NH listeners (Meister *et al.*, 2009; Most and Peled, 2007; Peng *et al.*, 2008). It is possible that the young CI users in the present experiment used a combination of acoustic cues to differentiate the happy from sad utterances although amplitude cues would have been the most prominent of these. In the absence of feedback, however, it is clear that CI users had reasonable representations of happy and sad vocal qualities on the basis of their everyday experience. Their ability to differentiate music excerpts expressing happy and sad emotions in conventional ways was assessed in Experiment 2.

Experiment 2

There is increasing interest in the non-musical and musical consequences of children's short- and long-term involvement in musical activities (Kirschner and Tomasello, 2010; Schellenberg, 2006). For example, listening to pleasant music has consequences for children's prosocial behavior (Kirschner and Tomasello, 2010) and their performance on a variety of tasks (Schellenberg and Hallam, 2005; Schellenberg *et al.*, 2007). Moreover, music lessons have been linked to long-term cognitive outcomes (Schellenberg, 2006, 2011).

In addition, there is long-standing interest in adults' and children's ability to understand emotion in music. Typical tasks in this realm require children to link musical excerpts to discrete emotional categories such as happiness, sadness, anger, and fear. Despite the ubiquity of this type of task, it seems less appropriate for describing emotional aspects of music than it does for speech or facial expressions (Trehub *et al.*, 2010). Such labels aptly describe the feelings associated with vocal or facial emotional expressions, but they are much less suitable for instrumental musical excerpts. Instead of discerning the emotion of the performer or composer or the emotional consequences on the listener, which may involve non-specific arousal or general feelings of pleasure (Salimpoor *et al.*, 2009), listeners are typically expected to discern the emotional intentions of the performance. To do so requires familiarity with culturally typical uses of emotional labels in relation to music.

Adults use some combination of tempo, loudness, pitch level, mode (major or minor), and consonance or dissonance to judge the emotional intentions of music excerpts, but tempo and mode have received the most empirical attention (see Hunter and Schellenberg, 2010, for a review). In general, western adults judge music in the major mode and with rapid tempo as happy and music in the minor mode and with slow tempo as sad (Peretz *et al.*, 1998). Unlike tempo, which often differentiates happy- from sad-sounding music across cultures, mode does not (e.g. Balkwill and Thompson, 1999). Despite the cross-cultural importance of tempo to musical emotions, it is still necessary to learn the cultural conventions involving tempo as well as mode. With excerpts from the classical repertoire, 4-year-old children fail to link musical mode, tempo, or their combination with happiness and sadness, 5-year-old children use tempo to differentiate happy from sad musical excerpts, and 6- to 8-year-old children link both cues, separately and in combination, to happiness and sadness (Dalla Bella *et al.*, 2001). When the stimuli consist of children's songs rather than classical music, 4-year-old children are able to use tempo to distinguish happy from sad excerpts (Mote, 2011). In general, however, listeners of all ages find it easier to differentiate high-arousal emotions (e.g. happiness, anger) in music from low-arousal emotions (e.g. sadness or peacefulness; Hunter *et al.*, 2011).

Given the musical pitch processing difficulties of CI users (Brockmeier *et al.*, 2011; Cooper *et al.*, 2008; McDermott, 2004), it is not surprising that they are unable to discriminate major from minor melodic patterns (Vongpaisal *et al.*, 2006). Although CI users can resolve timing differences and they use such differences in music-recognition tasks (Hsiao, 2008; Kong *et al.*, 2004; Stordahl, 2002), it is unclear when child CI users first link those and other acoustic cues with musical emotions. Hopyan *et al.* (2011) found that child CI users from it 7–13 years of age reliably differentiated happy from sad excerpts of classical music that had been used in emotion recognition research with NH adults and children (Dalla Bella *et al.*, 2001; Peretz *et al.*, 1998). Not surprisingly, the CI users performed more poorly than same age NH children. Although the stimuli had tempo, mode, and other cues to emotion, the authors speculated that CI users based their judgments primarily on tempo. When timing cues are available, adult CI users distinguish happy from sad musical excerpts, as do their NH counterparts (Brockmeier *et al.*, 2011).

According to Dalla Bella *et al.* (2001), children younger than 5 years of age are unable to use tempo and those younger than 6 years are unable to use mode to differentiate conventionally happy from sad musical excerpts from the classical repertoire. In the present experiment, selected piano excerpts from Vieillard *et al.* (2008), which were composed in the style of romantic film music, were used to evaluate the ability of 4- to 7-year-old child CI users to distinguish happy- from sad-sounding music. On average, these children had roughly 4 years of implant experience compared with the 7 years of average implant experience of children in Hopyan *et al.*'s (2011) study of emotion recognition. Clinically, the hearing age of the present sample would be considered 4 years, corresponding to their years of auditory input. In line with the findings by Dalla Bella *et al.* (2001), child CI users with less than 5 years of auditory experience might be unable to associate music excerpts with happy and sad emotions. Aside from a reduced quantity of auditory input relative to same-age peers, child CI users would have had a reduced quality of musical input. In view of their limited auditory and musical experience, it was important to ascertain whether they could differentiate emotions from samples of music. Unlike the happy and sad speech samples in Experiment 1, which differed in amplitude variability, the tones in all musical excerpts were equivalent in amplitude, eliminating an important cue to happiness and sadness. Mode cues were available in the present stimuli, but they were expected to be potentially useful only for the control sample of NH listeners.

Method

Participants

The participants were 14 CI users (six girls and eight boys), 12 of whom took part in Experiment 1. The average age of participants was 5.8 years (SD = 0.8, range: 4.1–6.9), and the average duration of implant experience was 4.2 years (SD = 1.0, range: 2.8–5.9). The two additional participants, both girls, were congenitally deaf and satisfied the criteria for participants in Experiment 1. The control sample consisted of the same 18 NH children who were tested in Experiment 1. At the time of testing, three children, all CI users, were taking piano lessons, one (CI-8) for 2 years, one (CI-6) for 6 months, and one (CI-2) for 2 months. Two NH children participated in early child music programs when they were approximately 3 years of age. In general, parents expressed favorable attitudes toward music, both for themselves and for their children, and all of them were providing informal musical exposure at home. In addition, all children were getting some exposure to music in their preschool or elementary school programs. Nevertheless, child CI users would have had less overall exposure to music by virtue of their lesser ability to profit from overhearing.

Apparatus and Stimuli

The musical stimuli consisted of 10 10-second synthesized piano excerpts, five happy and five sad excerpts, from the corpus of Vieillard *et al.* (2008), which includes excerpts for both emotions (as well as scary- and peaceful-sounding excerpts). The excerpts (copyright Bernard Bouchard, 1998) were composed in the style of romantic film music to express these four emotions. Each excerpt comprised different tones and tone durations but all tones were of equal amplitude. The 10 excerpts used here were selected from a larger sample because of adults' relative ease of identifying their emotional status (Hunter *et al.*, 2011). The happy-sounding excerpts were in the major mode and had a rapid tempo (mean of 137 beats per minute), in contrast to the sad-sounding excerpts, which were in the minor mode and had a slow tempo (mean of 46 beats per minute). Visual depictions of happiness and sadness consisted of close-ups of a frame from each of two animated films by Hirao Miyazaki, 'My Neighbor Totoro' (1988), and 'Spirited Away' (2001). Whereas the photographs of a man and woman with happy and sad facial expressions in Experiment 1 could be linked logically to the male and female talkers, they had no natural link to the instrumental music. Instead, imaginative cartoon portrayals, one of a child laughing and another of a child wiping tears from her eyes, were chosen for use with the expressive musical portrayals of happiness and sadness. The apparatus was identical to that of Experiment 1.

Procedure

The procedure was similar to that of Experiment 1 except that no practice trials were deemed necessary (i.e. unlike Experiment 1, there were no potentially distracting words). First, children were asked whether the children in the pictures looked happy or sad, which every child answered correctly without hesitation. NH children heard five happy- and five sad-sounding excerpts presented randomly for a total of 10 trials. They were not given a second block of trials, as originally intended, because their performance was near-perfect. Children with CIs heard two blocks of the 10 trials with the order of stimuli randomized in both blocks. No feedback was provided for correct or incorrect responses.

Results and discussion

NH children, who only received one block of trials, performed near ceiling (97.8% correct), and they were much less variable than the child CI users, $F(13, 17) = 22.84$, $P < 0.001$ (Levene's test). Moreover, performance was not distributed normally ($P = 0.005$) (Kolmogorov–Smirnov test). Thus, nonparametric analyses were used, as in Experiment

1. We initially examined how many children exceeded chance levels, which required scores of 9 or 10 correct on the 10 trials in each block (binomial test, $P < 0.05$, one-tailed). On the first block, 5 of 14 children with CIs and 17 of 18 NH children exceeded chance. On the second block of trials, 10 of 14 CI children surpassed chance. For the first block (i.e. the only block completed by both groups), the difference between groups in the proportion of children exceeding chance was significant ($\chi^2(1,\ N = 32) = 12.64,\ P < 0.001$). A nonparametric comparison of actual scores, contrasting the NH children with CI users (first block, see Fig. 5), also confirmed an advantage for the NH group ($P < 0.001$) (Mann–Whitney U test). For child CI users, improvement across trial blocks was not significant ($P = .283$) (Wilcoxon signed rank test). Individual differences among the CI children, collapsed across blocks, are illustrated in Fig. 6. In general, the performance of 4- to 7-year-old CI users (first block: 76.4% correct; second block: 83.6%) was comparable to that reported by Hopyan *et al.* (2011) for 7- to 13-year-old CI users (78% correct), who were tested on a similar (happy/sad) task with a different set of music excerpts.

Child CI users readily distinguished happy- from sad-sounding music although not with the extraordinary accuracy shown by their hearing peers, who could capitalize on pitch structure as well as tempo cues. Despite their limited auditory and musical exposure, young CI users' ability to identify happy and sad emotions in samples of instrumental music devoid of amplitude cues implies that (1) they would fare far better with real-world samples that feature amplitude as well as tempo cues, including children's music (Mote, 2011); and (2) cognitive or task factors must underlie the inability of some 4-year-old hearing children to identify musical emotions (Dalla Bella *et al.*, 2001).

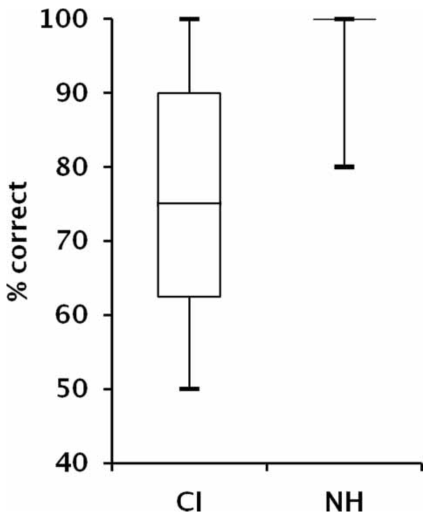

Figure 5 Box plots showing performance of child CI users and NH children on happy and sad music (Experiment 2). Lower and upper borders of each box indicate first and third quartiles, respectively; line within each box indicates the median; and whiskers depict the highest and lowest scores. Ceiling performance for most NH children resulted in identical values (100%) for the median, lower, and upper quartiles, and maximum.

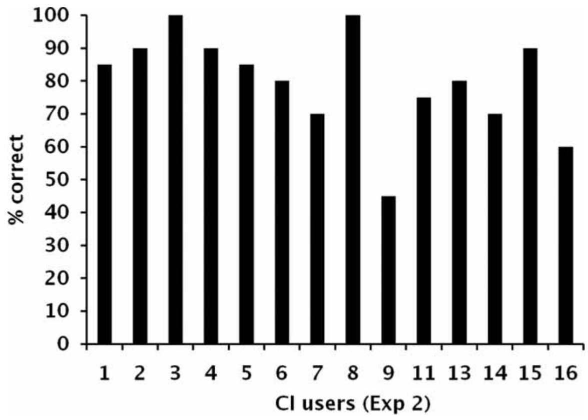

Figure 6 Performance of individual child CI users on happy and sad music (Experiment 2).

For child CI users, the association between duration of implant use (i.e. months since first implant activation) and performance collapsed across blocks was significant ($r = 0.49,\ N = 14,\ P = 0.038$) (one-tailed), as it was in Experiment 1, which is impressive once again in light of the small sample size. It is likely that excellent cognitive skills in combination with auditory experience enabled some children to learn which emotional labels are linked to which acoustic cues in speech and music. We also correlated performance on the first block of trials with performance from Experiment 1 separately for the 18 NH children and the 12 CI users who participated in both experiments. For the NH group, the correlation was not significant ($P = 0.367$), presumably because of high levels of performance and little variation in either experiment. For the CI group, however, there was a positive association ($r = 0.51,\ N = 12,\ P = 0.043$) (one-tailed). Although there are some common cues to emotion in speech and music, such as pitch, tempo, and amplitude (Juslin and Laukka, 2003), pitch cues to happiness and sadness were unlikely to have been useful for child CI users, tempo cues were inconsistent in the speech samples of Experiment 1, and amplitude cues were unavailable in the music excerpts.

CI-8, the only CI user who performed perfectly in the speech and music tasks, had the typical profile of so-called 'star' performers, which includes genetic, non-syndromic congenital deafness (Kawasaki *et al.*, 2006; Wu *et al.*, 2008), well-educated and highly involved parents (Geers and Brenner, 2003; Teagle and Eskridge, 2010), and at 6.3 years of age, over 5 years of implant experience. At the time of testing, moreover, he had been taking piano lessons for 2 years.

General discussion

In two experiments, 4- to 7-year-old deaf children with bilateral CIs and age-matched NH children identified happiness and sadness in speech and music in the context of a two-alternative forced-choice task. CI users performed well above chance levels but significantly below their hearing peers. The present findings with speech stimuli contrast with the poor performance of child CI users in previous studies of emotion identification in speech (Hopyan-Misakyan *et al.*, 2009; Luo *et al.*, 2007; Most and Aviner, 2009). Note, however, that listeners in those studies were required to identify vocal emotions from four or more alternatives, some of which had overlapping acoustic cues arising from similar arousal levels (e.g. happiness and anger). Variations in pitch or intonation, which contribute to emotion identification, are more likely to pose difficulty for CI users, but intonation differences are often accompanied by differences in speaking rate and amplitude, which would be accessible to CI users. Although hearing listeners can usually count on reliable pitch cues to emotion, CI listeners are likely to make use of alternative cues to emotion in the speech signal. Young CI users' ability to identify basic vocal emotions such as happiness or sadness on the basis of incidental exposure suggests that training in this realm could lead to enhanced perception and production of emotional prosody. In light of the consequences of vocal emotion identification for socialization and well-being in young CI users (Schorr *et al.*, 2009), the addition of such training to the current habilitation agenda for child CI users seems warranted.

The ability of 4- to 7-year-old CI users to differentiate happy from sad classical music excerpts extends the findings of Hopyan *et al.* (2011) to younger children with lesser implant experience and adds to the growing literature on the accessibility of music to prelingually deaf implant users. Although the present findings indicate that young CI users can discern the emotional intentions expressed in music, they shed no light on the emotional consequences of music for CI users. There are indications that young CI users enjoy music (Mitani *et al.*, 2007; Stordahl, 2002; Vongpaisal *et al.*, 2006), but it remains to be determined whether they experience changes in arousal and mood comparable to those experienced by individuals with NH (e.g. Balkwill and Thompson, 1999; Husain *et al.*, 2002). Music training is associated with enhanced speech perception (Moreno *et al.*, 2009), executive function (Degé *et al.*, 2011; Kraus and Chandrasekaran, 2010; Moreno *et al.*, 2011), and IQ (Schellenberg, 2004) in NH children. Such training may have even greater benefits for children with CIs, a possibility that awaits further research.

Acknowledgements

The authors gratefully acknowledge the parents and children who participated in this research, Deanna Feltracco for assistance with data collection, and Vicky Papaioannou and Jerome Valero for general assistance and encouragement. Funding for this research was provided by the Natural Sciences and Engineering Research Council of Canada.

References

Bachorowski J.A. 1999. Vocal expression and perception of emotion. *Current Directions in Psychological Science*, 8: 53–57.

Balkwill L.-L., Thompson W.F. 1999. A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Perception*, 17: 43–64.

Bänziger T., Scherer K.R. 2005. The role of intonation in emotional expressions. *Speech Communication*, 46: 252–267.

Bergeson T.R. 2011. Maternal speech to hearing-impaired infants in the first year of hearing aid or cochlear implant use: a preliminary report. *Cochlear Implants International*, 12: S101–S104.

Boersma P., Weenink D. 2005. Praat: doing phonetics by computer (Version 4.3.01) [Computer program]. Available from: http://www.praat.org/.

Brockmeier S.J., Fitzgerald D., Searle O., Fitzgerald H., Grasmeder M., Hilbig S., Vermiere K., Peterreins M., Heydner S., Arnold W. 2011. The music perception test: a novel battery for testing music perception of cochlear implant users. *Cochlear Implants International*, 12: 618–626.

Bryant G.A., Barrett H.C. 2007. Recognizing intentions in infant-directed speech: evidence for universals. *Psychological Science*, 18: 746–751.

Chatterjee M., Peng S.C. 2007. Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hearing Research*, 235: 143–156.

Cooper R.P., Aslin R.N. 1990. Preference for infant-directed speech in the first month after birth. *Child Development*, 61: 1584–1595.

Cooper W.B., Tobey E., Loizou P. 2008. Music perception by cochlear implant and normal hearing listeners as measured by the Montreal Battery for the Evaluation of Amusia. *Ear & Hearing*, 29: 618–626.

Cullington H.E., Zeng F.-G. 2011. Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. *Ear & Hearing*, 32: 16–30.

Dalla Bella S., Peretz I., Rousseau L., Gosselin N. 2001. A developmental study of the affective value of tempo and mode in music. *Cognition*, 80: B1–B10.

Dammeyer J. 2010. Psychosocial development in a Danish population of children with cochlear implants and deaf and hard-of-hearing children. *Journal of Deaf Studies and Deaf Education*, 15: 50–58.

Degé F., Kubicek C., Schwarzer G. 2011. Music lessons and intelligence: A relation mediated by executive functions. *Music Perception*, 29: 195–201.

DesJardin J.L., Eisenberg L. 2007. Maternal contributions: supporting language development in deaf children with cochlear implants. *Ear & Hearing*, 28: 456–469.

Fernald A. 1985. Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development*, 8: 181–195.

Fernald A. 1989. Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 60: 1497–1510.

Fernald A. 1991. Prosody in speech to children: Prelinguistic and linguistic functions. *Annals of Child Development*, 8: 43–80.

Galvin J.J., Fu Q.J., Nogaki G. 2007. Melodic contour identification by cochlear implant listeners. *Ear & Hearing*, 28: 302–319.

Geers A.E. 2004. Speech, language, and reading skills after early cochlear implantation. *Head & Neck Surgery*, 130: 634–638.

Geers A., Brenner C. 2003. Background and educational characteristics of prelingually deaf children implanted by five years of age. *Ear & Hearing*, 24: 2S–14S.

Geers A.E., Nicholas J.G., Sedey A.L. 2003. Language skills of children with early cochlear implantation. *Ear & Hearing*, 24: 46S–58S.

Geurts L., Wouters J. 2001. Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *Journal of the Acoustical Society of America*, 109: 713–726.

Gfeller K., Christ A., Knutson J.F., Witt S., Murray K.T., Tyler R.S. 2000. Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *Journal of the American Academy of Audiology*, 11: 390–406.

Holt R.F., Svirsky M.A. 2008. An exploratory look at pediatric cochlear implantation: is earliest always best? *Ear & Hearing*, 29: 492–511.

Hopyan T., Gordon K.A., Papsin B.C. 2011. Identifying emotions in music through electrical hearing in deaf children using cochlear implants. *Cochlear Implants International*, 12: 21–26.

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: 136–146.

Hsiao F. 2008. Mandarin melody recognition by pediatric cochlear implant recipients. *Journal of Music Therapy*, 45: 390–404.

Hunter P.G., Schellenberg E.G. 2010. Music and emotion. In: Jones M.R., Fay R.R., Popper A.N., (eds). *Music perception*. New York: Springer, p. 129–164.

Hunter P.G., Schellenberg E.G., Stalinski S.M. 2011. Liking and identifying emotionally expressive music: Age and gender differences. *Journal of Experimental Child Psychology*, 110: 80–93.

Husain G., Thompson W.F., Schellenberg E.G. 2002. Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception*, 20: 151–171.

Johnstone T., Scherer K.R., 2000. Vocal communication of emotion. In: Lewis M., Haviland J., (eds.) *Handbook of emotion*. 2nd ed. New York: Guilford, p. 220–235.

Juslin P.N., Laukka P. 2003. Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129: 770–814.

Kawasaki A., Fukushima K., Kataoka Y., Fukuda S., Nishizaki K. 2006. Using assessment of higher brain functions of children with GJB2-associated deafness and cochlear implants as a procedure to evaluate language development. *International Journal of Pediatric Otorhinolaryngology*, 70: 1343–1349.

Kirschner S., Tomasello M. 2010. Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, 31: 354–364.

Kondaurova M.V., Bergeson T.R. 2011. The effects of age and infant hearing status on maternal use of prosodic cues for phrase boundaries in speech. *Journal of Speech, Language, and Hearing Research*, 54: 740–754.

Kong Y.-Y., Cruz R., Jones J. A., Zeng F.-G. 2004. Music perception with temporal cues in acoustic and electric hearing. *Ear & Hearing*, 25: 173–185.

Kraus N., Skoe E., Parbery-Clark A., Ashley R. 2009. Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Annals of the New York Academy of Sciences*, 1169: 543–557.

Kraus N., Chandrasekaran B. 2010. Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11: 599–605.

Lassaletta L., Castro A., Bastarrica M., Pérez-Mora R., Madero R., De Sarriá J., Gavilán J. 2007. Does music perception have an impact on quality of life following cochlear implantation? *Acta Oto-Laryngologica*, 127: 682–686.

Laukka P., Juslin P.N., Bresin R. 2005. A dimensional approach to vocal expression of emotion. *Cognition and Emotion*, 19: 633–653.

Leal M.C., Young J., Laborde M.-L., Calmels M.-N., Verges S., Lugardon S., *et al.* 2003. Music perception in adult cochlear implant recipients. *Acta Oto-Laryngologica*, 123: 826–835.

Loizou P. 1998. Mimicking the human ear. *IEEE Signal Processing Magazine*, 15: 101–130.

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

Masataka N. 1999. Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental Psychology*, 35: 1001–1005

McDermott H.J. 2004. Music perception with cochlear implants: A review. *Trends in Amplification*, 8: 49–82.

Meister H., Landwehr M., Pyschny V., Walger M., von Wedel H. 2009. The perception of prosody and speaker gender in normal-hearing listeners and cochlear implant recipients. *International Journal of Audiology*, 48: 38–48.

Mitani C., Nakata T., Trehub S.E., Kanda Y., Kumagami H., Takasaki K., Miyamoto I., Takahashi H. 2007. Music recognition, music listening, and word recognition by deaf children with cochlear implants. *Ear and Hearing*, 28: 29S–33S.

Moreno S., Bialystok E., Barac R., Schellenberg E.G., Cepeda N.J., Chau T. 2011. Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, 22: 1425–1433.

Moreno S., Marques C., Santos A., Santos M., Castro S.L., Besson M. 2009. Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, 19: 712–723.

Morton J.B., Trehub S.E. 2001. Children's understanding of emotion in speech. *Child Development*, 72: 834–843.

Most T., Aviner C. 2009. Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *Journal of Deaf Studies and Deaf Education*, 14: 449–464.

Most T., Peled M. 2007. Perception of suprasegmental features of speech by children with cochlear implants and children with hearing aids. *Journal of Deaf Studies and Deaf Education*, 12: 350–361.

Mote J. 2011. The effects of tempo and familiarity on children's affective interpretation of music. *Emotion*, 11: 618–622.

Nakata T., Trehub S.E., Mitani C., Kanda Y. 2006. Pitch and timing in the songs of deaf children with cochlear implants. *Music Perception*, 24: 147–154.

Nicholas J.G., Geers A.E. 2006. Effects of early auditory experience on the spoken language of deaf children at 3 years of age. *Ear and Hearing*, 27: 286–298.

Nienhuys T.G., Cross T.G., Horsborough K.M. 1984. Child variables influencing maternal speech style: Deaf and hearing children. *Journal of Communication Disorders*, 17: 189–207.

Nimmons G.L., Kang R.S., Drennan W.R., Longnion J., Ruffin C., Worman T., Yueh B., Rubinstein J.T. 2007. Clinical assessment of music perception in cochlear implant listeners. *Otology and Neurotology*, 29: 149–155.

Papoušek M. 1992. Early ontogeny of vocal communication in parent–infant interactions. In: Papoušek H., Jürgens U., Papoušek M., (eds.) *Nonverbal vocal communication: comparative and developmental approaches*. New York: Cambridge University Press, p. 230–261.

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 & Hearing*, 29: 336–351.

Percy-Smith L., Jensen J.H., Caye-Thomasen P., Thomsen J., Gudman M., Lopez A.G. 2008. Factors that affect the social well-being of children with cochlear implants. *Cochlear Implants International*, 9: 199–214.

Peretz I., Gagnon L., Bouchard B. 1998. Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68, 111–141.

Salimpoor V.N., Benovoy M., Longo G., Cooperstock J.R., Zatorre R.J. 2009. The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS ONE*, 4: e7487.

Schellenberg E.G. 2004. Music lessons enhance IQ. *Psychological Science*, 15: 511–514.

Schellenberg E.G. 2006. Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology*, 98: 457–468.

Schellenberg E.G. 2011. Examining the association between music lessons and intelligence. *British Journal of Psychology*, 102: 283–302.

Schellenberg E.G., Hallam S. 2005. Music listening and cognitive abilities in 10- and 11-year-olds: the Blur effect. *Annals of the New York Academy of Sciences*, 1060: 202–209.

Schellenberg E.G., Nakata T., Hunter P.G., Tamoto S. 2007. Exposure to music and cognitive performance: Tests of children and adults. *Psychology of Music*, 35: 5–19.

Scherer K.R. 1986. Vocal affect expression: a review and a model for future research. *Psychological Bulletin*, 99: 143–165.
Scherer K.R. 2003.Vocal communication of emotion: a review of research paradigms. *Speech Communication*, 40: 227–256.
Scherer K.R., Banse R., Wallbott H.G. 2001. Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, 32: 76–92.
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 & Hearing Research*, 52: 141–152.
Singh L., Morgan J.L., Best C.T. 2002. Infants' listening preferences: Baby talk or happy talk? *Infancy*, 3: 365–394.
Stordahl J. 2002. Song recognition and appraisal: a comparison of children who use cochlear implants and normally hearing children. *Journal of Music Therapy*, 39: 2–19.
Teagle H.F.B., Eskridge H. 2010. Predictors of success for children with cochlear implants: the impact of individual differences. In: Weiss A.L. (ed.). *Perspectives on individual differences affecting therapeutic change in communication disorders*. New York: Psychology Press, p. 251–272.
Trainor L.J. 1996. Infant preferences for infant-directed versus non-infant-directed playsongs and lullabies. *Infant Behavior & Development*, 19: 83–92.
Trainor L.J., Austin C.M., Desjardins R.N. 2000. Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, 11: 188–195.
Trehub S.E., Hannon E.E., Schachner A. 2010. Perspectives on music and affect in the early years. In: Juslin P.N., Sloboda J.A., (eds.). *Handbook of music and emotion*. New York: Oxford University Press, p. 645–668.
Trehub S.E., Trainor L.J. 1998. Singing to infants: Lullabies and play songs. *Advances in Infancy Research*, 12, 43–77.
Trehub S.E., Trainor L.J., Unyk A.M. 1993. Music and speech processing in the first year of life. *Advances in Child Development and Behavior*, 24: 1–35.
van Besouw R.M., Grasmeder M.L., Hamilton M.E., Baumann S.E. 2011. Music activities and responses of young cochlear implant recipients. *International Journal of Audiology*, 50: 340–348.
Vieillard S., Peretz I., Gosselin N., Khalfa S., Gagnon L., Bouchard B. 2008. Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion*, 22: 720–752.
Vongpaisal T., Trehub S.E., Schellenberg E.G. 2006. Song recognition by children and adolescents with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 49: 1091–1103.
Vongpaisal T., Trehub S.E., Schellenberg E.G. 2009. Identification of TV tunes by children with cochlear implants. *Music Perception*, 27: 17–24.
Vongpaisal T., Trehub S.E., Schellenberg E.G., van Lieshout P., Papsin B.C. 2010. Children with cochlear implants recognize their mother's voice. *Ear & Hearing*, 31: 555–566.
Waltzman S.B., Cohen N.L. 1998. Cochlear implantation in children younger than 2 years old. *American Journal of Otology*, 19: 158–162.
Wu C.C., Lee Y.C., Chen P.J., Hsu C.J. 2008. Predominance of genetic diagnosis and imaging results as predictors in determining the speech perception performance outcome after cochlear implantation in children. *Archives of Pediatrics & Adolescent Medicine*, 162: 269–276.
Zentner M., Grandjean D., Scherer K.R. 2008. Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, 8: 494–521.