

Vocal emotion identification by children using cochlear implants, relations to voice quality and musical interests

Teija Waaramaa¹, Tarja Kukkonen², Sari Mykkänen³, Ahmed Geneid⁴

¹Tampere Research Centre for Journalism, Media and Communication (COMET), Faculty of Communication Sciences, University of Tampere, Tampere, Finland, e-mail: teija.waaramaa@uta.fi.

²Faculty of Social Sciences/Logopedics, University of Tampere, Tampere, Finland

³Hearing Centre, Tampere University Hospital, Tampere, Finland

⁴Department of Otorhinolaryngology and Phoniatics – Head and Neck Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Abstract

Purpose: Listening tests for emotion identification were conducted with 8–17 year-old children with hearing impairment (HI) ($N=25$) using cochlear implants (CIs), and their 12-year-old peers with normal hearing (NH) ($N=18$). The study examined the impact of musical interests and acoustics of the stimuli on correct emotion identification.

Method: The children completed a questionnaire with their background information and noting musical interests. They then listened to vocal stimuli produced by actors ($N=5$) and consisting of nonsense sentences and prolonged vowels ([a:], [i:] and [u:]; $N=32$) expressing excitement, anger, contentment, and fear. The children's task was to identify the emotions they heard in the sample by choosing from the provided options. Acoustics of the samples were studied using Praat software, and statistics were examined using SPSS24 software.

Results: The HI children identified the emotions with 57% accuracy, and the NH children with 75% accuracy. Female listeners were more accurate than males in both groups. Those who were implanted before age of three identified emotions more accurately than others ($p < .05$). No connection between the child's audiogram and correct identification was observed. Musical interests and voice quality parameters were found to be related to correct identification.

Conclusions: Implantation age, musical interests and voice quality tended to have an impact on correct emotion identification. Thus, in developing the CIs, it may be worth paying attention to the acoustic structures of vocal emotional expressions, especially the formant frequency of F3. Supporting the HI children's musical interests may help their emotional development and improve their social lives.

Keywords: acoustics, children, cochlear implant, voice

Vocal emotion identification by children using cochlear implants, relations to voice quality and musical interests

Childhood is an important time for learning social interaction, expressing and perceiving emotions, and understanding humor and double meanings. These expressions are not necessarily conveyed by actual words but are often conveyed by changing the tone of one's voice (i.e., voice quality; Fant, 1970). It is important that children learn to understand the meaningful nuances of the voice as early as possible in childhood. Restrictions in the ability to recognize vocal and nonverbal communication may disturb the normal development of emotion expression and perception as well as that of speech itself. A developmental delay due to hearing impairment (HI), for example, may lead to complications in interpersonal interactions and communication, which can result in withdrawal and exclusion from social situations. Therefore, it is essential to examine how the children with cochlear implants (CIs) perform in detecting and identifying vocal emotional cues from speech samples that reflect different voice qualities and express a variety of emotions.

Infants learn emotion expression and regulation in interaction with their caretakers at a very early stage, during the first years of life (Flavell, 2004; Fox & Calkins, 2000). Thereafter and into adulthood, emotion regulation develops more slowly (Eisenberg, Spinrad & Eggum, 2010). Emotion regulation, defined as "a skill that involves coping with emotions (i.e. internally regulating their intensity) as well as expressing them" (Wiefferink, Rieffe, Ketelaar & Frijns, 2012a), includes situation selection, situation modification, attentional deployment, cognitive change, and response modulation (Eisenberg et al., 2010). Even though emotion expression and emotion regulation, which are profound elements in an infant's primary secure attachment and social interaction, can be seen as normative systems, there are individual differences between infants. Both the temperament of the infant and the caretaker's feedback

have a remarkable impact on the child's future interaction relationships (Eisenberg et al. 2010; Fox & Calkins, 2000). The caretaker's ability to support the infant in stressful situations may affect the infant's ability to be alone and to be relatively independent of the caretaker. A lack of caretaker support may also affect the development of the child's sense of security: An inability to tolerate stress may lead to a sense of insecurity and the withdrawal from social situations (Fox & Calkins, 2000). Caretakers can support the development of their children's emotions by validating and labelling the emotions, teach coping strategies, and helping them to improve their regulatory skills (Eisenberg et al., 2010; Ketelaar & Wiefferink, 2015). Self-awareness develops by the age of two, and starting from the age of three, children with normal hearing (NH) begin to internalize moral standards and to channel their emotional behavior without their caretakers' guidance (Ketelaar & Wiefferink, 2015).

The interaction between children and their caretakers is also influenced by their skills with respect to language, emotion expression, and perception (Shorr, Roth & Fox, 2009; Wiefferink et al. 2012; Wiefferink, Rieffe, Ketelaar, De Raeve & Frijns, 2012). Wiefferink et al. (2012) reported that the language skills of NH children were stronger than those of HI children. If an infant has HI, the development of understanding the caretaker's vocal expressions and connecting them to facial expressions is unavoidably disturbed (Scheiner, Hammerschmidt, Jürgens & Zwirner, 2006), which further affects the children's social behavior and social skills, their ability to demonstrate empathy as well as to understand others' emotions (Netten, Rieffe, Theunissen, Soede, Dirks, Briaire & Frijns, 2015; Tamoepeau & Ruffman, 2008). Wiefferink et al. (2012) reported a lower level of social competence and a reduced ability to use adequate coping strategies in children with HI when compared with NH children. According to the researchers, better social skills are related to more frequent expressions of positive emotions (Wiefferink et al. 2012). Ketelaar and Wiefferink (2015) observed a poorer understanding of moral emotions, such as shame and guilt, and poorer ability to relate them to

social functioning in children with HI when compared with NH children. Wiefferink et al. (2012) stated that if emotion-regulation strategies do not develop well in the early years of life, optimal strategies cannot be learned later in life. Therefore, children with HI benefit profoundly from a very early implantation, as it improves the child's social interaction, sense of coherence, mental health (Anmyr, Olsson, Freijd & Larsson, 2015), and development of language skills, while decreasing emotional problems (De Giacomo, Graig, D'Elia, Giagnotti, Matera & Quaranta, 2013; Ketelaar & Wiefferink, 2015; Wiefferink et al., 2012). Researchers stress the importance of early intervention to ensure an enhanced quality of life for children with HI (Shorr et al., 2009) and the follow-up of hearing and behavioral performance to enable the design of tailored rehabilitation programs (Chao, Lee, Liu, Tsou & Chan, 2015). Research has shown that a better quality of life is related to emotion identification by HI children (Shorr et al., 2009). However, there is significant variability in speech perception (Shorr et al., 2009), and individual differences between HI children are notable in social competence, behaviors, and academic performance.

Earlier studies on emotion recognition by children and adolescents with HI have suggested that they experience significant delays or deficits in their ability to recognize or to understand vocal or facial expressions of emotion and to judge the expected emotional response of others to a particular event (Dyck, Farrugia, Shochet, & Holmes-Brown, 2004; Gray, Hosie, Russell, Scott & Hunter, 2007; Ludlow, Heaton, Rosset, Hills & Deruelle, 2010; Most & Michaelis, 2012; Wake, Hughes, Poulakis, Collins, & Rickards., 2004; Wiefferink, Rieffe, Ketelaar, De Raeve, et al., 2012), or to recognize emotions from a speaking voice (Chatterjee et al., 2015; Luo, Fu, & Galvin, 2007; Most & Aviner, 2009; Peng, Tomblin, & Turner, 2008; Scheiner et al., 2006). In particular, perception of spectral information and the duration of the phonemes have been shown to be inadequate in people who have HI and CIs (Hopyan-Misakyan, Gordon, Dennis & Papsin, 2009; Jiam, Caldwell, Deroche, Chatterjee & Limb, in

press; Luo et al., 2007; Stiles & Nadler, 2013; Wei, Cao, & Zeng, 2004). Listeners with HI seem to rely more on tempo, speech rate and its variations, and intensity in emotion perception (Giannantonio, Polonenko, Papsin, Paludetti & Gordon, 2015; Hopyan, Manno III, Papsin & Gordon, 2016; Jiam et al., in press; Nakata, Trehub & Kanda, 2012).

The auditory processing of HI people has been found to be slower than that of NH people without hearing aids (Lasak, Allen, McVay, & Lewis, 2014). The materials used in processing may be partly related to the reaction time of CIs and hearing aids (Stone, Moore, Meisenbacher, & Derleth, 2008). Consequently, social interaction may be more difficult for people with HI than for those with NH.

Emotions are conveyed through body language, facial expressions, and the voice, the quality of which varies according to the conveyed emotion. Voice quality is defined as a combination of the voice source and vocal tract functions (Fant, 1970). The voice source, or vocal fold vibration, determines the fundamental frequency (F_0) of the voice. Vocal tract functions (formants) are vocal tract resonances. Both voice source and formants affect the manner in which sound energy is distributed along the frequency range on the spectrum. In hyperfunctional voice production, the spectral slope is flatter than that in hypofunctional voice production, where the slope is steeper (Gauffin & Sundberg, 1989). Thus, a hyperfunctional phonation type carries more energy in the higher frequency area than that carried in a hypofunctional phonation type. These two phonation types, hyperfunctional and hypofunctional, are perceived respectively as pressed and breathy voice qualities. Perceptually, more energy in the higher frequency area means louder voice. Typically, with respect to joy or anger, there is a large amount of energy in the higher frequency area. In terms of emotions such as tenderness or sadness, the spectral slope is more tilted; thus, the energy peaks are lower and, perceptually, not as loud as in emotions conveyed with hyperfunctional voice quality. Therefore, the fine nuances in the emotions with hypofunctional voice quality are harder to hear

or may even be perceived as silent (Gauffin & Sundberg, 1989).

The development of hearing devices has progressed rapidly since 1990s when physicians were pleased if an implanted child could hear the sound of an alarm vehicle while using the CI device. The development of CIs has since concentrated on word discrimination, i.e., speech recognition. Today, we are able to test the more subtle characteristics of vocal expression, for example, emotional nuances. Vocal nuances, like those associated with emotions are much less studied than speech production and perception of HI children. Thus, the present study aimed to examine whether HI children can generally identify emotional colorings of the voice (different voice qualities) in nonsense sentences and prolonged vowel sounds without linguistic content, and if so, whether the identification differs from that by NH children. It was also of interest to study children's own subjective understanding of their relations to music and musical interests, and whether their self-reported (i.e., not reported, for example, by their parents) musical interests are connected to their ability to perceive emotional cues from speech signals. According to Juslin and Laukka (2003), the emotional expressiveness of music is based on the similarities of the emotional acoustic cues in vocal expressions (see also Waaramaa and Leisiö, 2013). Thus, speech and emotional music may engage the same neural processes (Juslin and Västfäll, 2008).

Additionally, the acoustic structure of the stimuli was studied, and the connection between the acoustics and emotion identification was examined. In the present study, we also wanted to examine whether there were gender differences among the listeners with respect to perceptions of emotions from vocal utterances. Gender differences may be expected since voice quality parameters have been shown to be connected to our socio-linguistic consciousness and personality characteristics (Addington, 1968; Lukkarila, Laukkanen, & Palo, 2012). Depending on whether the speaker is a male or a female, the same voice quality may be perceived as positive or negative (Addington 1968), indicating that voice quality may also be gender-related.

The present study addressed the following research questions:

RQ1: Does the child's chronological or implantation age affect his/her emotion or valence identification from vocal emotional samples?

RQ2: Are the perceptual differences based on gender?

RQ3: Are there connections between the child's audiogram, the acoustic structures of the stimuli, and the emotion identification?

RQ4: Do the child's musical interests have an effect on the emotion recognition?

Methods

Recordings and the stimuli

In the present study, the recordings of the emotional expressions were made at the University of Tampere in a soundproof studio using Logic Pro X workstation and a Brüel & Kjær 4006 microphone, which was placed at a distance of 40 cm from the speakers' lips. Thirty volunteering professional and student actors participated in the production of the stimuli for the present study. The researchers chose the most expressive samples, produced by five speakers, for use in the listening test. The selected speakers comprised professional actors and student actors ($n = 5$; 2 males, 3 females), each of whom produced the same number of samples.

The stimuli consisted of nonsense sentences [*Komma taaku tuuka teekas. Pyyta rakku manki loibanea tiika kassy.*] and prolonged vowel sounds /a:/, /i:/ and /u:/ ($n = 32$, including 8 x nonsense sentences, 8 x [a:], 8 x [i:], and 8 x [u:].) that expressed excitement, anger, contentment, and fear. These emotions were chosen since they were considered to represent everyday-life emotions, which may be faced frequently. They also represent positive (excitement and contentment), and negative (fear and anger) emotions, with a high activity or arousal level (excitement and anger), and a low arousal level (contentment and fear). These

contrasts are purposeful in analyzing the connections between the acoustic structures of the stimuli and the perceptions of those stimuli.

Nonsense sentences and prolonged vowel sounds were used only to eliminate the linguistic content of the speech, thus ensuring that the focus of the test was purely on voice quality characteristics and not on the words and meanings they carry. The nonsense sentences were created by the researchers. Typical phonemes in the Finnish language were used in order to make the sentences easy to read for the Finnish speakers. The duration of each of the nonsense sentences was approximately 6571 milliseconds (ms) and the vowel sounds were approximately 629 ms.

Acoustic Analyses

The acoustics of the samples were studied using Praat software (Boersma & Weenink, 2010). The parameters measured were fundamental frequency (F0), formant frequencies F1, F2, F3, F4, sound pressure level (SPL), sample duration, number of periods, number and degree of voice breaks, jitter, shimmer, and harmonics to noise. Praat software uses an automatic calculation for all these voice parameters. Differences between the emotions and parameters measured were studied using Mann-Whitney U tests. It is well known that F0 and SPL vary in relation to emotional expressions. Frequencies of F1 and F2 are known to be related to the vowel expressed, whereas F3 and F4 may vary more freely. Frequencies of F3 and F4 have been shown to differ between emotional expressions (Laukkanen, Vilkman, Alku & Oksanen, 1997; Waaramaa et al., 2006; Waaramaa et al., 2008; Waaramaa, Palo, & Kankare, 2014). Voice breaks, jitter, shimmer, and harmonics to noise are also shown to be related to emotional expressions (Laukkanen et al., 2008; Toivanen, Waaramaa, Alku, Laukkanen, Seppänen, Väyrynen & Airas, 2006; Waaramaa et al., 2010; Waaramaa & Leisiö, 2013).

Participants

Twenty five ($n = 25$; 10 males, 15 females) HI children and with no other diagnosed symptoms participated in the study. Twelve of the children used bilateral CIs, ten used a single CI and three used a single CI and a hearing aid in the other ear. The children participated singularly in a listening test conducted in a soundproof studio at Tampere University Hospital. The participating children were 8–17 years old, with a mean age of 11 years for males and 12 years for females (standard deviation [SD] 2.47). Thirteen children were implanted during their second year of life, five children were implanted between their second and third year of life, and seven children were implanted after their third year of life; the oldest child at the time of implantation was nine years old.

Healthy, NH peers, with a mean age of 12 years (SD 0.76; $n = 18$; 8 males, 10 females), participated in the identification test. Six of the children were bilingual.

Identification Test

Emotion identification tests were conducted when the HI children came to the clinic for their twice-a-year device-control visit. It took two years to collect this data. Permission to participate in the study was given by the children and parents by proxy.

The NH peers participated in the identification test as a group in their classroom during a school day. Listening circumstances were thus less ideal (e.g. due to reverberation), and different when compared to HI listeners since it was not possible to use the soundproof studio with the NH children at the hospital. Nevertheless, the University of Tampere has an agreement of co-operation with the school, and therefore has permission to test NH children at that site.

The study was granted the approval of the Ethical Board of the University of Tampere, which was applied as the subjects were underage participants, and the HI children had medical diagnoses. Thus, the study met ethical requirements.

At the beginning of the listening test for emotion identification, the children completed a questionnaire by providing their personal background information and responding to nine questions about their musical interests. The children reported their age, gender and hearing, whether their hearing was normal or impaired. They also answered a question about alertness by choosing one of the following options: I am 1) alert, 2) tired, 3) in between of these two. The children described their mood by answering whether they were in a 1) good, 2) bad, or 3) neutral mood. Concerning the musical interests, the children were asked whether

- they liked to listen to music
- it was easy for them to respond to music
- they were interested in singing
- they played a musical instrument
- they were interested in dancing
- it is easy for them to dance in the correct rhythm
- it was easy for them to learn a new melody
- music affected their moods
- music caused them physical reactions.

The children chose their answer from the alternatives, Yes/Not anymore/Sometimes/No. The answer “Not any more” was an option to the question whether they played a musical instrument. The questionnaire was designed to examine the children’s own understanding and how they perceive their relation to music.

Second, the children listened to four emotional samples in order to ‘calibrate’ their ears for the test, and they were told which emotion each of the four samples represented. To ensure proper hearing of each stimuli, the children were asked to tell when they found the loudness to be at a comfortable level. The identification test was a forced-choice test. The listening instructions were

given orally to the children by the researchers. While listening, the HI children were faced toward three loudspeakers which were set at -60, 0 and 60 degrees, each placed at the same distance from the listener. It was necessary to set the loudspeakers in front of the listeners since those using a bilateral CI do not necessarily have direction hearing (Dunn, Tyler, Oakley, Gantz & Noble, 2008; Faulkner & Pisoni, 2013; Lovett, Kitterick, Hewitt & Summerfield, 2010; Neuman, Haravon, Sislian & Waltzman, 2007; Litovsky, Johnstone, Godar, Agrawal, Parkinson, Peters & Lake, 2006; van Hoesel & Tyler, 2003). The stimuli were presented to the three loudspeakers simultaneously, and the children were allowed to listen to the samples as many times as they wished (in general, once or twice per each sample). The children responded orally to the researchers regarding the emotion they perceived to be associated with the four options presented to them. One of the researchers played the samples in a random order, and the other wrote down the children's answers. The test took about 15 minutes for each participant.

The NH children used a questionnaire for responding to the identification task because they were simultaneously tested as a group in the classroom. The test was conducted using the computer and the loudspeakers in the classroom. The loudspeakers were situated in front of the classroom, both sides above the whiteboard.

Statistical Analyses

The acoustic parameters measured, the answers provided on the questionnaire, and the children's responses to the listening test were studied using SPSS software (IBM SPSS Statistics v. 22 for Windows, Armonk, NY, USA). Mann-Whitney U tests were conducted on the acoustic parameters measured in order to see whether there were significant differences between the emotional expressions for their acoustical structures. Correlations between the share (%) of the correctly identified emotional samples and their acoustic parameters were investigated.

When emotion identification is tested, researchers often group the emotions also according to their valence on an axis positive–neutral–negative emotion, and their activity or arousal level on an axis low–moderate–high (measured by dBs; Laukkanen, Alku, Airas & Waaramaa, 2008; Waaramaa, Alku, & Laukkanen, 2006; Waaramaa, Laukkanen, Airas & Alku 2010; Waaramaa, Laukkanen, Alku, & Väyrynen, 2008). If the listeners are not able to name the actual emotion, this grouping allows the researchers to analyze whether the listeners still can distinguish between the valences and arousal levels of the stimuli.

Recognition of the emotions and valence was studied by a crosstabulation of the expressed and perceived emotions. The valence and the arousal level groups were formed by the researchers; Excitement and contentment were considered positive emotions, and fear and anger were considered negative emotions; excitement and anger were considered high-arousal emotions; contentment and fear were considered low-arousal emotions. Pearson's chi square test and contingency coefficient tests were examined. Cronbach's alpha was calculated. The effect of the sample type (whether a nonsense sentence or a vowel sound) on the identification was evaluated via the use of Mann-Whitney U tests, which were also conducted on the emotion and valence identification and the answers provided on the questionnaire. The connection between correct emotion identification and the chronological age of the child, implantation age, and the gender of the child were calculated.

Pure tone average (PTA, 0.5–4 kHz) from each of the HI children's audiograms was studied and compared with the individual results of each listening task.

Results

Acoustic Analyses

The acoustic parameters measured from the emotions (Table 1) differed from each other significantly. There were statistically significant between excitement and contentment for

F0 and SPL ($U = 6.000$, $p < .01$; $U = 2.000$, $p < .01$). Anger and contentment differed significantly for F0 ($U = 6.500$, $p < 0.1$), SPL ($U = 0.000$, $p < .001$), duration ($U = 13.000$, $p < .05$), and shimmer ($U = 13.000$, $p < .05$). Fear differed significantly from excitement and anger for SPL ($U = 6.000$, $p < .01$). No significant differences were found between anger and excitement or between fear and contentment, most likely due to their similar arousal level.

Emotion	Average Standard deviation	F0	F1	F2	F3	F4	SPL	Duration (ms)	Number of pulses	Number of periods	Number of voice breaks	Degree of voice breaks (%)	Jitter (Hz)	Shimmer (%) (dB)	Mean harmonics to-noise ratio (dB)
excitement	av	243	627	1621	2878	4103	76	1881	327	321	4	10	1.6	7.3	12.8
excitement	sd	56	185	546	102	343	5	2410	362	354	6.5	18	0.7	2.1	3.3
fear	av	200	595	1607	2900	4036	68	2089	215	210	3.6	12	2.1	7.5	14.3
fear	sd	101	181	534	263	245	4	2920	234	227	6.7	23	0.8	2.4	2.9
anger	av	281	725	1743	2958	3986	79	1104	255	249	4	13	2.1	10.9	9.4
anger	sd	73	242	397	185	397	4	1694	267	260	6.5	18	1	5	4.9
contentment	av	151	621	1604	2830	3771	64	2672	244	240	3.6	12	2.5	6.4	14.4
contentment	sd	54	262	753	197	204	4	4713	398	389	9.2	37	0.7	2.1	1.4

Table 1. Averages and standard deviations of the acoustic parameters measured from the emotion expressions.

A statistically significant correlation was found between the frequency of F1 and the share (%) of the emotion ($r = .547$, $p < .01$) and valence ($r = .494$, $p < .01$) identified by the HI children. No other significant correlations were found between the acoustic parameters and the share of emotions or valence correctly identified by the HI children.

Among the NH children, the correct identification of emotions correlated statistically significantly with F0 ($r = .349$, $p = .05$); F1 ($r = .427$, $p < .05$); and F3 ($r = .354$, $p < .05$). The correct identification of valences correlated statistically significantly with F1 ($r = .396$, $p < .05$), and F3 ($r = .375$, $p < .05$). The frequencies of F0 (281 Hz), F1 (725 Hz), and F3 (2958 Hz) were highest in anger, which was the most accurately recognized emotion.

Identification Test

The HI children identified the emotions with 57% accuracy and the valences with 70% accuracy. Cronbach's alpha for the emotion identification was .378, which suggests a large dispersion in the answers provided. Gender differences were statistically significant for both

emotion and valence identification, with females demonstrating a higher level of accuracy (emotions: $df = 9$; $X^2(2) = 476.964$; $p < .001$); valence: ($df = 1$; $X^2(2) = 4.024$; $p < .05$; see also Brown & Dunn, 1996). The percentage for correct identification was higher for females with respect to each of the four emotions and for each sample type (nonsense sentences, vowels [a:], [i:] and [u:]) (Tables 2, 3, and 4). Hence, females demonstrated a higher level of accuracy than males in emotion and valence identification.

Emotion perception, Emotion expression, and Gender crosstabulation

		Gender	Emotion expressed				Total
			excitement	fear	anger	contentment	
Emotion perceived	Male	excitement	38	9	14	7	68
		fear	13	33	11	22	79
		anger	17	15	47	11	90
		contentment	12	23	8	40	83
	Female	excitement	62	7	11	11	91
		fear	18	55	11	16	100
		anger	27	24	96	8	155
		contentment	13	34	2	85	134
Total			200	200	200	200	800

Table 2. Crosstabulation of emotion expressed and emotion perceived by gender, HI children.

Emotion and valence identification by HI children

Perception	Males (N = 320)		Females (N = 480)		Pearson		Contingency coefficient	
	n	(%)	n	(%)	df	p	variable	p
Emotion perception	158	(49.4%)	298	(62.1%)	9	<0.001	.611	<0.001
Valence perception	203	(63.4%)	357	(74.4%)	2	<.05	.071	<.05

Table 3. Results of the Pearson's chi-square tests and contingency coefficient tests for emotion and valence identification, HI children. N = the total number of answers given by males and females separately. Percentages for the matching answers are in the paragraphs.

The NH children identified the emotions with 75% accuracy and valences with 84% accuracy. Cronbach's alpha for the emotion identification was .70 (Tables 5, 6, and 7). Females were also more accurate than males in the NH group with respect to emotion identification (df

=12; $X^2(15) = 804.095$; $p < .001$). Identification of valences did not differ significantly between genders in the NH group of children. Again, the percentage of correct identifications of emotions was higher for females with respect to each of the four emotions and for each sample type (nonsense sentences, vowels [a:], [i:] and [u:]).

There was a statistically significant difference between the HI and NH groups in the identification of emotion ($df = 1$; $X^2(2) = 44.581$; $p < .001$), and valence: ($df = 1$; $X^2(2) = 35.997$; $p < .001$). The identification of the arousal level did not differ between the HI and the NH children.

Emotion identified, Sample type, and Gender crosstabulation

Gender	Sample type					Total
	emotion identification	nonsense sentence	vowel [i:]	vowel [a:]	vowel [u:]	
Male	non-identified	35	43	33	51	162
	identified	45	37	47	29	158
Female	non-identified	38	51	32	61	182
	identified	82	69	88	59	298
Total		200	200	200	200	800

Table 4. Crosstabulation of the sample types, gender, and emotion identification, HI children.

From the emotional stimuli played for the HI children, anger was the most accurately recognized emotion, and it was also the most frequently chosen answer, whereas excitement was the least chosen answer (Table 2). Fear and contentment were frequently confused with one another, and they were most accurately recognized from nonsense sentences (62.5% and 76.9% identified, respectively; Table 4). Excitement and anger were most accurately recognized from vowel [a:] (74.3% and 78.9% identified, respectively). Perception of emotions was statistically significantly more accurately recognized in nonsense sentences compared to vowel [i:] ($U = 17.900$, $p < .05$), and vowel [u:] ($U = 16100.000$, $p < .001$). A statistically significant difference was also found between vowel [i:] and vowel [a:] ($U = 17100.000$, $p < .01$), with vowel [a:] being identified with better accuracy than vowel [i:]. Emotions in vowel [a:] were

also more accurately identified than emotions in vowel [u:] (U = 15300.000, $p < .001$). Perception of valence was statistically significantly more accurately recognized in nonsense sentences and in vowel [a:] (U = 16.400, $p < .001$) than in vowel [u:] (U = 15500.000, $p < .001$). In addition, vowels [i:] and [u:] differed from each other significantly (U = 18100.000, $p = .05$), with HI children perceiving vowel [i:] more accurately than vowel [u:].

Emotion perception, Emotion expression and Gender Crosstabulation

Gender		Emotion expressed				Total	
		excitement	fear	anger	contentment		
Emotion perceived	Male	missing	3	1	0	1	5
		excitement	35	5	2	6	48
	fear	9	36	2	8	55	
	anger	7	7	59	1	74	
	contentment	10	15	1	48	74	
Female	missing	0	0	2	1	3	
	excitement	57	5	0	3	65	
	fear	0	58	3	14	75	
	anger	5	5	75	2	87	
	contentment	18	12	0	60	90	
Total		144	144	144	144	576	

Table 5. Crosstabulation of emotion expressed and emotion perceived by gender, NH peers.

Emotion and valence identification by NH peers

Perception	Males (N = 256)		Females (N = 320)		Pearson chi2		Contingency coefficient	
	n	(%)	n	(%)	df	p	variable	p
Emotion perception	179	(69.9%)	250	(78.1%)	12	< .001	.255	< .001
Valence perception	203	(79.3%)	281	(87.8%)	12	.363	.038	.363

Table 6. Results of the Pearson's chi-square tests and contingency coefficient tests for emotion and valence identification, NH peers. N = the total number of answers given by males and females separately. Percentages for the matching answers are in the paragraphs.

In the NH group, anger was the most accurately recognized emotion of those expressed. However, it was not the most frequently chosen answer. In the NH group, the most frequently chosen answer was contentment (Table 5), which was frequently confused with excitement and fear. The emotions were most accurately recognized from the nonsense sentences and vowel [a:] (80.6% and 75.7% identified, respectively; Table 7). Perception of emotions was statistically significantly more accurately recognized in nonsense sentences than in vowel [i:] ($U = 17.900$, $p < .05$). The NH children's perception of valence was statistically significantly more accurately recognized in nonsense sentences and in vowel [a:] compared with vowel [u:] ($U = 9144.000$, $p < .01$; $U = 9288.000$, $p < .05$, respectively).

Emotion identified, Sample type, and Gender crosstabulation

Gender	Sample type					Total
	emotion identification	nonsense sentence	vowel [i:]	vowel [a:]	vowel [u:]	
Male	non-identified	16	23	17	21	77
	identified	48	41	47	43	179
Female	non-identified	12	19	18	21	70
	identified	68	61	62	59	250
Total		144	144	144	144	576

Table 7. Crosstabulation of the sample types, gender, and emotion identification, NH peers.

The results showed no statistically significant correlation between the HI children's chronological age and emotion or valence identification. The mean age for implantation of the children was three years (SD 2.4). As the participating children were implanted at very different ages (between one and nine years), the data was split into three parts — those who were implanted at 13–18 months, 19–36 months, and at 37+ months. The results showed that those

who were implanted before the age of 36 months identified emotions statistically significantly more accurately than those who were implanted after age of 36 months ($p < .05$).

PTAs were studied from the children's audiograms and compared to the results of the emotion and valence identification. It was found that the thresholds of hearing in the HI children did not correlate with the ability to discriminate between the emotions and valences.

Questionnaire

Among the HI children, the emotion identification scores were significantly connected to self-reported good mood ($U = 27408.000$, $p < .01$), alertness ($U = 30784.000$, $p < .05$), and the questions whether they liked to listen to music ($U = 8320.000$, $p < .05$), whether they were interested in dancing ($U = 26832.000$, $p < .05$), and whether it was easy for them to dance in the correct rhythm ($U = 12688.000$, $p < .01$). Valence identification scores of the HI children were significantly connected to the question whether it was easy for them to dance in the correct rhythm ($U = 13552.000$, $p < .01$).

Among the NH children, the emotions identified were significantly connected to the questions whether they played a musical instrument ($U = 13840.000$, $p < .05$), whether it was easy for them to learn a new melody ($U = 4896.000$, $p < .001$), and whether music affected their mood ($U = 6720.000$, $p = .01$). Correct valence identification by the NH children was significantly connected to the questions whether it was easy for them to learn a new melody ($U = 5184.000$, $p < .01$), whether music affected their mood ($U = 6976.000$, $p < .01$), and whether music caused them physical reactions ($U = 13040.000$, $p < .05$).

Discussion

The results showed that the HI children's chronological age had no connection to the correct identification of emotion or valence. However, the age of implantation was significantly

connected to emotion identification. This result is consistent with earlier findings (Anmyr et al., 2015; De Giacomo et al., 2013; Ketelaar & Wiefferink, 2015; Wiefferink, Rieffe, Ketelaar, De Raeve, et al., 2012), suggesting that the earlier children are implanted, the more accurate their emotion recognition. The result for the chronological age may suggest that children as young as eight-years-old (as the youngest child in this study was) are still in the process of developing an understanding of the concepts of emotions. Young children may understand single basic emotions such as joy or anger; however, the whole ‘map’ of emotions, including social emotions and how they differ from each other, is not yet wholly developed. The understanding of emotional concepts in relation to social interaction is in the developmental stage until the age of 13+ years (Saarni, 1999), and the understanding of emotional facial expressions develops even later, until the age of 16 (Lawrence, Campbell & Skuse, 2015). In the present study, statistically significant gender differences among HI children may partly imply that the average chronological age in females was one year older than that of males, which also indicates that the maturity stage of a child is a relevant criterion in emotion identification. There was not such age difference between genders among the NH children.

Fear and contentment were frequently confused with one another by the HI children due to their similar arousal level (here: low). Confusion between the negative emotion of fear and the positive emotion of contentment may cause fatal consequences in some cases, for example in a case of a threat. The NH children confused contentment with excitement due to their positive valence. Anger was the most accurately recognized emotion in both the HI and the NH groups. Anger was also the answer most frequently chosen, as the HI children often confused it with excitement. These emotions (anger and excitement) share similar arousal levels (here: a high arousal level), which may be why they were confused with one another (see Pakosz, 1982; 1983). Thus, contemplating the confusion, it seems that the HI children tended to rely mainly on the arousal level they perceived and chose their answers based on that in the

listening test. This type of confusion could be expected, as the CIs today are well adjusted to transfer the verbal content of the speech but less so for the finest spectral nuances, such as emotional variations in speech (Luo et al., 2007). In a recent study, Luo (2016) examined the effects of talker variability in emotion perception using NH subjects and CI simulations; the result suggested that, in the perception of CI simulations, the NH subjects relied mainly on the F0 range in emotion recognition and not on the finer acoustic nuances. Again, the NH subjects in Luo's study, may have relied on vocal parameters that were different from those on which listeners with HI would have relied on while detecting emotions from voice. Jiam et al. (in press), maintained that in emotion identification listeners with HI are forced to rely on cues other than frequency, i.e. intensity, tempo and duration. Mildner and Koska (2014) investigated bilaterally deaf children who had received CIs before age of two in terms of their ability to recognize auditory emotions. These children were compared with their NH peers of the same age (six to seven years old). The researchers concluded that children with HI and children with NH may rely on different acoustic cues in emotion perception. The researchers also reported that listeners with HI could more accurately recognize anger and disgust (see also Hosie, Gray, Russell, Scott & Hunter, 1998) than their NH peers. In the present study, disgust was not tested, but anger was the most accurately recognized emotion in the both groups tested, suggesting that HI children may base their assessments mainly on the arousal levels reflected in F0 and SPL. Furthermore, it seems to be an evolutionary/survival-related reaction that people interpret a loud and high voice to be a negative, rather than a positive, expression (Waaramaa-Mäki-Kulmala, 2009).

The voice quality parameters measured from the samples confirm the earlier findings of anger: the emotion of anger had the highest values for F0, F1, F2, F3 and SPL, and the shortest duration of the four emotions studied (Laukkanen et al., 1997; Waaramaa et al., 2006; Waaramaa et al., 2008; Waaramaa et al., 2014). Among the NH children, the correct

identification was related to F0, F1 and F3; hence, F3 also tended to be associated with the emotion perception (see also Waaramaa et al., 2006).

The high frequency of F4 for the positive emotion of excitement was confirmed in the present study: Excitement had the highest value for F4. On the one hand, it has been shown that positive emotions have higher values than negative emotions, especially for F4, for example, due to the articulatory setting (shortened vocal tract) in smiling (Laukkanen et al., 1997; Waaramaa et al., 2014). On the other hand, it is known that long opening time of the glottis (as it tends to be in hypofunctional voice production) may heighten the frequency of F4 (Laukkanen et al., 1997; Waaramaa et al., 2014). This phenomenon may be seen in the value of F4 in the emotion of fear, which is considered to be a negative emotion; nevertheless, F4 in the emotion of fear had a higher frequency than F4 in the positive emotion of contentment in the present study. Thus, the opening time of the glottis in the expressions of fear may be relatively long; however, the contact quotient of the glottis was not measured.

The acoustic measurements revealed that both fear and contentment were produced with hypofunctional voice quality, which indicates that the amount of energy was low in the higher frequency area. This, in turn, affects the perception and hearing of the overtones. Fear and contentment also required a longer sample duration for correct identification, as these emotions were most accurately recognized from long-sentence samples. The correlation between the frequency of F1 and the share of the emotion correctly identified points to the possibility that HI children rely on the lower-frequency energy in the voice spectrum when assessing emotions from the speech voice. This can be expected particularly when there is a relatively low amount of energy in the higher frequency area. Perhaps the perceptual meaning of the formant frequencies of F3 and F4 can be considered when developing CIs in the future.

A significant connection between mood, alertness, and emotion identification was found in the group of HI children. Volkle, Ebner, Lindenberger, and Riediger (2014) reported

similar results for older adults based on a study in which they examined the perception of emotions from faces. It is notable that earlier studies with both adult NH and HI listeners, revealed no similar correlation with respect to either auditory or from visual stimuli (Waaramaa, 2016; Waaramaa, Kukkonen, Stoltz, & Geneid, 2016). It can be then speculated that perhaps the tendency to control emotions is greater during adulthood. A person's own emotional state may, however, affect the interpretation of the perceived signal. According to Damasio (2005), some emotional state is always present in our minds. This state is called a background emotion. Izard (2007) has suggested that a continuous emotional state of one's mind functions two ways: proactively and reactively. Thus, it selectively affects the perception and the processing of the information, and may therefore have an impact on the interpretation.

Some of the questions of musical interest on the questionnaire were related to correct emotion identification. Among the HI children, the correct identification of emotions showed connections to listening to music, interest in dancing, and dancing in the correct rhythm. Hence, the identification of emotions may be related to the recognition of the rhythm of the speech. In the present study, the voice quality parameters of the expressions of anger and contentment differed substantially from each other, and there was a statistically significant difference in their duration. It is possible that the identification of these emotions may have been based on their different durations.

Correct identification of emotions was connected to the NH children's ability to learn a new melody. This result may refer to a perceptual importance of the 'flow' of the speech, its regularity, predictability, or melodicity (see Fónagy, 1981). The melodicity of the speech was found to be connected to positive vocal expressions rather than to negative expression, the latter having more voice breaks, irregularities, and unpredictability (Waaramaa, 2015).

Mildner and Koska (2014) suggested that HI children benefit from early implantation and regular auditory/oral-based therapy, as this combination enables these children to express

and perceive emotional content that is comparable with their NH peers. Today, when possible, both ears are implanted for the children with hearing loss during their first year or at the age of one in order to enhance speech and localization perception. The earlier the children are treated, the better. In the future, even vocal emotional nuances may be better recognized in accordance with development of the CIs, early treatment and support from audiologists, speech therapists, phoniaticians, and the environment.

Conclusion

The implantation age of a child tends to have an impact on the ability to identify emotions. HI females were statistically significantly more accurate than HI males in emotion and valence identification, and NH females were statistically significantly more accurate than NH males in emotion identification. No connection between the children's audiograms and correct identification was found. Musical interests and the voice quality parameters affected emotion identification, and therefore, in developing CIs, it may be beneficial to pay attention to the acoustic structures of vocal emotional expressions, especially the frequency of the formant F3, and F4 as well.

Acknowledgements

The authors extend their gratitude to the study participants, their parents and caretakers, and to Tampere University Hospital for their cooperation. Special thanks also go to the volunteer actors for providing the samples, to Tiina Syrjä, D.A., for recruiting the actors, and to Nuutti Vapaavuori, M.A., for recording and editing the samples. This study was supported by the Emil Aaltonen Foundation and the Finnish Cultural Foundation.

References

- Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monographs*, 35, 492-508. ERIC Number: ED026624
- Anmyr, L., Olsson, M., Freijd, A. & Larsson, K. (2015). Sense of coherence social networks, and mental health among children with a cochlear implant. *International Journal of Pediatric Otorhinolaryngology*, 79, 610-615. doi: 10.1016/j.ijporl.2015.02.009
- Boersma, P., Weenink, D. (2010). Praat: Downloading Praat for Windows [Computer program]. Version 5.2.18. http://www.fon.hum.uva.nl/praat/download_win.html
- Brown, J. R. & Dunn, J. (1996). Continuities in emotion understanding from three to six years. *Child Development*, 67(3), 789-802. (<http://www.jstor.org/stable/1131861>)
- Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C., Goren, A. P., Aditya, M., Kulkarni, A. M., & Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing Research*, 322, 151–62. doi:10.1016/j.heares.2014.10.003
- Chao, W-C., Lee, L-A, Liu, T-C., Tsou, Y-T. & Chan, K-C. (2015). Behavior problems in children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology*, 79, 648-653. doi: 10.1016/j.ijporl.2015.02.006
- Damasio, A. (2005). Brain and mind: From medicine to society. Videoed presentation at a conference, “Brain and mind: From medicine to society”, Barcelona, Spain. <https://www.youtube.com/watch?v=KbacW1HVZVk>, <https://www.youtube.com/watch?v=agxMmhHn5G4>
- De Giacomo, A., Graig, F., D’Elia, A., Giagnotti, F., Matera, E. & Quaranta, N. (2013). Children with cochlear implants: Cognitive skills, adaptive behaviors, social and emotional skills. *International Journal of Pediatric Otorhinolaryngology*, 77, 1975-1979. doi: 10.1016/j.ijporl.2013.09.015

- Dunn, C. C., Tyler, R. S., Oakley, S., Gantz, B. J., & Noble, W. (2008). Comparison of speech recognition and localization performance in bilateral and unilateral cochlear implant users matched on duration of deafness and age at implantation. *Ear and Hearing, 29*(3), 352–359. doi:10.1097/AUD.0b013e318167b870
- Dyck, M. J., Farrugia, C., Shochet, I. M., & Holmes-Brown, M. (2004). Emotion recognition/understanding ability in hearing or vision-impaired children: Do sounds, sights, or words make the difference? *Journal of Child Psychology and Psychiatry, 45*(4), 789–800. doi: 10.1111/j.1469-610.2004.00272.x
- Eisenberg, N., Spinrad, T. L. & Eggum, N. D. (2010). Emotion-related self-regulation and its relation to children's maladjustment. *Annual Review of Clinical Psychology, 27*(6), 495-525. doi: 10.1146/annurev.clinpsy.121208.13120
- Faulkner K. F. & Pisoni, D. B. (2013). Some observations about cochlear implants: challenges and future directions. *Neuroscience discovery, open access article.* <http://www.hoajonline.com/neuroscience/2052-6946/1/9>. doi: 10.7243/2052-6946-1-9
- Fant, G. (1970). *Acoustic theory of speech production. With calculations based on X-ray studies of Russian articulations.* (2nd ed.). The Hague: Mouton.
- Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quaterly, 50*, 274-290. doi:10.1353/mpq.2004.0018
- Fónagy I. (1981). Emotions, voice and music. In: Sundberg, J. (Ed.), *Research aspects on singing*, Royal Swedish Academy of Music, 51–79.
- Fox, N. A., Calkins, S. D. (2000). Multiple-measure approaches to the study of infant emotion. In: Lewis, M. and Haviland-Jones, J. M. (Eds.), *Handbook of emotions.* (2nd ed.). 203-219. New York: The Guilford Press.

- Gauffin, J., & Sundberg, J. (1989). Spectral correlates of glottal voice source waveform characteristics. *Journal of Speech and Hearing Research*, 32, 556–565. doi:10.1044/jshr.3203.556
- Giannantonio, S., Polonenko, M. J., Papsin, B. C., Paludetti, G. & Gordon, K. A. (2015). Experience changes how emotion in music is judged: Evidence from children listening with bilateral cochlear implants, bimodal devices, and normal hearing. An open access article. *PLoS ONE*, 1-29. doi: 10.1371/journal.pone.0136685
- Gray, C., Hosie, J., Russell, P., Scott, C. & Hunter, N. (2007). Attribution of emotions to story characteristics by severely and profoundly deaf children. *Journal of Developmental and Physical Disabilities*, 19(2), 145-159. doi:10.1007/s10882-006-9029-1.
- Hopyan, T., Manno III, F. A. M., Papsin, B. C. & Gordon, K. A. (2016). Sad and Happy emotion discrimination in music by children with cochlear implants. *Child Neuropsychology* 22(3), 366-380. doi: 10.1080/09297049.2014.992400
- Hopyan-Misakyanab TM, Gordonc KA, Dennisabd M. & Papsin BC. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychology* 15, 136-146. doi.org/10.1080/09297040802403682
- Hosie, J. A., Gray, C. D., Russell, P. A., Scott, C. & Hunter, N. (1998). The matching of facial expressions by deaf and hearing children and their production and comprehension of emotional labels. *Motivation and Emotion*, 22 (4), 293-313. doi:10.1023/A:1021352323157
- Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, 2(3), 260–280. doi: 10.1111/j.1745-6916.2007.00044.x

- Jiam, N.T., Caldwell, M., Deroche, M. L., Chatterjee, M. & Limb, C. J. (in press). Voice emotion perception and production in cochlear implant users. *Hearing Research*. Article in Press. doi: 10.1016/j.heares.2017.01.006
- Juslin, P.N., Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129, 770-814. doi: 10.1037/0033-2909.129.5.770
- Juslin, P.N., Västfäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31, 559-575. doi: 10.1017/S0140525X08005293
- Ketelaar, L. & Wiefferink, C.H. (2015). Preliminary findings on associations between moral emotions and social behaviour in young children with normal hearing and with cochlear implants. *European Child and Adolescent Psychiatry* 24, 1369-1380. doi10.1007/s00787-015-0688-2
- Lasak, J.M., Allen, P., McVay, T. & Lewis, D. (2014). Hearing loss: diagnosis and management. *Primary care* 41, (1), 19–31. doi:10.1016/j.pop.2013.10.003
- Laukkanen, A.-M., Alku, P., Airas, M., Waaramaa, T. (2008). The role of voice in the expression and perception of emotions. In: Krzysztof Izdebski (ed.) *Emotions in the human voice*. Volume I. San Diego: Plural Publishing, 171-184.
- Laukkanen, A.-M., Vilkman, E., Alku, P., & Oksanen, H. (1997). On the perception of emotions in speech: The role of voice quality. *Logopedics Phoniatrics Vocology*, 22(4), 157–168. doi: N/A
- Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology*, 6, 761. doi: 10.3389/fpsyg.2015.00761

- Litovsky, R. Y., Johnstone, P. M., Godar, S., Agrawal, S., Parkinson, A., Peters, R. & Lake, J. (2006). Bilateral cochlear implants in children: localization acuity measured with minimum audible angle. *Ear and Hearing*, 27(1), 43-59. doi: 10.1097/01.aud.0000194515.28023.4b
- Lovett, R. E., Kitterick, P. T., Hewitt, C. E., Summerfield, A. Q. (2010). Bilateral or unilateral cochlear implantation for deaf children: an observational study. *Archives of Disease in Childhood*, 95(2), 107-12. doi: 10.1136/adc.2009.160325
- Ludlow, A., Heaton, P., Rosset, D., Hills, P. & Deruelle, C. (2010). Emotion recognition in children with profound and severe deafness: Do they have a deficit in perceptual processing? *Journal of Clinical and Experimental Neuropsychology*, 32(10), 923–928. doi:10.1080/13803391003596447
- Lukkarila, P., Laukkanen, A.-M., Palo, P. (2012). Influence of the Intentional Voice Quality on the Impression of the Female Speaker. *Logopedics Phoniatrics Vocology*, 37, 158-166. doi: 10.3109/14015439.2012.687762
- Luo, X. (2016). Talker variability effects on vocal emotion recognition in acoustic and simulated electric hearing. *Journal of The Acoustical Society of America*, 140(6), 497-503. doi: 10.1121/1.4971758
- Luo, X., Fu, Q.-J., & Galvin, J. J. (2007). Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends in Amplification*, 11, 301–315. doi: 10.1177/1084713807305301
- Mildner, V., & Koska, T. (2014). Recognition and production of emotions in children with Cochlear Implants. *Clinical Linguistics & Phonetics*, published online: 2014, Informa. doi: 10.3109/02699206.2014.927000

- 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*, 449–64. doi:10.1093/deafed/enp007
- Most, T. & Michaelis, H. (2012). Auditory, visual, and auditory-visual perceptions of emotions by young children with hearing loss versus children with normal hearing. *Journal of Speech, Language and Hearing Research*, *55*(44), 1148—1162. doi:10.1044/1092—4388(2011/11-0060)
- Nakata, T., Trehub, S.E. & Kanda, Y. (2012.) Effect of cochlear implants on children’s perception and production of speech prosody. *Journal of The Acoustical Society of America*, *131*(2), 1307-1314. doi: 10.1121/1.3672697.
- Netten, A. P., Rieffe, C., Theunissen, S. C. P. M., Soede, W., Dirks, E., Briare, J. J. & Frijns, J. H. M. (2015). Low empathy in deaf and hard of hearing (pre)adolescents compared to normal hearing controls. *PLoS ONE*, *10*(4), 1-15. doi: 10.1371/journal.pone.0124102
- Neuman, A., C., Haravon, A., Sislian, N. & Waltzman, S. B. (2007). Sound-direction identification with bilateral cochlear implants. *Ear and Hearing*, *28*(1), 73-82. doi: 10.1097/01.aud.0000249910.80803.b9
- Pakosz, M. (1983). Attitudinal judgements in intonation: some evidence for a theory. *Journal of Psycholinguistic Research*, *12*(3), 311–326. doi:10.1007/BF01067673
- Pakosz, M. (1982). Intonation and attitude. *Lingua*, *56*, 153–178. doi:10.1016/0024-3841(82)90028-6
- 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*, 336–51. doi: 10.1097/AUD.0b013e318168d94d
- Saarni, C. (1999). *The development of emotional competence*. New York: The Guilford Press.

- Scheiner, E., Hammerschmidt, K., Jürgens, U., & Zwirner, P. (2006). Vocal expression of emotions in normally hearing and hearing-impaired infants. *Journal of Voice*, 20(4), 585–604. doi:10.1016/j.jvoice.2005.09.001
- Shorr, 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. doi: 10.1044/1092-4388(2008/07-0213)
- Stiles, D., J. & Nadler, L. J. (2013). Sarcasm recognition in children with hearing loss: The role of context and intonation. *Journal of Educational Audiology* 19, 3-11. doi: N/A
- Stone, M. A., Moore, B. C., Meisenbacher, K., & Derleth, R. P. (2008). Tolerable hearing aid delays v. estimation of limits for open canal fittings. *Ear and Hearing*, 29(4), 601–17. doi: 10.1097/AUD.0b013e3181734ef2
- Taumoepeau, M. & Ruffman, T. (2008). Stepping stones to other's minds: Maternal talk relates to child mental state language and emotion understanding at 15, 24 and 33 months. *Child Development*, 79, 284–302. doi:10.1111/j.1467—8624.2007.01126.x
- Toivanen, J., Waaramaa, T., Alku, P., Laukkanen, A.-M., Seppänen, T., Väyrynen, E., Airas, M. (2006). Emotions in [a:]: A perceptual and acoustic study. *Logopedics Phoniatrics Vocology* 31(1), 43-48. doi: 10.1080/14015430500293926
- van Hoesel, R. J. & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *Journal of The Acoustical Society of America*, 113(3), 1617-30. doi: 10.1121/1.1539520
- Voelkle, M. C., Ebner, N. C., Lindenberger, U., & Riediger, M. (2014). A note on age differences in mood-congruent vs. mood-incongruent emotion processing in faces. *Frontiers in Psychology*, 5(Article 635). doi: 10.3389/fpsyg.2014.00635

- Waaramaa T. (2016). Gender differences in identifying emotions from auditory and visual stimuli. *Logopedics Phoniatics Vocology*. doi: 10.1080/14015439.2016.1243725.
- Waaramaa, T. (2015). Perception of emotional nonsense sentences in China, Egypt, Estonia, Finland, Russia, Sweden and the USA. *Logopedics Phoniatics Vocology*, 40(3), 129–35. doi: 10.3109/14015439.2014.915982.
- Waaramaa, T., Alku, P., & Laukkanen, A.-M. (2006). The role of F3 in the vocal expression of emotions. *Logopedics Phoniatics Vocology*, 31(4), 153–156. doi: 10.1080/14015430500456739
- Waaramaa, T., Kukkonen, T., Stoltz, M., & Geneid, A. (2016). Hearing impairment and emotion identification from auditory and visual stimuli. *International Journal of Listening*. doi: 10.1080/10904018.2016.1250633
- Waaramaa, T., Laukkanen, A.-M., Airas, M., Alku, P. (2010). Perception of Emotional Valences and Activity Levels from Vowel Segments of Continuous Speech. *Journal of Voice* 24(1), 30-38. doi: 10.1016/j.jvoice.2008.04.004
- Waaramaa, T., Laukkanen, A.-M., Alku, P., & Väyrynen, E. (2008). Mono-pitched expression of emotions in different vowels. *Folia Phoniatica et Logopaedica*, 60(5), 249–255. doi:10.1159/000151762
- Waaramaa, T., Leisiö, T. (2013). Perception of emotionally loaded vocal expressions and its connection to responses to music. A cross-cultural investigation: Estonia, Finland, Sweden, Russia and the USA. *Frontiers in Emotion Science*. Open access publication. http://www.frontiersin.org/Emotion_Science/10.3389/fpsyg.2013.00344/abstract.
- Waaramaa, T., Palo, P., & Kankare, E. (2014). Emotions in freely varying and mono-pitched vowels, acoustic and EGG analyses. *Logopedics Phoniatics Vocology*, July 7, 1–15. doi: 10.3109/14015439.2014.934277

- Waaramaa-Mäki-Kulmala T. (2009). Emotions in Voice. Acoustic and perceptual analysis of voice quality in the vocal expression of emotions. Tampere, Finland: Tampere University Press. (Acta Universitatis Tamperensis 1399).
<https://tampub.uta.fi/handle/10024/66396/browse?value=Waaramaa-M%C3%A4ki-Kulmala%2C+Teija&type=author>
- Wake, M., Hughes, E. K., Poulakis, Z., Collins, C., & Rickards, F. W. (2004). Outcomes of children with mild-profound congenital hearing loss at 7 to 8 years: A population study. *Ear & Hearing*, 25(1), 1–8. doi:10.1097/01.AUD.0000111262.12219.2F
- Wei, C.-G., Cao, K., & Zeng, F.-G. (2004). Mandarin tone recognition in cochlear-implant subjects. *Hearing Research*, 197, 87–95. doi:10.1016/j.heares.2004.06.002
- Wiefferink, C.H., Rieffe, C., Ketelaar, L. & Frijns, J.H.M. (2012a). Predicting social functioning in children with a cochlear implant and in normal-hearing children: The role of emotion regulation. *International Journal of Pediatric Otorhinolaryngology* 76, 883-889. doi:10.1016/j.ijporl.2012.02.065.
- Wiefferink, C.H., Rieffe, C., Ketelaar, L., De Raeve, L. & Frijns, J.H.M. (2012b). Emotion understanding in deaf children with a cochlear implant. *Journal of Deaf Studies and Deaf Education*, 18(2), 175-186. doi: 10.1093/deafed/ens042