Clinical Linguistics & Phonetics, 2014; Early Online: 1–12

ISSN: 0269-9206 print / 1464-5076 online
DOI: 10.3109/02699206.2014.927000

informa
healthcare

Recognition and production of emotions in children with cochlear implants

VESNA MILDNER & TENA KOSKA

Faculty of Humanities and Social Sciences, Department of Phonetics, University of Zagreb, Zagreb, Croatia

(*Received 13 May 2014; accepted 19 May 2014*)

Abstract

The aim of this study was to examine auditory recognition and vocal production of emotions in three prelingually bilaterally profoundly deaf children aged 6–7 who received cochlear implants before age 2, and compare them with age-matched normally hearing children. No consistent advantage was found for the normally hearing participants. In both groups, sadness was recognized best and disgust was the most difficult. Confusion matrices among other emotions (anger, happiness, and fear) showed that children with and without hearing impairment may rely on different cues. Both groups of children showed that perception is superior to production. Normally hearing children were more successful in the production of sadness, happiness, and fear, but not anger or disgust. The data set is too small to draw any definite conclusions, but it seems that a combination of early implantation and regular auditory–oral-based therapy enables children with cochlear implants to process and produce emotional content comparable with children with normal hearing.

Keywords: Auditory processing of emotion, cochlear implants, children

Introduction

Emotion is frequently described in a three-dimensional space defined by arousal, valence, and control (Peters, 2006; Scherer, 2003). Processing emotional expressions is crucial to social interactions. From very early on, infants are able to detect relevant visual and auditory information in faces and voices of people around them. With experience and the maturation of sensory and perceptual systems, this eventually develops into the ability to recognize and discriminate emotions. This recognition has exhibited a robust multisensory effect that seems to be automatic (de Gelder, Stienen, & Van den Stock, 2013; Kreifelts, Wildgruber, & Ethofer, 2013; Pourois & Dhar, 2013) and apparently occurs at the early stage after stimulus onset, revealing an early perceptual (Pourois & Dhar, 2013) or a later cognitive process (de Gelder et al., 2013). Not all emotions are perceived with equal ease (Abelin & Allwood, 2000; McAllister & Hansson, 1995; Most, 1994), and their recognition may be affected by various factors, such as context (e.g. Feldman Barrett, Lindquist, & Gendron, 2007), cross-modal combinations (de Gelder et al., 2013), and method of presentation (Scherer, 2003). There is electrophysiological evidence that by

Correspondence: Professor Vesna Mildner, Ph.D., Faculty of Humanities and Social Sciences, Department of Phonetics, I. Lučića 3, 10000 Zagreb, Croatia. Tel: +385 1 6120097. Fax: +385 1 6120096. E-mail: vmildner@ffzg.hr

7 months of age infants recognize anger and happiness across modalities (Grossman, 2013). Behavioral data indicate that this may occur earlier in interaction with primary caregivers, i.e. mothers rather than strangers (Kahana-Kalman & Walker-Andrews, 2001). Depending on its severity, onset, etiology, and the extent of remediation (i.e. hearing aids and/or therapy), hearing impairment may change or disable this multisensory integration forcing the individual to rely solely or predominantly on the visual modality in communication, including emotional perception and expression. With respect to general processing preferences, contrary to adults, who prefer the visual modality (Scherer, 2003), infants and young children exhibit auditory processing preference. In children with congenital hearing impairment, this early auditory dominance (Lewkowicz, 1988; Zupan, 2013) is absent. In processing emotional content, a combination of auditory and visual cues yields best results, but some authors report that visual mode alone elicits comparable responses. Contrary to normally hearing individuals, those with hearing impairments do not benefit from the addition of the auditory cues to the visual mode (e.g. Most & Aviner, 2009 and review of the literature therein).

Similar to perception, vocal expression of emotions seems to be (at least partially) innate, universal, and even common to other primates (Abelin & Allwood, 2000; Briefer, 2012; Hammerschmidt & Jürgens, 2007; Oudeyer, 2003; Scherer, 2003). This does not mean, however, that there are no language and culture-specific expressions of and attitudes toward emotions.

Although there is no exhaustive list of parameters that reliably differentiate among various emotions, and although their characteristics depend on the source of stimuli (e.g. spontaneous versus acted speech), definition of a particular emotion (e.g. ''hot'' versus ''cold'' anger) and related arousal differences (Oudeyer, 2003; Scherer, 2003), certain vocal cues have emerged as important indicators: fundamental frequency, i.e. its mean value, range, and the rate of change; duration and (changes in) intensity. Murray and Arnott (1993) have summarized acoustic characteristics of different emotions (among other things) in terms of speech rate, pitch (average, range, and changes), and intensity. Relative to neutral emotions *anger* is characterized by slightly faster speech rate, very much higher pitch average, much wider pitch range, abrupt pitch changes, and higher intensity; *sadness* is characterized by slightly slower speech rate, slightly lower pitch average, slightly narrower pitch range with downward inflections, and lower intensity; *happiness* is characterized by faster or slower speech rate, much higher pitch average, and much wider pitch range with smooth upward inflections and higher intensity; *disgust* is characterized by very much slower speech rate, very much lower pitch average, slightly wider pitch range with downward terminal inflections, and lower intensity; and characteristics of *fear* are much faster speech rate, very much higher pitch average, and much wider pitch range with normal changes and normal intensity. This generally corresponds to van Bezooijen's (1984) descriptions, with slight variations in the description of happiness (according to him happiness is characterized by faster speech rate), disgust (characterized by smaller pitch range), and fear (characterized by low intensity). Slightly different acoustic characteristics were described for some emotions by Abelin and Allwood (2000) with respect to intensity and duration: they claim that disgust is expressed with overall highest intensity and that fear is associated with slow tempo. Scherer (2003) describes the intensity in expressing fear as higher. In physiological terms, anger, fear, and joy are related to the sympathetic nervous system activity and sadness to the parasympathetic system (Oudeyer, 2003).

In aided hearing conditions (e.g. cochlear implant or traditional hearing aids) auditory processing will be affected by the characteristics of the device. For instance, changes in pitch, which are among key cues to emotional content may be distorted or difficult to detect. In contrast, another cue, namely change in intensity, is more easily captured by hearing aids (House, 1991). Duration of individual speech segments as well as the overall duration of an utterance (including pauses), which is inversely proportional to speech rate, is another element of emotional perception and expression and it is the least affected by hearing aid characteristics. Obviously, pitch, i.e.

frequency sensitivity presents the greatest challenge to hearing aids and problems for their users. However, even in such cases auditory cues still play a role (Zupan, 2013).

Since the deviations in these three major parameters of speech are typically correlated with the severity of hearing loss, it is not surprising that production and perception of emotions present serious problems for individuals with hearing impairment. For example, Most (1994) compared the production of fear, sadness, anger, happiness, and neutral emotion by 12–14-year old children with severe hearing loss and children with normal hearing and found (mostly significant) differences in F0 range (greater for the hearing impaired), intensity (greater for the hearing impaired), and duration (with the exception of fear, longer in the hearing impaired). Consequently, with the exception of anger, the emotions expressed by the children with normal hearing were significantly more correctly perceived. McAllister and Hansson (1995) reported differences between their hearing impaired and normally hearing participants, but these were not strongly or significantly correlated with the level of hearing loss. Most and Michaelis (2012) found that children with prelingual sensory-neural hearing impairment ranging from moderate to profound exhibited lower accuracy of emotion perception than control children without hearing impairment in auditory only, visual only and auditory–visual conditions, and did not find any significant correlation with the level of hearing impairment. Hopyan-Misakyan, Gordon, Dennis, and Papsin (2009) reported poorer recognition of emotions presented auditorily but not visually by their implanted participants. Good effects of cochlear implants were found in the studies of Bat-Chava, Martin, and Kosciw (2005) and Volkova, Trehub, Schellenberg, Papsin, and Gordon (2012). Sanders (1985) reports conflicting evidence.

The aim of this research was to study the recognition and production of emotions in children who are using cochlear implants and compare their performance with that of the children without hearing impairments.

Material and methods

Participants

Nine children participated in this study. Three children (two aged 7 and one 6-year old) had prelingual bilateral profound hearing loss, they had received cochlear implants (right ear) at the age 1;08 (±4 months) and their contralateral ears remained unaided (CI group). At the time of study, the two older children had had between 5;07 and 6 years of postoperative therapy, and the younger participant had had 4;09 months of therapy using the verbotonal method (auditory–oral type). All children were mainstreamed and were attending therapy as outpatients. The first control group (NH1) was three children with normal hearing who were matched with the study group in their chronological age (two 7-year olds and one 6-year old). The second control group (NH2) was three children with normal hearing who were matched with the hearing age of the study group (between four and five). All children were otherwise healthy.

Upon obtaining parental consent, all children were tested individually in their respective preschools or at home, in a quiet room.

Test material

The emotions studied were anger, sadness, happiness, disgust, and fear. In the perception tests, two sets of stimuli were used (recorded by an actress): for the purpose of perception test 1, each child’s name was pronounced in combination with the five emotions; and for the purpose of perception test 2, nonsense 3-syllable word pairs with the stress on the first syllables (bábaba bábaba) were also combined with the five emotions. In order to make certain that the model in fact pronounced the nonsense syllables with the emotions extended, the latter material was presented to 28 naïve listeners prior to the study. All emotions were recognized as intended in more than 90% instances, which is a higher recognition rate than usually reported in literature (McAllister & Hansson, 1995; Oudeyer, 2003; Scherer, 2003), but may be explained by the limited number of stimuli presented in ideal conditions and forced choice paradigm.

Preparation

Before the test, one of the authors spent several hours (distributed over several weeks) with the children, talking about different topics in order to make them feel comfortable and relaxed. With each child, the study session began with presenting two sets of pictures depicting the five studied emotions. One set consisted of pictures of masks expressing the emotions (Brodock, 2010) and the other set was made up of drawings of a boy expressing those emotions. The children were asked to say what emotion each of the illustrations expressed. Additionally, they were asked to give examples of these emotions from their experience and were encouraged to use drawings to illustrate them. Based on these conversations, we can safely assume that all children were familiar with the concept of all five emotions. It needs to be mentioned here that disgust was the most difficult for the children. They associated it mostly with food (NH2 group exclusively so) and aesthetically unappealing situations, such as filth, picking one's nose, etc. This is hardly surprising since (in evolutionary context) this expression's underlying functional action is preventing conspecifics from eating rotten food (Scherer, 2003).

Procedure 1 – perception tests

As mentioned above, there were two perceptual tests. In perceptual test 1 (PER1), the children had to match the emotion expressed in the audio recording of their name with the appropriate picture, and in perceptual test 2 (PER2), they had to match the emotion expressed in the audio recording of the nonsense syllables with the appropriate picture. In both the tests, the recordings were presented two times in random order.

Procedure 2 – production test

In the production test, each child was presented with the audio recording of the nonsense strings (once for each emotion) and asked to repeat them. Their renditions were recorded and the recordings were analyzed in Praat (Boersma & Weenink, 2013). Unfortunately, in this test, we got usable recordings only from the CI and NH1 groups. The younger hearing subjects who were chosen to match CIs' hearing age (NH2) produced a lot of giggling and noise and did not seem to understand what they were supposed to do.

The obtained renditions were presented to 33 undergraduate phonetic students (as a part of their credit requirements) in a forced choice paradigm where they were asked to choose one of the five emotions offered that were supposedly expressed. The original recordings of the model adult female voice were included as well, and the total of 35 tokens were presented in random order via loudspeakers to three groups of 11 students each.

Results and discussion

Perceptual tests

The results of perceptual tests are summarized in Table 1. Due to the small number of participants, our results have to be treated as preliminary and the discussion is mostly qualitative. Rather than

Table 1. Responses of all children to the combination of their name and the five emotions (PER1) and the combinations of nonsense syllables and the five emotions (PER2).

		Response (%)																	
		Anger			Sadness			Happiness			Disgust			Fear			Don't know		
		CI	NH1	NH2	CI	NH1	NH2	CI	NH1	NH2	CI	NH1	NH2	CI	NH1	NH2	CI	NH1	NH2
Intended																			
PER1	Anger	67	67	50	17				33								17		50
PER2		50	67	50	17		17	17	33		17	17							33
PER1	Sadness				100	67	50								33	17			33
PER2				17	100	100	50									33			
PER1	Happiness	17	17	17				33	67	17	17			33	17	33		33	
PER2								67	83	50				33	17	33			17
PER1	Disgust		17	50	33	17	17				33	17					33	50	33
PER2			17	50	17		17				33	50		33			17	33	33
PER1	Fear							33	67	17				67	33	33			50
PER2		17						33	33	50				50	67	33			17

referring to responses as correct or incorrect, we label them as intended or unintended, depending on whether they correspond to the emotion the model wanted to produce/elicit or not. Shaded cells represent correspondence between intended emotion and response.

With respect to the type of emotion, it may be seen that disgust was the most difficult for all children. It elicited the greatest number of ''Don't know'' responses. NH2 children reported no ''disgust'' responses at all (in this case, when it was intended, or as an unintended emotion), which is along the lines of their association of this emotion exclusively with food and, hence, inability to process it in a more abstract context. The CI children identified it as intended in 33% of occurrences in both PER1 and PER2, whereas NH1 children were somewhat better on PER2 (50% as opposed to 17% on PER1). The CI children confused it most frequently with sadness (33% on PER1 and 17% on PER2) and fear (33% on PER2), but never with anger. In contrast, NH1 children never confused it with fear but there were 17% ''anger'' responses on both PER1 and PER2. They also had 17% ''sadness'' responses on PER1. Both sadness and disgust are characterized by low frequencies, slow tempo and low intensity (Murray & Arnott, 1993; van Bezooijen, 1984) and it is, therefore, not surprising that the two are most frequently confused in this study as well. The finding that there is a sort of double dissociation in the confusions with anger and fear in NH1 and CI children, respectively, is difficult to explain on the basis of such small amount of data, but it may indicate that CI children rely more on intensity (expression of fear shares low intensity with the expressions of disgust), which would be in line with reasoning in the introduction. Disgust has been found to be the most difficult to convey in other studies as well (e.g. Abelin & Allwood, 2000).

Sadness was recognized with the highest percentage of agreement with the intended emotion. CI children had 100% recognition on both PER1 and PER2 and NH1 had 67% on PER1 and 100% on PER2. NH2 recognized the intended emotion at the rate of 50% on both tests. In the instances when the intended emotion was not recognized, it was confused with fear (33% on PER1 in the NH1 group; and 17% and 33% on PER1 and PER2, respectively, in the NH2 group), which could be attributed to the similarity in low intensity (van Bezooijen, 1984). Most and Michaelis (2012) found sadness to be the easiest to recognize by hearing impaired and normally hearing individuals in the auditory mode.

Anger was recognized in 67% instances on PER1 and 50% on PER2 by the CI group, in 67% on each test by the NH1 group and in 50% on each test by the NH2 group. The CI group confused it most frequently with sadness (17% on each test), whereas the NH1 group confused it most frequently with happiness (33% on each test), which is the most frequently found confusion in the literature (e.g. Oudeyer, 2003). Of the three groups, the NH2 group had the highest percentage of ''Don't know'' responses (50% on PER1 and 33% on PER2). Whereas anger and happiness share the characteristics of all three discussed major cues (consistent with high arousal), making the NH1 confusions more or less expected, none of the cues are shared with sadness, which was the most frequent confusion in the CI group.

The results for sadness, anger, and disgust are in line with the studies that report the first two emotions as being most frequently perceived as intended and the third one as generally the most difficult to recognize correctly, a robust cross-linguistic phenomenon (e.g. Abelin & Allwood, 2000; Scherer, 2003). These general relations between ''easy'' and ''difficult'' have been reported for adult cochlear implant users as well (Peters, 2006). In the auditory mode of presentation in normally hearing participants as well as in early and late cochlear implant users (age 6 being the early versus late cutoff), but not in traditional hearing aid users, Most and Aviner (2009) found anger and sadness to be the easiest to recognize as intended and fear the most difficult.

Happiness was recognized in 33% instances on PER1 and 67% on PER2 by the CI group, in 67% on PER1 and 83% on PER2 by the NH1 group and in 17% on PER1 and 50% on PER2 by the NH2 group. It was most frequently confused with fear: equally frequently by the CI and NH2 groups – 33% on both tests; and in 17% instances on both tests by the NH1 group. If Scherer's (2003) characterization of fear as having high intensity is adopted, then these two emotions have in common all the above-described characteristics (see introduction), making this confusion less surprising. Peters (2006) also reports that her adult cochlear implant users confuse happiness with fear, whereas neither normally hearing adults nor children had any difficulty with identification or discrimination of emotional expressions, exhibiting above 90% recognition rate.

Fear was among the difficult emotions to recognize, similarly to the results reported by Most and Aviner (2009). The CI group was the most successful (67% on PER1 and 50% on PER2) followed by the NH1 group (33% on PER1 and 67% on PER2) and the NH2 (with 33% instances on both tests). Only the NH2 group had ''Don't know'' responses (50% on PER1 and 17% on PER2). In all three groups, the most frequent confusions were with happiness: 33% responses on both tests in the CI group, 67% and 33% responses on PER1 and PER2, respectively, in the NH1 group, and 17% on PER1 and 50% on PER2 in the NH2 group. These confusions are justified by the faster speech rate, higher pitch, and wider pitch range common to the two emotions (the difference being in intensity). Since in the recognition of this emotion the CI group was slightly more successful than any of the NH groups, this may again be attributed to their reliance on intensity (as mentioned above).

When the responses on the two tests are averaged it becomes obvious that groups CI and NH1 recognized the intended emotions reasonably well, i.e. the highest percentage of their responses to each emotion corresponds to that intended by the model. Anger was recognized in 59% instances by the CI group and in 67% by the NH1 group. Sadness was recognized in 100% by the CI group and in 84% by the NH1 group. Happiness was recognized in 50% instances by the CI group and in 75% by the NH1 group. Even disgust, which was the most difficult one, was chosen in the highest percentage among the five possible responses: 33% by the CI group and 34% by the NH1 group. Fear was recognized in 59% instances by the CI group and in 50% by the NH1 group. There is no consistent advantage of any of these two groups, in contrast to McAllister and Hansson (1995) who reported that their hearing impaired subjects had lower correct recognition rate of the intended emotions (comparable emotions from their study are sadness, happiness, and anger) than their counterparts with normal hearing. However, they do not provide detailed data (or information about possible hearing aids) on their hearing impaired subjects (other than age), although it may be inferred that their hearing on the better ear was around 60 dB or better.

In contrast, the NH2 group was more inconsistent and their responses were more evenly distributed across all five possible emotions, with a high ratio of ‘‘Don’t know’’ responses (see Table 1 and discussion above). Apparently they were simply too young for participation in this kind of study (Morton & Trehub, 2001; Peters, 2006).

It may be worth mentioning at this point that adult controls presented with the nonsense syllables combined with the five emotions recognized anger and sadness with 91% success, disgust and fear with 70%, and happiness with only 51%. In case of disgust, fear, and especially happiness, that is much lower than in the initial panel of 28 naïve listeners who were consulted during preparation of the stimuli, and who recognized all emotions with above 90% success rate. Obviously, even adult controls had problems in the context of numerous stimuli from various sources, as opposed to the situation where there was just one speaker in the preparatory phase. This remark is in line with the point made in the introduction that the responses are sensitive to the context of testing. However, with the exception of happiness, they were within the usually reported success rate in this type of studies: between 55% and 65% (McAllister & Hansson, 1995; Oudeyer, 2003; Scherer, 2003).

Production

The model produced anger with F0 ranging between 162 Hz and 478 Hz (peak to lowest pitch ratio: 2.95), at high intensity and took 1.3 s. NH1 children mimicked intensity and speech tempo more closely than the CI children. They repeated the pair of nonsense words correctly, pronouncing all syllables with correct word stress. Their ratio of peak to lowest pitch was between 1.72 and 2.21. There was considerable variation in the repetitions of the CI children. Only one child – L. K. – repeated all syllables, but instead of two stresses produced only one (on the first nonsense word). He also produced correct intensity and tempo. His peak to lowest pitch ratio was 1.82. The other two CI children’s intonations were flatter with a peak to lowest pitch ratio of 1.30 and 1.48. They pronounced only four or five syllables (instead of six). This difference was recognized by the listeners – the first child’s production was recognized as anger by 91% listeners, which is even higher than any of the NH1 children.

The model produced sadness with F0 ranging between 88 Hz and 319 Hz (peak to lowest pitch ratio: 3.62) at low intensity and took 2 s. All children correctly identified low intensity and slow tempo and took them as their main cues to produce this emotion. NH1 children mimicked the model more closely in all details (which is reflected in high recognition of their productions as intended (78%). CI children shortened their productions to only three or four syllables, which made their productions shorter without change in speech tempo. It is possible that due to low intensity they did not hear the unstressed parts well. They also had very little variation in pitch. As shown in Table 2, although the most frequent response to their productions was indeed sadness (61%), there were more disgust (20%) and fear (16%) responses than in case of the NH1 group.

The model produced happiness with F0 ranging between 214 Hz and 678 Hz (peak to lowest pitch ratio: 3.17) at high intensity and took 1.4 s. NH1 children mimicked this emotion very closely in all respects, and it is, therefore, somewhat surprising that they elicited only 45% ‘‘happiness’’ responses with considerable confusion with anger (32%). It is possible that the model did not convey the emotion in a most prototypical way, because only 55% of the listeners recognized model’s rendition of happiness, whereas the rest confused it with anger, fear, and disgust with almost equal frequency (between 12% and 15%). The listeners’ responses to CI children’s productions were almost evenly distributed across happiness, sadness, fear, and disgust (in that order) ranging from 24% for happiness to 19% for disgust. In this sense, we may conclude that the children in fact mimicked the intended emotion considerably well, since their renditions were confused similarly as the model's. CI children again shortened their productions and had problems with word stress, mainly producing syllables with equal stress on all syllables. Their peak to lowest pitch ratio ranged between 1.58 and 2.77, whereas that ratio for the NH1 group ranged between 2.25 and 3.28.

The model produced disgust with F0 ranging between 149 Hz and 230 Hz (peak to lowest pitch ratio: 1.5), low intensity and took 2.24 s, which is the slowest tempo of the five.

CI children again had problems with repeating all syllables – with the exception of L. K. they shortened their productions considerably (even down to two or three syllables, to less than 1 s) but adhered to slow tempo and low intensity. Here too it seems that, while low intensity and slow tempo are important cues for recognition of emotions, they are at the same time difficult to process actively. This group's peak to lowest pitch ratio ranged from 2.80 to 3.14, which is considerably higher than in the model. NH1 children were overall more successful than the CI group. In two children, the peak to lowest pitch ratio was similar to the model's (1.51 and 1.68) and, in one child, it was 3.17. They were more similar to the model in tempo as well and correctly repeated all syllables. Although the model's intended emotion was correctly identified as disgust by 70% of the listeners, even the NH1 children who acoustically appeared very similar to the model, elicited more ''sadness'' and ''happiness'' responses than ''disgust''.

The model produced fear with F0 ranging between 250 Hz and 473 Hz (peak to lowest pitch ratio: 1.89), low intensity and fast tempo (1.4 s). CI children – with the exception of L. K. – shortened their productions to one four-syllable nonsense words. This made their renditions shorter without increasing the tempo. Two children had the peak to lowest pitch ratio 1.58 and 1.60, respectively, while one child's intonation varied seemingly uncontrollably. NH1 children spoke more slowly than the model, in the two slower children, the peak to lowest pitch ratio was 1.83 and 1.89, respectively, while one child had a much higher ratio: 3.53. There were only 34% ''fear'' responses. This may be attributed to the two children who had slower tempo (in line with Murray & Arnott, 1993), since inspection of individual data (not shown here) reveals that the third child's rendition, who's tempo was comparable with the model's, was predominantly perceived as intended.

Summarizing acoustical data and observations of the CI and NH1 children suggests two possible major issues that are not new to the field of hearing impairments and seem to persist in CI children's development of speech production. First, it is possible that they do not benefit from the full intensity range, particularly at the low end, which is manifested in their problems with hearing unstressed parts of utterances. Second, they seem to have occasional problems with voice control, which results in uncontrolled variations in intonation (see also discussion in Most, 1994). As reported by other authors (e.g. House, 1991; Most & Aviner, 2009; Peters, 2006), there was considerable variation in CI group performance, but variation was found in the NH1 group as well, so we tend to attribute it to their young age. Most and Aviner (2009) also found age to be an important factor in perception of emotion, as did a number of other authors (e.g. Snow & Ertmer, 2009; Volkova et al., 2012).

In the second part of the production study, the recordings were presented to students of phonetics in a forced choice recognition test. The results of students' responses (PRO) are presented in Table 2, together with the responses of the CI and NH1 groups averaged across the two perceptual tests (PER). Shaded cells represent correspondence between intended emotion and response. Bold numerals represent highest response percentage for a particular intended emotion in each group of children. The discrepancies between the maximum numbers in the PER and PRO columns for each emotion are indicative of the differences in children's ability to perceive and produce given emotions.

It can be readily seen that sadness is the emotion that (apart from eliciting the highest correspondence between the intended and the perceived emotion) is also produced most

Table 2. Responses of adult listeners to children's productions of emotions (PRO) and averaged children's responses on perceptual tests (PER).

		Response (%)									
		Anger		Sadness		Happiness		Disgust		Fear	
		PER	PRO	PER	PRO	PER	PRO	PER	PRO	PER	PRO
Intended											
CI	Anger	**59**	**36**	17	8	9	20	9	20		16
NH1		**67**	29		2	33	**53**	9	2		14
CI	Sadness		1	**100**	**61**		2		20		16
NH1			2	**84**	**78**		5		9	17	6
CI	Happiness	9	10		24	**50**	**26**	9	19	33	20
NH1		9	32		2	**75**	**45**		4	17	16
CI	Disgust		3	25	**40**		24	**33**	21	17	12
NH1		17	12	9	**37**		30	**34**	11		9
CI	Fear	9	8		**43**	33	8		21	**59**	20
NH1			12		9	**50**	**36**		8	**50**	34

successfully (61% in the CI group and 78% in the NH1 group). Happiness is another emotion that elicited the highest percentage of intended responses of the possible five. In the NH1 group, the recognition rate was 45% and, in the CI group, it was 26%. The most frequent confusion was with anger (32%) in the NH1 group and with sadness (24%) in the CI group. Other emotions did not exhibit such congruence between perception and production. While anger was most frequently recognized as intended by both groups, the NH1 group's productions were most frequently recognized as happiness by the listeners (53%), with anger coming in second with 29% rate of recognition as intended. In this case, the CI group's productions were most frequently recognized as intended (36%), with happiness and disgust sharing the second place (20%). Just as they were difficult in perception, fear and disgust proved to be the most difficult to mimic successfully, i.e. in such a way as to convey the intended emotion to the listeners. While fear was recognized as intended by CI children in 59% of instances, their productions were most frequently perceived as sadness (43%) and recognized as intended with the rate of 20%. Just as they were divided in perception of fear (50% fear and 50% happiness responses), NH1 children produced it in such a way that it was to elicit almost equal percentage of fear and happiness responses (34% and 36%, respectively). Finally, relatively poor perception of disgust and its most frequent confusion with sadness was mirrored in children's production: listeners recognized it as sadness in 40% CI children's productions and 37% of NH1 children. Even happiness was a more frequent response to both CI and NH1 (24% and 30%, respectively) than the intended disgust (21% for CI and 11% for NH1). The relatively high rate of ''happiness'' responses to children's productions, especially when the intended emotions were so different in valence and commonly used acoustic descriptions (e.g. disgust) warrants a comment. Table 2 reveals that disgust was never recognized as happiness (compare columns PER and PRO) by any group of children. However, it was apparent that children had a lot of fun with the study and their enjoyment was difficult to conceal in their voices. This, we believe, was reflected in the listeners' responses.

There seem to be no consistent differences between the two groups of children who participated in this study. Acoustic analysis did not reveal consistently longer, more intense, or pitch-wise more variable productions in the CI children than in the NH1 children, as was the case in the Most (1994) study. With respect to listeners' success in recognizing intended emotions, we found that some emotions (e.g. anger and disgust) were better perceived as intended by the CI group and the rest were better perceived as intended by the NH1 group.

In any study of speech production where the task includes repetition after a model, it is a question to what extent the subjects' responses reflect the perceptual process/outcome, on one hand, and production abilities, on the other hand. On the basis of the data presented in Table 2, we believe we can say that the children's perceptual abilities surpass their production (which is in line with common notion that comprehension/perception is ahead of production in typical speech acquisition). In spite of uncertainties with some emotions (especially disgust), on perceptual tests, in both groups the highest percentage of responses corresponded with the intended emotions (PER columns), which indicates that children were successful in recognizing emotions. Moreover, which is very important for the evaluation of therapy, it may be seen that children with cochlear implants are comparable in all respects to children without hearing impairments. In that respect, our results are in line with the authors who found that cochlear implants have favorable effects in processing emotional content (e.g. Bat-Chava et al., 2005; Snow & Ertmer, 2009). At this time, we cannot explain why one CI child, L. K., seemed to be more successful than his two peers in mimicking the model. It is true that of the three he was implanted at the youngest age (1;04), but he was also involved in therapy the shortest time (4;09) and his pre-implant hearing status was comparable with others.

Poorer performance of the NH2 group on the perceptual test compared with the other two groups is further mirrored by their inability to understand the task in the second (production) part of the study. We attribute these difficulties to the level of their cognitive development and maturation insufficient for this kind of task. It is frequently a subject of discussion in studies of children with hearing impairment, especially since the advancements and increasing sophistication and quality in all types of hearing aids, how to treat the period before functional use of the aids. The behavior of younger children in this study, albeit more anecdotal than conclusive, speaks in favor of those who believe that both the chronological and the hearing age should be taken into consideration (e.g. Snow & Ertmer, 2009; Volkova et al., 2012). On one hand, hearing experience is crucial in developing good speech perception and production skills but, clearly, general cognitive development is an important factor as well.

Conclusions

Since, as already discussed above, the amount of data presented in this study is limited, our results and discussion must be considered preliminary. Hopefully, with a larger number of participants and the more rigorous statistical analyses they will be corroborated. Also, further studies in more natural situations may provide additional information. However, we believe we have shown that children who are profoundly deaf may be successful in mastering such intricate speech tools that are necessary to process and express the fine-grained and sophisticated nuances in communication that are conveyed by emotions, given early intervention (both surgical and rehabilitation) and regular therapy based on the auditory–oral approach.

Acknowledgements

The authors thank the children, their parents, and educators for participation in the study.

Declaration of interest

The authors report no conflict of interest. This work was supported by Grant 130-0000000-3096 of the Ministry of Science, Education and Sports of the Republic of Croatia. The authors alone are responsible for the content and writing of the paper.

References

Abelin, A., & Allwood, J. (2000). *Cross-linguistic interpretation of emotional prosody.* ISCA Archive. Retrieved December 10, 2013, from http://www.isca-speech.org/archive.

Bat-Chava, Y., Martin, D., & Kosciw, J. G. (2005). Longitudinal improvements in communication and socialization of deaf children with cochlear implants and hearing aids: Evidence from parental reports. *Journal of Child Psychology and Psychiatry*, *46*, 1287–1296.

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer* [*Computer program*]. Version 5.3.66, Retrieved June 9, 2013, from http://www.praat.org/.

Briefer, E. F. (2012). Vocal expression of emotions in mammals: Mechanisms of production and evidence. *Journal of Zoology*, *288*, 1–20.

Brodock, K. (2010). *Emotional brand – It's also important*. Retrieved August 15, 2011, from http://www.otherside-group.com/2010/07/emotional-brand-also-important.

de Gelder, B., Stienen, B. M. C., & Van den Stock, J. (2013). Emotions by ear and by eye. In P. Belin, S. Campanella, & T. Ethofer (Eds.), *Integrating face and voice in person perception* (pp. 253–268). New York: Springer.

Feldman Barrett, L., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in Cognitive Sciences*, *11*, 327–332.

Grossman, T. (2013). The early development of processing emotions in face and voice. In P. Belin, S. Campanella, & T. Ethofer (Eds.), *Integrating face and voice in person perception* (pp. 95–116). New York: Springer.

Hammerschmidt, K., & Jürgens, U. (2007). Acoustical correlates of affective prosody. *Journal of Voice*, *21*, 531–540. doi:10.1016/j.jvoice.2006.03.002.

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.

House, D. (1991). Cochlear implants and the perception of mood in speech. *British Journal of Audiology*, *26*, 198.

Kahana-Kalman, R., & Walker-Andrews, A. S. (2001). The role of person familiarity in young infants' perception of emotional expression. *Child Development*, *72*, 352–369.

Kreifelts, B., Wildgruber, D., & Ethofer, T. (2013). Audiovisual integration of emotional information from voice and face. In P. Belin, S. Campanella, & T. Ethofer (Eds.), *Integrating face and voice in person perception* (pp. 225–252). New York: Springer.

Lewkowicz, D. J. (1988). Sensory dominance in infants 1: Six-month-old infants' response do auditory–visual compounds. *Developmental Psychology*, *24*, 172–182.

McAllister, R., & Hansson, H. (1995). The perception of speakers emotions by hard of hearing children and adolescents. Retrieved December 10, 2013, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.28.3867&rep=rep1&type=pdf.

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

Most, T. (1994). Production of emotions by hearing-impaired and normal-hearing children. *Scandinavian Audiology*, *23*, 147–151.

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., & 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*, 1148–1162.

Murray, I. R., & Arnott, J. L. (1993). Towards the simulation of emotion in synthetic speech: A review of the literature of human vocal emotion. *Journal of Acoustic Society of America, 93*, 1097–1198.

Oudeyer, P. -Y. (2003). The production and recognition of emotions in speech: Features and algorithms. *International Journal of Human–Computer Studies*, *59*, 157–183.

Peters, K. P. (2006). Emotion perception in speech: Discrimination, identification, and the effects of talker and sentence variability. Independent studies and capstones. Paper 121. Program in audiology and communication sciences, Washington University School of Medicine. Retrieved December 10, from http://digitalcommons.wustl.edu/pacs_capstones/121.

Pourois, G., & Dhar, M. (2013). Integration of face and voice during emotion perception: Is there anything gained for the perceptual system beyond stimulus modality redundancy? In P. Belin, S. Campanella, & T. Ethofer (Eds.), *Integrating face and voice in person perception* (pp. 181–206). New York: Springer.

Sanders, G. (1985). The perception and decoding of expressive emotional information by hearing and hearing-impaired children. *Early Child Development and Care*, *21*, 11–26.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*, 227–256.

Snow, D., & Ertmer, D. (2009). The development of intonation in young children with cochlear implants: A preliminary study of the influence of age at implantation and length of implant experience. *Clinical Linguistics & Phonetics*, *23*, 665–679.

van Bezooijen, R. (1984). *Characteristics and recognizability of vocal expressions of emotion*. Dordrecht: Foris Publications.

Volkova, A., Trehub, S. E., Schellenberg, E. G., Papsin, B. C., & Gordon, K. A. (2012). Children with bilateral cochlear implants identify emotion in speech and music. *Cochlear Implants International*, *14*, 80–91.

Zupan, B. (2013). The role of audition in audiovisual perception of speech and emotion in children with hearing loss. In P. Belin, S. Campanella, & T. Ethofer (Eds.), *Integrating face and voice in person perception* (pp. 299–324). New York: Springer.