Developmental Links Between Speech Perception in Noise, Singing, and Cortical Processing of Music in Children with Cochlear Implants

RITVA TORPPA
University of Helsinki, Helsinki, Finland

ANDREW FAULKNER
University College London, London, United Kingdom

TEIJA KUJALA, MINNA HUOTILAINEN, &
JARI LIPSANEN
University of Helsinki, Helsinki, Finland

THE PERCEPTION OF SPEECH IN NOISE IS challenging for children with cochlear implants (CIs). Singing and musical instrument playing have been associated with improved auditory skills in normal-hearing (NH) children. Therefore, we assessed how children with CIs who sing informally develop in the perception of speech in noise compared to those who do not. We also sought evidence of links of speech perception in noise with MMN and P3a brain responses to musical sounds and studied effects of age and changes over a 14–17 month time period in the speech-in-noise performance of children with CIs. Compared to the NH group, the entire CI group was less tolerant of noise in speech perception, but both groups improved similarly. The CI singing group showed better speech-in-noise perception than the CI non-singing group. The perception of speech in noise in children with CIs was associated with the amplitude of MMN to a change of sound from piano to cymbal, and in the CI singing group only, with earlier P3a for changes in timbre. While our results cannot address causality, they suggest that singing and musical instrument playing may have a potential to enhance the perception of speech in noise in children with CIs.

Background noise is pervasive in the everyday environments of children (Bradley & Sato, 2008; Fu & Galvin, 2008). This can severely affect their learning, especially because the perception of speech in noisy situations matures fairly late, and they reach adult-level performance in speech perception in noise at 11–15 years of age (Baker, Buss, Jacks, Taylor, & Leibold, 2014; Fallon, Trehub, & Schneider, 2000; Hall, Grose, Buss, & Dev, 2002; Stuart, 2005). Background noise is particularly disruptive for congenitally deaf children hearing with cochlear implants (CIs). With CIs, spectral detail is largely lost (Moore, 2003), leading to poorer speech-in-noise performance than in normal-hearing (NH) listeners (adults: Friesen, Shannon, Baskent, & Wang, 2001; Fu & Nogaki, 2005 children: Asp et al., 2012; Caldwell & Nittrouer, 2013; Geers, Davidson, Uchanski, & Nicholas, 2013; Mishra, Boddypally, & Rayapati, 2015). This underlines the need to improve this perceptual skill in children with CIs.

Musical activities might be a way to achieve this goal (Patel, 2014; Shahin, 2011). This suggestion is supported by the finding that NH adult musicians and children given music training are advantaged over other NH listeners in perceiving speech in noise (Coffey, Mogilever, & Zatorre, 2017; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark, Strait, & Kraus, 2011; Strait, Parbery-Clark, Hittner, & Kraus, 2012), and that speech-in-noise performance improves with duration of musical activity (adults: Ruggles, Freyman, & Oxenham, 2014; children: Strait et al., 2012). Further, there is recent longitudinal evidence of enhancement of NH children's speech-in-noise performance with long-term musical education (Slater et al., 2015). This effect might be connected to the beneficial effects of musical activities on speech segmentation (François, Chobert, Besson, & Schön, 2013). One cue for speech segmentation is the pattern of prosodic stress, which forms the rhythm of speech and assists in the identification of the beginning of words (Jusczyk, 1999); adults with deficits in hearing have been shown to use this so-called metrical segmentation strategy in the presence of background noise (Woodfield & Ackeroyd, 2010). Good ability to detect word stress patterns requires good

perception of vocal pitch and intensity (Torppa, Faulkner, et al., 2014), and is also linked to good perception of rhythm (Hausen, Torppa, Salmela, Vainio & Särkämö, 2013). These abilities all appear to be better in those children who participate in musical activities (rhythm: NH, Flaugnacco et al., 2015; stress, vocal pitch and intensity: CIs, Torppa, Faulkner, et al., 2014).

The musical activities in the studies above involved the playing of musical instruments and singing. It is plausible that singing would be more beneficial for improving speech perception in noise than instrument playing because songs contain not only musical elements of melody, meter, and predictable rhythm, but also lyrics (Patel, 2014). Thus, the processing of musical features when singing is likely to be closely linked to and assists phonological processing to a greater degree than when playing or listening to instrumental music. The perception of speech relies partially on cortical oscillations entrained to temporal modulation patterns at different timescales. One important timescale is that of speech rhythm (Leong, Kalashnikova, Burnham, & Goswami, 2017; Mattys, 1997). Such entrainment is also found to the meter of music (Cason & Schön, 2012). These oscillations lead to enhanced attention at the moments when the strong musical beat or speech stress is expected (Bolger, Coull, & Schön, 2014; Schön & Tillmann, 2015), and assist in binding information together in the final percept (Leong et al., 2017). The meter of music in song is more predictable than the meter of speech, and in the lyrics of song, stressed syllables are typically in strong metrical positions, leading to particularly strong changes in brain activity, to enhanced lexical decision, and enhanced speech processing (Gordon, Magne, & Large, 2011; Schön & Tillmann, 2015). In a multisensory context, for example, when fingers are tapped in time with predictable stressed syllables, there is even better sensitivity to changes in speech during these predictable syllables (Falk & Dalla Bella, 2016). As in "clear speech" (Picheny, Durlach, & Braida, 1985), motor movements during singing directly realize stressed syllables (Falk & Dalla Bella, 2016). Further, slow word rate and repetition in song gives more time to process words (Patel, 2014). It is plausible then that singing aloud can assist the development of speech perception in noise for children using CIs.

The review above emphasizes the role of rhythm and meter in the perception of speech, and in line with this, perception of speech in noise has been shown to be linked to behavioral discrimination of rhythm (Slater & Kraus, 2016). The review above also points out that fast attention shift towards sounds such as predictable stressed syllables, and well-functioning networks for attention processing are important. Correspondingly, it has been shown that attention in general is important in the perception of degraded speech; for example, with CIs and in background noise (Adank, Davis, & Hagoort, 2012; Beer, Kronenberger, & Pisoni, 2011; Houston & Bergeson, 2014; Strait et al., 2012; Wild et al., 2012). Notably, we have found that children with CIs who sang regularly, although without formal singing training, were advantaged over others who did not sing in both rhythmic and attentional abilities. Those who sang were better in the production of rhythms of song (a Finnish version of "Twinkle Twinkle Little Star"), and they developed more in the P3a event-related potential (ERP) brain responses associated with attention (Torppa, Huotilainen, Leminen, Lipsanen, & Tervaniemi, 2014). P3a reflects attention shift towards sound changes, and also activity in the neural networks for auditory attention (Alho et al., 1998; Horvath, Winkler, & Bendixen, 2008; for a more detailed explanation of P3a, see below). Comparable findings in NH children showed that informal musical activity including singing was linked to more advanced auditory attention functions (Putkinen, Tervaniemi, & Huotilainen, 2013), echoing findings in children receiving singing and musical instrument training (Strait et al., 2012). Furthermore, we have previously found that those children with CIs who sing frequently have also been sung to more frequently by their parents early in their life (Torppa, 2015). Parental singing may not only encourage the children to sing themselves, but may play a direct role in the linkage between child's own singing and P3a, since parental singing is known to regulate the attention of young children (Rock, Trainor, & Addison, 1999).

While formal singing training might be an ideal example of singing activity, this is rare and difficult to organize in children with CIs. However, informal music engagement is relatively common (at least in Finland), so is a feasible basis for dividing the child CI participants according to exposure to music. Correlational evidence associating singing with perception of speech in noise would be a basis to propose a randomized intervention study that addressed the issue of causality of singing in children with CIs. Therefore, the current study determined whether engagement in singing at home was associated with speech perception in noise.

We also studied whether neural discrimination of changes of musical instrument sounds, and attention shift toward such changes, show links to the perception of speech in noise in children using CIs. Such links would be consistent with common or shared processing of acoustic features between music and speech, as is evident in NH listeners for pitch, duration, and stimulus

onset (the last being necessary for encoding voice onset, attack, and rhythm) (for a review, Besson, Chobert, & Marie, 2011; Hausen et al., 2013). The shared processing of acoustic features may also explain the effects of musical activities to perception of speech in noise (Patel, 2014; Slater et al., 2015). In particular, we might expect that children who directly manipulate the sounds of musical instruments through playing them become more acutely aware of such sound differences than those who merely listen to instruments. Hence, as Patel as argued, experience of playing instruments may enhance the brain's capacity to process speech. Thus, if we find links between processing of changes in musical instrument tones and the perception of speech in noise, this would suggest that activities that enhance music perception, and perhaps playing musical instruments in particular, have the potential to improve speech-in-noise perception of children with CIs (see also Patel, 2014).

There is some evidence for the common or shared processing between music and speech in noise in adult CI recipients. Here, the perception of speech in noise is linked to perception of the direction of pitch change and melody recognition for piano-like tones, and also to musical instrument (timbre) perception (Won, Drennan, Kang, & Rubinstein, 2010). Moreover, for adults with CIs, there is some evidence linking speech perception in noise to pre-attentive discrimination of changes in pitch and intensity (Sandmann et al., 2010). However, such links may not develop in prelingually deaf children with CIs, whose brain development may be affected by the period of deafness before implantation, and after that, from the degraded sound signal from a CI (for an overview, Torppa, 2015). To address this, this study examined the connections of the perception of speech in noise to music processing, the latter assessed with mismatch negativity (MMN) and with P3a brain responses to changes in the intensity, pitch, and timbral quality of musical instrument tones.

The MMN reflects the auditory system's response to violations of regularities in sound input (Kujala & Näätänen, 2010; Winkler, Denham, & Nelken, 2009). The MMN becomes stronger in amplitude and shorter in latency with increasing physical difference between the standard stimulus and a deviant sound. Further, the MMN reflects discrimination accuracy both in NH listeners (Kujala & Näätänen, 2010; Näätänen, Paavilainen, Rinne, & Alho, 2007) and in CI users (Lonka et al., 2004; Ponton et al., 2000; for a review, see Näätänen, Petersen, Torppa, Lonka, & Vuust, 2017). The P3a typically follows the MMN if the change in the sounds is clearly detectable or significant for the listener and if the attention shifts toward the sound change. Furthermore,

the P3a is thought to reflect the reconfiguration of a brain network involved in updating task set information for goal-directed action selection (Escera & Corral, 2007; Horvath et al., 2008; P3a in CI recipients, Kileny, Boerst & Zwolan, 1997; Kelly, Purdy, & Thorne, 2005; Nager et al., 2007; Torppa et al., 2012; Torppa, Huotilainen, et al. 2014). Like the MMN, in NH listeners, P3a amplitude increases with increasing physical difference between the deviant and standard (Wetzel, Widmann, Berti, & Schröger, 2006; Winkler, Tervaniemi, Schröger, Wolff, & Näätänen, 1998) and with auditory training (Uther, Kujala, Huotilainen, Shtyrov, & Näätänen, 2006). In addition, P3a for speech is both larger and shorter in latency in children with CIs who show relatively good speech recognition than in other CI children (Kileny et al., 1997). Besides the ERP measures, we also looked at the connections of perception of speech in noise to behavioral discrimination of pitch (f_0) and intensity in synthesized speech. With this, we assessed whether the links of pitch and intensity processing to the perception of speech in noise are similar for speech and musical stimuli.

Previously, we found larger and shorter latency P3a, but smaller MMN responses (presumably due to the early P3a overlapping the MMN), for the CI singing than CI non-singing group, and found that brain responses developed differently in these groups (Torppa, Huotilainen, et al., 2014). As these results indicated differences between the CI singing vs. CI non-singing groups in the development of cortical processing of musical sound changes, this group factor was added to the current statistical analyses when links of perception of speech in noise to MMN and P3a brain responses were studied.

Additionally, we compared the development of speech-in-noise perception of child CI participants to the development of NH peers. While NH children show age-related improvements in speech-in-noise perception (Bradley & Sato, 2008; Hall et al., 2002; Nittrouer & Boothroyd, 1990; Stuart, 2005), according to our best knowledge, comparable effects of age have not been found for children with CIs (Jung et al., 2012; Looi & Radford, 2011; Ruffin, Kronenberger, Colson, Henning, & Pisoni, 2013). However, none of these previous studies has directly compared age-related development between children with CIs and NH. Therefore, this study has a NH control group, allowing statistical comparisons of age related development between children with CIs and NH. Further, the present study entailed two measurement points, 14 to 17 months apart, and it is the first to examine changes in speech-in-noise performance over time in children with CIs.

TABLE 1. *Details of the Participants*

| Children with cochlear implants | | | | | | | Normal-hearing children | | |
ID	Gender	Age at T1	Etiology	Age at switch-on (months)	CI use prior T1 (months)	CI processor type[4]	ID	Gender	Age at T1
CIs 01	F	5y 11m	U	18	53	NF	NH02	F	7y 11m
CIs 03	F	9y 2m	U	32	77	MT	NH 03	M	4y 6m
CIs 04	F	7y 10m	U	25	69	MT	NH 04	F	8y 2m
CIns 09	F	7y 4m	C	19	69	MO	NH 05	M	10y 0m
CIs 13	M	5y 5m	U	18	47	NE	NH 06	F	5y 8m
CIs 14	F	4y 4m	U	18	34	NF	NH 07	F	6y 9m
CIs 15	M	5y 1m	C	17	44	NE	NH 08	F	5y 7m
CIns 16	M	7y 2m	C	25	61	NF	NH 09	F	4y 6m
CIns 17	M	9y 4m	U	19	93	NF	NH 10	M	4y 0m
CIns 18	F	12y 1m	U	27	118	NF	NH 11	F	5y 6m
CIns 19	F	7y 5m	U	29	60	NE	NH 13	M	5y 0m
CIs 20	M	5y 8m	U	20	48	NF	NH 14	M	4y 6m
CIs 21	F	5y 7m	C	19	48	NF	NH 15	F	12y 0m
CIs 22	F	7y 1m	U	21	48	NE	NH 16	M	8y 5m
CIns 23	F	7y 10m	U	18	76	MT	NH 17	M	9y 8m
CIns 24	F	4y 2m	C	14	36	NF	NH 18	M	6y 9m
CIs 26	M	4y 2m	C	20	30	NF	NH 19	F	7y 0m
CIns 27	F	4y 2m	C	13	37	NF	NH 20	F	4y 6m
CIs 28	M	6y 2m	U	22	52	NF	NH 21	M	6y 5m
CIns 29	M	8y 7m	C	37	66	NF	NH 22	M	6y 11m
CIs 30	M	6y 7m	C	25	54	NF	NH 23	M	5y 5m
							NH 30	F	11y 2m
N CI = 21 N CIs = 12 N CIns = 9	N M+F= 9+12	Mean = 6y 7m	N U = 12 N C = 9	Mean = 21.7	Mean = 58.1	N NF = 13 N NE = 4 N MO = 1 N MT = 3	N NH = 22	N M+F = 11+11	Mean = 6y 9m

Note: ID = Identification number. CI = a child with cochlear implant (CI). NH = normal-hearing (NH) child. CIs = CI singing group, where children sang regularly. CIns = CI non-singing group, where children did not sing regularly. N = number. F = female, M = male. T1, T2 = the first and second time point of measurements. y = years, m = months. R = right, L = left. U = unknown, C = Connexin 26. NF = Nucleus Freedom, implant type CIC4 (coding strategy: ACE). NE = Nucleus ESPrit 3G, implant type CIC3 (coding strategy: ACE). MT = Medel Tempo + (coding strategy: CIS). MO = Medel Opus 2 (coding strategy: CIS).

This is essential, since while development over age is generally assessed between individuals, development over time reflects development within individual participants. Thus, the latter is unaffected by individual differences, and is thus a more reliable measure of development.

Based on the above considerations, we tested the following hypotheses: 1) compared to their NH peers, children with CIs will perform more poorly and show slower development of performance over age and time in the perception of speech in noise; 2) children with CIs who sing regularly at home will be advantaged in the perception of speech in noise compared to those who do not; 3) in children with CIs, performance in the perception of speech in noise will correlate with a) ERP measures of discrimination of pitch, intensity, and timbre for musical sounds, and b) behavioral discrimination of pitch and intensity in synthesized speech.

Method

PARTICIPANTS

Twenty-one unilaterally implanted children participated (ages 4–13 years) (Table 1; see also Torppa, Huotilainen, et al., 2014). All had been implanted prior to the age of 3:1 (years:months), had full insertion of the electrode array, and had more than 6 CI electrodes in use. Their hearing thresholds in the unimplanted ear were so high (in excess of 50 dB at 250 Hz, 60 dB at 500 Hz, and 70 dB at 1000 Hz) that they could not benefit from residual hearing in the measurements. They had been using their implants for at least 30 months prior to the first measurements (Table 1). None had any diagnosed additional developmental or language-related problems, and all attended typical day care or school and communicated with spoken language.

Twenty-two normal-hearing (NH) children served as a control group for the perception of speech in background noise. The NH group was matched to the CI group by age, gender, handedness, and social and musical background (Torppa, Huotilainen, et al., 2014; Table 1). None of the NH children had any diagnosed developmental or language-related problems. Their hearing had been screened and found to be normal at regular check-ups in child welfare clinics.

All participants were monolingual with Finnish as their native language. None were children of musicians. The parents gave their signed informed consent and the children gave verbal consent. The study was carried out in accordance with the Declaration of Helsinki and all procedures were approved by the ethical committees of the participating hospitals (University Hospitals of Helsinki, Kuopio, Tampere, and Turku).

Characteristics of the CI groups. We grouped the children with CIs into either a *CI singing* or a *CI non-singing group* according to their everyday habits of singing both before and during the study as reported in a parental questionnaire (Torppa, Faulkner, et al., 2014; Torppa, Huotilainen, et al., 2014). The questionnaire was administered twice, at the same times as the first and second test points of the study (T1, T2). Because systematic formal training in singing is rare for implanted children, this division was based on the regularity of informal singing at home. Since some questions were retrospective, we could not expect that parents would be able to report in detail on singing behavior, so they were simply asked how often they had heard the child sing (to them, to siblings, or alone). Twelve of the children with CIs were reported to have sung at home at least once a week before T1 and between T1 and T2, and were placed in the *CI singing group*. These children sang at home on average five times per week before T1 and 4 times per week between T1 and T2. Nine children with CIs sang less than once in a week or not at all and were placed in the *CI non-singing group* (Table 1).

According to analysis of variance (ANOVA), the CI groups did not differ significantly in age. ANCOVA (controlling for age) confirmed that the groups did not differ in frequency of other activities with parents (sports, handicraft, reading, etc.), or in the frequency of speech and language therapy. Nor did they differ in musical background before T1 or between T1 and T2, as assessed by a cluster analysis of questionnaire responses. Relevant dimensions extracted by this analysis were as follows: A/ *music at home*, measured according to frequency of siblings and parents singing or playing instruments with the child and whether or not the child played instruments him/herself; B/ frequency of the child listening to or watching children's music videos, DVDs or CDs, and the child's response to these; D/ frequency and duration of music lessons at school or in daycare; E/ time spent in musical activities or dancing lessons (before T1). Further, ANCOVA (controlling for age) confirmed the groups did not differ in respect to clinical records of hearing age, age at implantation, aided thresholds using the CI, or CI fitting parameters. Chi-square confirmed that the CI groups did not differ in the attendance at supervised musical or other activities outside of the home, socioeconomic background (assessed from the educational level of the parents and parental income), device type, gender, or etiology (see Table 1).

However, as reported earlier, a more detailed analysis of the aspects included in Cluster A (music at home) showed that the more the parents sang for the CI children prior to participation in this study (including the first year after implantation), the more the CI children sang by themselves. This shows also that parents sang more often for the CI singing group than for the CI non-singing group. None of the other factors falling into this cluster were related to the CI children's singing (Torppa, 2015). Moreover, we also previously found that, compared to the CI non-singing group, the CI singing group produced the rhythm of song more accurately, showed more change over time in the P3a ERP responses to a rhythmically predictable stimulus, and had significantly larger and shorter latency P3a responses for changes in pitch and timbre (Torppa, Huotilainen, et al., 2014).

PROCEDURE

Two measurement points (T1, T2) were used in order to analyze development in perception of speech in noise and connections between this and the development of ERP responses over a 14 to 17 month time period. A general overview of the experimental design is given in Table 2.

Speech-in-noise testing was done with a laptop and two loudspeakers. In the NH group, the testing took place either in an acoustically isolated and dampened room, or if the family was not able to travel to the laboratory, in a quiet room in the home of the participant. An experimenter traveled to the home of the participant, where she calibrated the test and took care that the procedure was always the same, the room acoustics were suitable for testing, and the environment was sufficiently quiet. Home testing was important for ensuring the continued participation of parents and children. For the CI group, tasks were always performed in an

TABLE 2. *Experimental Design*

| | | Duration (min) | NH group | CI group | |
				CI singing group	CI non-singing group
Behavioral experiments	Perception of speech in noise	15	T1, T2	T1, T2	T1, T2
	Discrimination of pitch and intensity	15	Not reported	T1, T2	T1, T2
ERP experiment	MMN, P3a for changes in pitch, timbre, intensity[1]	35	Not reported	T1, T2	T1, T2

CI singing group = Children with CIs who sang at home regularly. CI non-singing group = Children with CIs who sang at home less than weekly. T1, T2 = Two time points of measurements. [1]There were 3 degrees of change; only the statistically significant ERPs to change at T1 and/or T2 were included in testing hypotheses.

acoustically isolated and dampened room. During all experiments, the children with CIs used the everyday settings of their CI, without any acoustic hearing aid. All CI settings, including volume and sensitivity levels, were adjusted to the clinically recommended values.

For the ERP experiment (reported here only for the CI group), the stimuli were presented through two loudspeakers (OWI-202; OWI Inc. CA., USA) at 100 cm distance from the participant's ears. For behavioral tests, a pair of powered speakers (Edirol MA-15D) were each 70 cm from the subject. In all experiments, the loudspeakers were placed at a 45° angle to each side of the participant.

All stimuli were presented at a comfortable level, which was on average 60 dB for the NH group and 70 dB SPL for the CI group. For one CI child in the ERP experiment the level had to be lowered to 65 dB SPL at T1 because the standard 70 dB SPL was uncomfortable for her. The CI children watched a soundless video while the ERP responses were being collected.

Duration of the ERP experimental session was approximately 75 min including the placement and removal of the EEG cap. The duration of the behavioral experimental session was approximately 45 min for children with CIs and 30 min for children with NH, depending on the responsiveness of the child. The family had the option of dividing the experimental sessions into two days. There were also breaks and the children were provided food or juice and biscuits according to child's own choice.

BEHAVIORAL MEASUREMENTS
Speech reception threshold (SRT_{75}) in noise (for CI and NH groups). A Finnish translation of the Matrix sentence test (Tyler & Holstad, 1987) was recorded by an adult female native speaker of Finnish. The target sentences were based on a picture matrix containing 4 x 4 pictures, each of the pictures presenting one word in the target sentence (for example, in Finnish: /isä näkee vihreän auton/; /äiti ostaa punaisen pyörän/, in English, "father sees a green car"; "mother buys a red bike."

The children were required to point to the target pictures representing the words they heard, or to repeat the sentences as they heard them. The background noise was a steady speech-spectrum noise with male-weighted spectrum (ICRA-noise 1; Dreschler, Verschuure, Ludvigsen, & Westermann, 2001). The noise level was varied adaptively to find the signal-to-noise ratio (SNR) resulting in performance of 75% words correct, referred to subsequently as the speech reception threshold (SRT_{75}). On any trial, a score of four words correct led to an increased noise level, a score of two or fewer words correct led to a decrease of noise level, while SNR remained the same after a score of three words correct. The step size was 6 dB before the first reversal, 4 dB up to the second reversal, and 2 dB thereafter. The SRT_{75} was recorded as the average of the SNR from the final 8 reversals. If the child could not repeat any words, one repetition of the sentence was allowed to motivate the children to continue. In this event, the responses to the second sentence were used to control the staircase.

The children were familiarized with the pictures and vocabulary as well as the procedure with a live voice presentation of the sentences before the actual experiments, so that they did not begin before the experimenter was convinced that the child understood the task.

Discrimination of pitch and intensity in synthesized speech (for CI groups only). An adaptive two-interval same-different paradigm was used to assess behavioral discrimination thresholds for pitch and intensity change in a speech context (O'Halpin, 2010; Torppa, Faulkner, et al., 2014). Each trial included two synthetic speech bisyllables. These had either the same ("TAta"/"TAta") or different ("TAta"/ "taTA") stress patterns, with the stressed syllable marked by either a rise-and-fall of f_0 alone (see Figure 1) or a raised intensity alone. The
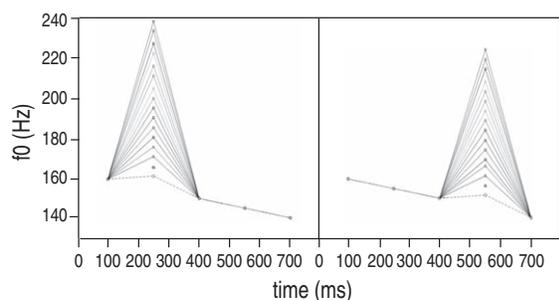
FIGURE 1. Example f_0 contours for the pitch (f_0) discrimination task (160-Hz baseline) (from Torppa, Faulkner, et al., 2014). The figure has been reprinted with the permission from Informa Healthcare.

duration of each syllable was always 300 ms and there was no temporal gap between the two syllables. The only cue was either a change in f_0 or intensity. A continuum of synthesized stimuli differing in either peak f_0 or intensity was generated using the KLATTSYN-88 software synthesizer (Klatt, 1980) and the Speech Filing System Tools for Speech Research (n.d.).

In the intensity discrimination experiment, the intensity varied between syllables by 1 to 15 dB. When pitch (f_0) was varied, the f_0 contour in the first or second syllable comprised a linear rise from syllable onset to the temporal mid-point, and continued with a linear fall from this mid-point to the end of the syllable (Figure 1). The f_0 rise-and-fall began from a baseline of 160 Hz (adult female f_0 range) or from 295 Hz (child f_0 range). The peak f_0 was higher than at onset according to 48 equally spaced multiplicative factors from 1.01 to 1.84. To make the f_0 contours more natural, a decline in f_0 was also introduced, as shown in Figure 1. For this, a linear fall in f_0 was added, such that the f_0 at syllable offset was 94% of that at onset (see Figure 1), and the onset f_0 of the second syllable was set at the same value as the offset f_0 of the first.

In these discrimination experiments, the adaptive procedure used a two-up one-down staircase that estimated the 71% correct discrimination threshold (Levitt, 1971). Repetitions of stimuli were not allowed. The CI children pointed to a picture representing either "same" or "different" or responded orally, depending on their own choice. Here also the experimenter registered the answers. As previously, participants were familiarized with the pictures as well as the procedure using live voice before the actual measurements, which began when the experimenter was convinced that the child had understood the task. To keep the interest of the child, pictorial feedback indicated whether or not the response was correct.

Because partial correlation analysis controlling for age showed that the f_0 discrimination thresholds for the two different baseline f_0 values were strongly correlated (at T1, $r_p = .91$, $p < .001$; T2, $r_p = .65$, $p = .002$), the thresholds were averaged over the female and child f_0 baselines for further analyses (Torppa, Faulkner, et al., 2014).

DETAILS OF THE ERP-EXPERIMENT (FOR CI GROUPS ONLY)

Stimuli. The stimuli for the ERP experiment were the same as in Torppa and colleagues (2012; Torppa, Huotilainen, et al., 2014). Piano, cembalo, cymbal, and violin sounds from the McGill University Master Samples DVD (Opolko & Wapnick, 2006) were cut to the desired duration and normalized in average intensity with Adobe Audition 2.0 (Adobe Systems Inc., San Jose, USA). The standard was a piano tone at 295 Hz and of 200 ms duration including a 20 ms offset ramp. The deviants always had three different degrees of change. Pitch (f_0) deviants were piano tones at 312, 351 and 441 Hz (1, 3, and 7 semitone changes). Both increments and decrements of intensity were included of 3, 6, and 9 dB. There were also three musical instrument (timbre) deviants with the sound being either a cembalo, violin, or cymbal (Figure 2). The timbre deviants matched the standard in average intensity and duration, and in f_0 for the cembalo and violin sounds (see Figure 2). Other deviants involving duration changes and gaps were also presented (see Torppa et al., 2012) but are not reported here. The multi-feature paradigm was used, in which the standard and deviant stimuli alternate (see Figure 3). Each stimulus sequence comprised 4500 stimuli, of which half were standards. Each deviant had a probability of .028, and was presented 125 times, with the order of deviants being random. The stimulus onset asynchrony (SOA) was always 480 ms. The duration of the ERP measurements, excluding fitting and removal of electrodes, was approximately 35 min.

EEG recording: Data preprocessing and data analysis. The procedure was similar to Torppa, Huotilainen, et al. (2014). EEG data were acquired from a 64-channel Biosemi cap with active electrodes and ActiveTwo mk1 amplifier (sampling rate of 512 Hz, on-line low-pass filtering at 102.4 Hz). Electrodes at the left and right mastoids were also used, and horizontal and vertical electro-oculograms were recorded to monitor eye movement artifacts. After recording, the data were re-referenced to an electrode placed at the nose tip. Data analyses were performed with EEGLAB 8 (Delorme & Makeig, 2004) after downsampling to 256 Hz and high-

(a) 295 Hz (standard) and 312 Hz tones (small change)   295 Hz (standard) and 351 Hz tones (medium change)   295 Hz (standard) and 441 Hz tones (large change)   (b)

Piano (standard) and cembalo tones (small change)   Piano (standard) and violin tones (medium change)   Piano (standard) and cymbal tones (large change)
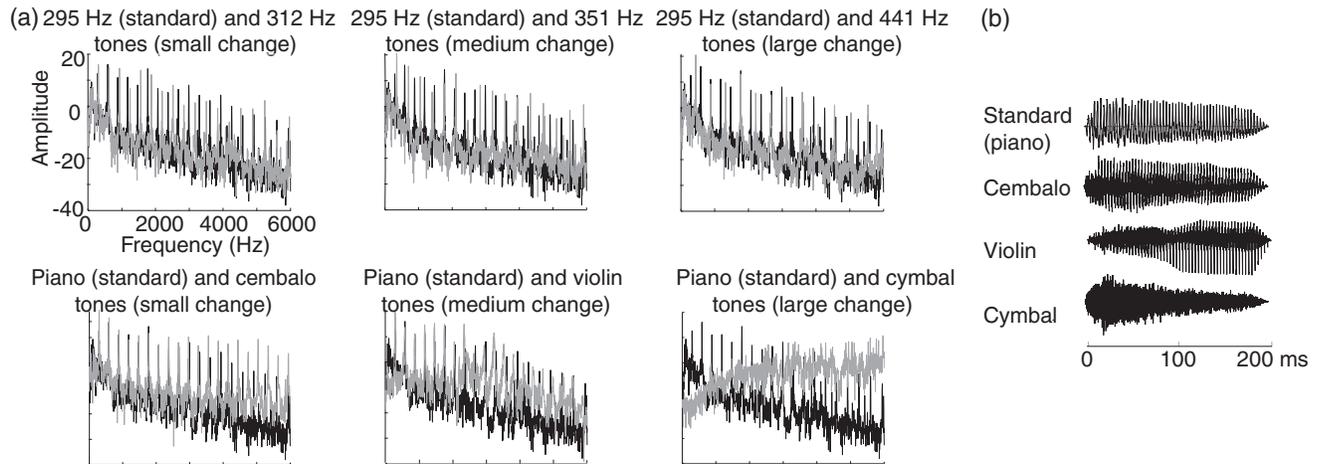
FIGURE 2. (a) Frequency spectra of the standard tone (black) in comparison to pitch and music instrument deviants (gray) (from Torppa et al., 2012). (b) Amplitude envelopes of the standard piano tone and the music instrument deviants (from Torppa et al., 2012). The figures have been reprinted with permission from Elsevier.

FIGURE 3. Visualization of the stimulus sequence (from Torppa et al., 2012). 1 = standard, 2 = pitch deviant, 3 = intensity deviant, 4 = gap deviant, 5 = musical instrument deviant, 6 = duration deviant. After every stimulus there was a pause, not marked in the visualisation. The SOA was a constant 480 ms. The figure has been reprinted with permission from Elsevier.

pass filtering at 0.5 Hz. The analysis epoch was 550 ms, starting 100 ms before stimulus onset.

Epochs with extreme amplitudes exceeding +/- 300 - +/- 400 μV were rejected. The limit was set individually to retain 85% of the epochs for effective independent component analysis (ICA). The Fastica ICA algorithm was applied to remove CI, ocular, and muscle artefacts (Makeig, Debener, Onton, & Delorme, 2004; for details, see Torppa et al., 2012). After ICA, the data from missing electrode positions coinciding with the location of the post-aural CI speech processor were interpolated. Data dimensionality was reduced in accordance with the number of interpolated channels. Epochs having amplitude exceeding $\pm$ 150 μV were rejected, followed by the analysis of the proportion of remaining epochs for each individual participant. In each child, an inclusion criterion of 75% (95) remaining epochs for each deviant was applied. One participant with a CI did not reach this criterion at T1 and so her T1 data were excluded from the analyses. All children reached the inclusion criterion at T2. The mean percentage of accepted epochs at T1 was 94% (119 deviants, 2348 standards) and at T2 93% (116 deviants, 2330 standards). Since latency measures are especially sensitive to noise, responses from F3, Fz, F4, C3, Cz, and C4 electrodes were averaged to form a ROI (region of interest) channel, which was used in further ERP analyses.

To reduce the influence of extreme values, which are common in studies of young children, we calculated the median instead of the average for each sample point of each individual child's EEG-signal (Yabe, Saito, & Fukushima, 1993). Then, the data were filtered using a 25 Hz low-pass filter. After this, the median signals for each deviant and standard were averaged over participants using the time period of –50 to 0 ms before the tone onsets as the baseline level for epochs, and the subtraction (deviant-standard) waveform was calculated. Based on visual inspection of the data and polarity changes in the mastoids, MMN and P3a time windows for the quantification of individual response amplitudes were chosen that were the same for all deviants. The latency of the MMN was determined at the most

negative peak in the group average (all CI children) within a 90–250 ms time window after deviant onset in the subtraction waveform. Similarly, the latency of the P3a was determined at the most positive peak in the group average during the time window of 145–300 ms after deviant onset. Finally, the mean amplitudes for MMN and P3a were calculated from the individual responses as averages over a time window of 30 ms surrounding the group average peak latency.

The time windows for individual response latencies were separately identified for each deviant and thus differed from those for group-level response amplitudes because some individual peak latencies did not fall in the group-level latency window. This is justified since both in the current data and in previous studies on children with CIs (Torppa et al., 2012; Torppa, Huotilainen, et al., 2014), the individual response latencies could be identified. After visual inspection of individual data, the individual peaks were automatically detected in windows as follows: 85–250 ms for pitch (f0) and timbre MMN, 100–400 ms for intensity MMN, 145–400 ms for timbre and pitch (f0) P3a, and 200–450 ms for intensity P3a (see also Torppa, Huotilainen, et al., 2014).

ERP amplitudes were subjected to one-sample, two-tailed t-tests in order to examine whether they differed significantly from zero. If the MMN/P3a was significant for the deviant at T1 and/or T2, and the response did not change polarity from T1 to T2 in the time window of the specific (MMN or P3a) response, it was taken into the further statistical analyses.

STATISTICAL ANALYSES FOR TESTING THE HYPOTHESES

The Linear Mixed Model (LMM; Singer & Wilett, 2003; West, 2009) was used for testing the hypotheses. LMM was chosen because, first, it allows the utilization of data from all participants even though the data at one of the two time points are missing (Ibrahim & Molenberghs, 2009). This was important because we had to exclude ERP data from one CI child at T1, and two CI children at T1 and one CI child at T2 were not able to perform the speech in noise experiment. Second, the LMM allows the inclusion of results from the two time points in a single analysis.

The factors used in the LMM always included age and measurement time. For hypothesis 1 (compared to their NH peers, children with CIs will perform more poorly and show slower development of performance over age and time in the perception of speech in noise), the dependent variable was speech reception threshold (SRT$_{75}$) and independent variables were group, measurement time, and age. The other hypotheses were

tested in the CI children only. For hypothesis 2 (children with CIs who sing regularly at home will be advantaged in the perception of speech in noise compared to those who do not), the dependent variable was SRT$_{75}$ and the independent variables were CI group (= CI singing vs. CI non-singing group), measurement time, and age. For hypothesis 3 (in children with CIs, performance in the perception of speech in noise will correlate with: a) ERP measures of discrimination of pitch, intensity, and timbre for musical sounds, and b) behavioral discrimination of pitch and intensity in synthesized speech), the dependent variables were a) the MMN or P3a latencies or amplitudes for changes in pitch (f$_0$), timbre, or intensity, or b) the thresholds for behavioral discrimination of pitch (f$_0$) or intensity. Independent variables were SRT$_{75}$, measurement time, and age, and in the case of ERP responses, CI group (due to differences between CI singing and CI non-singing groups in the development of cortical processing of musical sounds, please see introductory section of this paper). In addition, in the case of MMN and P3a, the degree of change (small, medium, and/or large) was a term in the model to find out whether the connections to SRT$_{75}$ were consistent across degrees of change.

Due to the relatively small number of participants, there was insufficient power to include interactions involving more than three factors. Models were refined by the dropping of non-significant interactions, and only the significant interactions in each final model are reported. Selection of the best fitting LMM was determined by Akaike's and Bayesian information criteria (AIC and BIC, respectively) (Bryk & Raudenbush, 2002). AIC is a measure of the relative quality of a statistical model for a given set of data (Akaike, 1974), while BIC is a criterion for model selection among a finite set of models. The results presented are from the model with the best-fitted covariance structure and only the significant main effects of SRT$_{75}$ and interactions with SRT$_{75}$ are reported. The critical level for significance for LMM analyses was .05. Bonferroni correction was used for LMM post hoc tests.

Results

COMPARISON OF SPEECH-IN-NOISE PERFORMANCE BETWEEN GROUPS (HYPOTHESES 1 AND 2)

The NH children had lower (better) speech reception thresholds (SRT$_{75}$) than the CI children ($F_{2,37} = 282.61$; $B = -7.13$, reference = CI children; $p < .001$) (Figure 4a). In both child groups, perception of speech in noise was better at T2 than at T1 ($F_{1,37} = 27.96$; $B = 1.09$, reference = T2; $p < .001$: Figure 4b). SRT$_{75}$ was
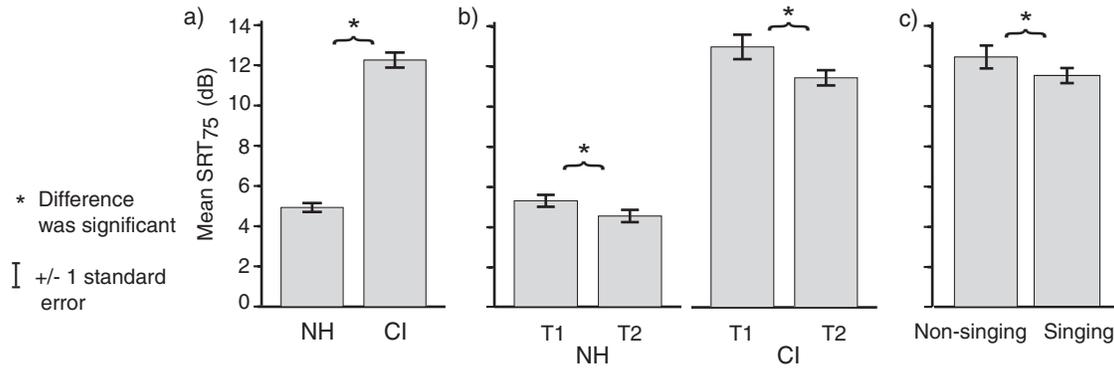
FIGURE 4. The illustration of a) difference between NH and CI children, b) development over time for NH and CI children, c) difference between CI non-singing ("non-singing") and CI singing ("singing") groups, in perception of speech in noise. Mean SRT_{75} = Mean of signal to noise-ratio for 75% correct responses (decibels).
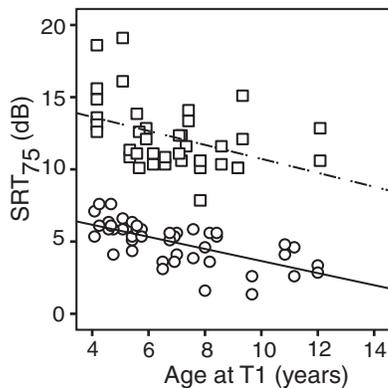
FIGURE 5. SRT_{75} (Signal to noise-ratio for 75% correct responses, decibels) as a function of age at T1, in the children with NH (circles and black line) and in the children with CIs (squares and dotted line). Responses for both T1 and T2 are included in the figure. Connection to age in LMM analysis for NH and CI group: $B = -0.44; p < .001$.

negatively related to age ($F_{1,37} = 19.66$, $B = -0.44$; $p < .001$), suggesting improving performance with age in both CI and NH children (Figure 5). There were no interactions of group with time or age suggesting that development was similar in both groups.

The analysis of connections to singing in the children with CIs showed that the CI singing group had better perception of speech in noise (lower SRT_{75}) than the CI non-singing group ($F_{1,19} = 4.98$; $B = 1.96$, reference = CI singing group; $p = .038$) (Figure 4c). While all children in the CI singing group were able to complete the task at T1 and T2, this was not the case for the CI non-singing group where one child could not complete the task at T1.

LINKS OF SPEECH-IN-NOISE PERFORMANCE TO ERPS FOR MUSICAL SOUND PROCESSING AND BEHAVIORAL DISCRIMINATION OF PITCH (F_0) AND INTENSITY (HYPOTHESIS 3)

MMN responses for all degrees of changes in pitch (f_0), for the most extreme change in timbre (from piano to cymbal) and for the 6 dB decrement, were significant at T1 and/or T2 (Table 3). P3a responses for all degrees of changes in pitch (f_0) and timbre were significant at T1 and/or T2. Only these MMN and P3a responses were included in statistical analyses (Table 3, Figure 6).

For P3a latency to changes in pitch, LMM analysis showed a significant three-way interaction of CI group (singing vs. non-singing group), age, and SRT_{75} ($F_{1,28} = 9.57$, $p = .004$), and two-way interactions of CI group with SRT_{75} ($F_{1,27} = 7.36$, $p = .012$) and of time with SRT_{75} ($F_{1,97} = 7.55$, $p = .007$). However, no significant pairwise differences within these interactions were found in post hoc tests.

MMN amplitudes in response to timbre change from piano to cymbal were larger with better speech-in-noise performance in the CI group as a whole ($F_{1,27} = 9.59$, $B = 0.79$, $p = .003$) (Figure 7a). P3a for all changes in timbre showed shorter latency with better speech-in-noise performance in the CI singing group only (Figure 7b) (the interaction of P3a latency with CI group was significant, $F_{1,32} = 7.86$, $p = .009$, and the connection between SRT_{75} and P3a latencies was significant only in the CI singing group, $B = 9.81$, $p = .006$).

The connections of perception of speech in noise to the amplitude or latency of the MMN for the 6 dB decrement were not significant, nor were there any significant connections of perception of speech in noise to behavioral discrimination of pitch (f_0) and intensity.

TABLE 3. *The MMN and P3a Mean Amplitudes and Latencies*

Stimulus eliciting the response:		CI group			
		T1 µV	T2 µV	T1 ms	T2 ms
Timbre	cembalo (S)	−1.06 (2.80)°	−0.82 (2.46)°	−	−
MMN	violin (M)	0.04 (2.17)	−0.12 (2.30)	−	−
	cymbal (L)	−2.44 (2.82)***	−02.17 (2.50)***	126 (40)	133 (31)
P3a	cembalo (S)	1.81 (1.98)***	2.20 (2.83)**	249 (60)	276 (54)
	violin (M)	2.82 (2.57)***	3.31 (2.88)***	218 (45)	248 (60)
	cymbal (L)	1.81 (1.88)***	1.49 (2.68)*	247 (52)	242 (60)
Pitch (f_0)	312 Hz (S)	−1.68 (2.69)*	−0.72 (1.67)°	147 (48)	158 (48)
MMN	351 Hz (M)	−1.47 (1.55)***	−1.37 (1.78)**	139 (26)	148 (41)
	441 Hz (L)	−1.46 (2.81)*	−1.81 (3.26)*	143 (44)	135 (46)
P3a	312 Hz (S)	0.94 (1.53)*	1.03 (2.58)°	265 (72)	307 (59)
	351 Hz (M)	1.49 (2.42)*	1.39 (2.65)*	266 (57)	283 (65)
	441 Hz (L)	0.64 (1.76)°	1.32 (2.60)*	248 (80)	274 (54)
Intensity	3 dB (S)	−0.43 (1.99)	−0.76 (1.92)°	−	−
decrement	6 dB (M)	−0.82 (1.86)*	−0.28 (2.14)	255 (83)	249 (84)
MMN	9 dB (L)	−0.19 (2.02)	−0.41 (2.04)	−	−
Intensity	3 dB (S)	−1.26 (.96)***	N	−	−
increment	6 dB (M)	−0.07 (1.61)	−0.90 (2.31)°	−	−
MMN	9 dB (L)	−0.20 (1.67)	−0.60 (1.94)	−	−
P3a	3 dB (S)	n	1.29 (1.81)**	−	−

S, M, L = small, medium and large degree of change. For both time points of the measurements (T1, T2), at first the mean amplitude (the standard deviation in parenthesis) and after that the significance of the response (°*p* < .10, *p < .05, **p ≤ .01, ***p < .001; two-tailed *t*-test against zero). Following these the mean latencies (and standard deviations) of the responses. The rows marked with gray present the values included in statistical analysis for testing the hypotheses. - = the latencies were not analyzed. n = the response was non-existent (wrong polarity in the time window of the response).

Discussion

The present results showed that children with CIs who sang regularly at home and whose parents sang for them at an early age (the CI singing group) perceived speech in noise better than the CI non-singing group, suggesting that further studies should address the question of whether singing improves this important perceptual skill. We also found that, compared to their normal-hearing (NH) peers, children with CIs developed similarly with age and over time (during 14 to 17 months) in the perception of speech in noise, but did not reach the performance of their NH peers. Further, better speech-in-noise performance was significantly associated in the CI group as a whole with better pre-attentive discrimination (larger MMN responses) of the change from piano to cymbal, and in the CI singing group only, with faster attention shifting (earlier latency of P3a responses) towards all changes in musical instrument timbre. These results suggest shared processing of timbre of musical instrument tones and speech in noise in children with CIs, particularly in the CI singing group, giving a baseline for future studies on the effects of musical instrument training in speech-in-noise performance. In the following, we will discuss these findings in more detail.

PERCEPTION OF SPEECH IN NOISE IN THE TWO CI GROUPS

As we predicted, the CI singing group, whose *auditory attention* shift towards sound changes was faster and who produced the *rhythm* of songs better compared to the CI non-singing group (Torppa, Huotilainen, et al., 2014), were better at perceiving speech in noise than the CI non-singing group. We assume that this result is not related purely to attention functions or to the perception of rhythm, but also to aspects of speech perception. For example, music training including singing has been shown to have beneficial effects on *speech segmentation* (François et al., 2013), a skill which is essential if hearing-impaired listeners are to achieve good perception of speech in noise (Woodfield & Ackeroyd, 2010; see introductory section of this paper). Perception of speech stress, an important cue for speech segmentation, is linked to the perception of musical rhythm (Hausen et al., 2013). Thus, the improved production of rhythm in the CI singing group can be related to better perception of speech stress and speech segmentation, leading to better perception of speech in noise.

Additionally, the *predictability of rhythm (meter) of songs and lyrics* may play a role in the present results. Meter can entrain cortical oscillations, leading to attention shift towards, and better speech perception of,
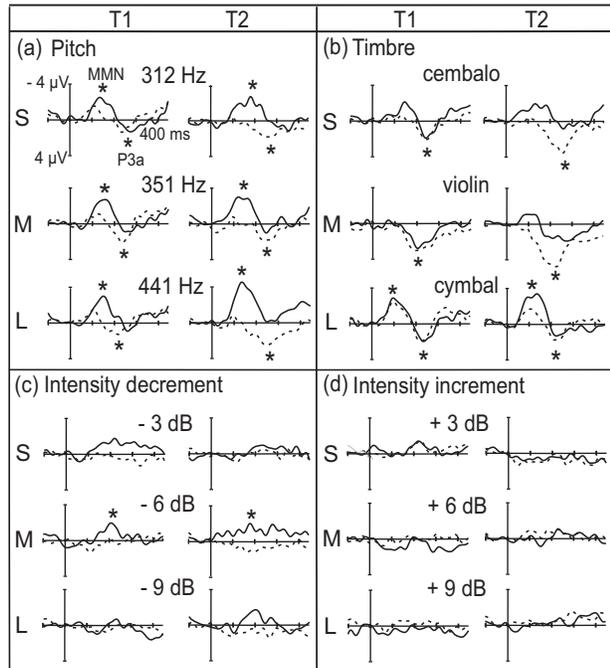
FIGURE 6. The subtraction (deviant - standard) ROI waveforms averaged across F3, Fz, F4, C3, Cz, and C4 electrodes for CI singing (dotted lines) and for CI non-singing (black lines) groups for A) pitch changes, B) timbre changes, C) intensity decrements, and D) intensity increments. S, M, L represent small, medium and large degrees of changes. * = The MMN or P3a was included in statistical analyses for testing hypotheses. The ERP waveforms are given for two time points of the measurements (T1 and T2 on the left and right in each panel, respectively. The figure is adapted from Torppa, Huotilainen, et al., 2014 (Frontiers in Psychology, http://journal.frontiersin.org/article/10.3389/fpsyg.2014.01389/full). Licensed under CC-BY.

FIGURE 7. Scatterplots of SRT_{75} (Signal to noise-ratio for 75% correct responses, decibels) against (panel a) MMN amplitudes for change to cymbal tone ($B = 0.79$, $p = .003$), and (panel b) P3a latencies for changes in timbre within the CI singing group ($B = 9.81$, $p = .006$). The plots show raw data without any correction for the effects of time and age. Panel a combines data for the both CI groups. Panel b combines data across all timbre deviants.

predicted stressed syllables (Gordon et al., 2011; Schön & Tillmann, 2015), and of lyrics in general (Leong et al., 2017), which may have a role in the development of speech perception in noise (as reviewed earlier). Notably, consistent and fast attention shifting to violations at rhythmically predictable moments in CI singing group (see introductory section of this paper and Torppa, Huotilainen, et al., 2014), suggests that the CI singing group is more efficient in extracting regularities from the auditory signal than the CI non-singing group. This would further improve the ability to predict regular time points, such as stressed syllables in rhythmically predictable songs. Thus, the ability to predict rhythmic regularity may be a factor in differences between the CI groups in speech-in-noise performance.

We assume that singing is important in the improved perception of speech in noise in the CI singing group because of the *multisensory context* of singing. For example, the motor movements that realize stressed syllables can particularly improve the sensitivity to speech at the predicted syllables (Falk & Dalla Bella, 2016; Picheny et al., 1985; see introductory section of this paper). Singing can also improve speech-related auditory-motor interactions and functional connectivity between auditory and speech motor regions. These both improve with more music training and may partially underlie enhanced perception of speech in noise (Du & Zatorre, 2017). Notably, in the aforementioned study more than half of the musicians participated also voice (singing) training. However, we cannot rule out that listening also makes a contribution since the parents sang more to the CI singing group at an early age. Conceivably, *parental singing* may be particularly effective at an early age when parents use child-directed singing with practically unchanged tempo in song repetitions (Bergeson & Trehub, 2002) and exaggerated stress patterns (Trainor, Clark, Huntley, & Adams, 1997). This singing style may be beneficial for directing attention to the phonemic content of songs (Lebedeva & Kuhl, 2010), for perception of stress patterns (Torppa, Faulkner, Järvikivi, & Vainio, 2010), and thus for speech segmentation (Jusczyk, 1999). It may also lead to enhanced attention (Rock et al., 1999). These all are thought to be important for the perception of speech in noise (segmentation, Woodfield & Ackeroyd, 2010; attention, Adank et al., 2012; Beer et al., 2011; Houston & Bergeson, 2014; Strait et al., 2012; Wild et al., 2012; see also introductory section of this paper).

Importantly, the CI singing and CI non-singing groups did not differ from each other in audiometric status, CI fitting, age, gender, socioeconomic background or aetiology, amount of general musical activities at home or outside of the home, other extracurricular or home activities, or the general engagement of the parents with their children. Former results have shown that speech-in-noise perception can improve with auditory training (number recognition in noise) despite the limitations of CIs (adults with CIs: Oba, Fu, & Galvin, 2011; children with CIs: Mishra et al., 2015). There also is longitudinal evidence of the beneficial effects of musical activities for the perception of speech in noise in NH children (Slater et al., 2015). Since the speech in noise performance of children with CIs can be improved with auditory training, the present results suggest that singing at an early age, by CI children by themselves and listening to parental singing, could lead to similar improvements of their speech in noise performance as seen with musical activities in NH children. Because the present results cannot confirm causality, further studies are needed to confirm this. Further studies are also needed to find out the mechanisms underlying the better speech in noise performance in the CI children who sing regularly and informally at home.

DEVELOPMENT OF PERCEPTION OF SPEECH IN NOISE IN CHILDREN WITH CIS AND NH

The implanted children in this study showed comparable improvements in their perception of speech in noise over time (during approximately 16 months), and with age to those of the NH group. This fits well with previous findings in children with NH, who show improving perception of speech in background noise with increasing age (Bradley & Sato, 2008; Hall et al., 2002; Nittrouer & Boothroyd, 1990; Stuart, 2005). However, as far as we know, this is the first time that such changes been found for children with CIs. The early age at implantation of the participants (no later than 3 years 1 month) might play a role here. Other studies failing to show effects of age have examined older and later implanted children; Looi and Radford (2011) examined a group with an age range of 11.1 to 14.4 and implantation age up to 8.5 years, while Jung and colleagues (2012) studied children aged between 8 and 16 and implanted up to 5 years of age. Moreover, Ruffin and others (2013) found that children with CIs who were older than 15 years, and who were implanted at a later age than the younger age groups, showed poorer perception of speech in noise than the younger ones. Early age at implantation allows more complete development of auditory function (Kral & Sharma, 2012) and has

been assumed to be beneficial for the perception of speech in noise (Caldwell & Nittrouer, 2013; Ruffin et al., 2013). Therefore, our findings are consistent with the assumption that early implantation accelerates the development of this perceptual ability.

As in previous studies, we found less tolerance of noise in the perception of speech in noise for children with CIs, also for the CI singing group, than for their NH peers. This suggests that neither early implantation, nor informal singing can completely compensate for the effects of the impoverished auditory input from the CI and the auditory pathology which may be associated with early deafness (Moore & Linthicum, 2007; for a review, Torppa, 2015).

CONNECTIONS OF MMN AND P3A TO PERCEPTION OF SPEECH IN NOISE FOR CHILDREN WITH CIS

Better perception of speech in noise was accompanied by a larger MMN to a change from piano to cymbal in the entire group of children with CIs and by earlier P3a for all changes of musical instrument timbre in the CI singing group. The linkage of speech in noise performance to these MMN and P3a outcomes is consistent with the view that for children with CIs, the perception of speech in noise and of musical instrument timbre share similar processes (for a review, Besson et al., 2011).

The MMN and P3a for timbre changes seen in children with CIs may be mediated by spectral differences or by temporal amplitude envelope differences (Won et al., 2010). Both classes of information are important in the perception of speech in quiet (Rosen, 1992). Evidence from adult CI users indicates that amplitude envelope (sound onset, attack, and decay) is a dominant cue for CI users' perception of the timbre of music instruments (Drennan & Rubinstein, 2008; Kong, Mullangi, Marozeau, & Epstein, 2011). However, there is no clear evidence that better envelope processing (in the timescale of temporal envelope cues for timbre) can contribute to better perception of speech in noise in CI users. Spectral cues, on the other hand, are clearly important for CI users' perception of speech in noise. For example, an increase of the number of speech processor spectral channels over the range of two to eight leads to improved perception of vowels, consonants, and sentences in noise for adult CI users (Friesen et al., 2001). In order to create efficient training methods, the importance of spectral and temporal cues in supporting the improvement of speech perception in noise needs further investigation.

That P3a latencies for all changes in music instrument timbre became shorter with better perception of speech in noise only in the CI singing group may be related to

their faster auditory attention shift (earlier P3a, especially for changes in pitch and timbre) compared to the CI non-singing group, found previously (see Torppa, Huotilainen, et al., 2014). It is even possible that for the CI singing group, attention shift is so fast that they can shift their attention to rapidly changing sounds both in the present music and speech contexts. Attention shift indexed by P3a is triggered only by clearly detectable or significant sound changes (Escera & Corral, 2007; Horvath et al., 2008), and reflects sound discrimination in CI children (Kileny et al., 1997). The specific process that underlies this connection cannot be determined from our results, but parental singing might lead to improved attention to the spectral content important for both perception of timbre and speech in noise. The parents sang for the CI singing group more, often face to face with the child. This could improve attention to spectral changes because the child can see the mouth and lip shapes and movements related to the spectral shape of consonant and vowel sounds. Former research indicates that CI users benefit more than NH listeners from lipreading, i.e., both seeing and hearing speech (Strelnikov, Rouger, Barone, & Deguine, 2009), and that integration of visual and auditory information is important for their speech perception (Anderson, Wiggins, Kitterick, & Hartley, 2017). Thus, seeing and hearing speech in the context of slow-rate and predictable songs may be particularly effective not only for the development of CI children's speech-in-noise perception, but also for their perception of spectral changes in general, and this could be related to the connection we found in the CI singing group only.

We also cannot rule out that the absence of connections between speech-in-noise performance and brain responses to changes in pitch and intensity is simply a null result. For example, it is possible that the changes were too large or too small, and therefore, the measurements were not sensitive enough to elicit differences between groups. However, the absent connections to discrimination of these similar sound dimensions in speech support the conclusion that the perception of speech in noise and processing of changes of pitch and intensity between static music instrument sounds do not share similar processes in children with CIs. Furthermore, recent results from Lo and colleagues (2015) showed that pitch-based (melodic contour) training that improved perception of melodic contours did not enhance perception of speech in noise of children with CIs. This suggests that pitch is not a limiting factor in these children's speech-in-noise perception. Future studies are also needed to show whether the connections are absent due to age of the present participants or due to CI input. This could be done, for example, by comparing the connections to those in age-matched NH children.

CAVEATS AND FUTURE DIRECTIONS

One caveat of the present study is that data on the regularity of singing was collected retrospectively, and lacked a detailed log of time spent in singing. The present study also can only show a correlational link between neural processing of musical instrument timbre and perception of speech in noise. A more substantial reason for caution is that group assignment could not be random and only a randomized controlled study can establish a causal link between singing or musical instrument playing and the perception of speech in noise. However, as Schön and Tillmann (2015) and Gfeller (2016) state, it is difficult to organize these kind of intervention studies even in normal-hearing populations, and even more challenging in children with CIs. While waiting for these intervention studies, it might be beneficial to encourage singing activities of children with CIs and their parents. The encouragement can easily be embedded in rehabilitation, as the Lindfors Foundation MUKULA-project (http://lindforsinsaatio.net/) in Finland, Auditory-Verbal Therapy (AVT; Estabrooks, 1994) and rehabilitation materials (like STEPS Together, http://www.earfoundation.org.uk/shop/items/76) have shown. Many CI clinics and speech and language therapists have adopted these materials and methods in their everyday work and have encouraged parents to engage in singing at home with their children.

Moreover, while waiting for the intervention studies, it might be beneficial to let the children with CIs play musical instruments or sing in a way that would enhance the perception of acoustic cues shared between timbre and speech. As Patel (2011) describes, musical instrument training targeted on the perception of amplitude envelope could employ musical rhythm, or detection of the sounds of musical instruments where envelope cues can signal timbre. It is likely to be useful also to include musical tasks that can assist the development of the perception of spectral differences—which might be best provided using sung vowel sounds supported by visual cues from the singer's lip and mouth shape. In line with the OPERA hypothesis' predictions, training should also be associated with strong positive emotion, extensive repetition, and focused attention.

Conclusions

The present findings show that early-implanted children with unilateral CIs, aged 4 to 13 years, and

consistently exposed to singing, show development of speech-in-noise perception over time, and with age, at rates similar to NH peers, even though they do not achieve the same performance. The children with CIs who sang regularly at home were better at the perception of speech in noise than other children with CIs. The results also show that the perception of speech in noise by children with CIs is connected to their pre-attentive discrimination of a change of musical instrument from piano to cymbal, and in those who sing regularly, to attention shift towards timbre changes in general (from piano to cembalo, violin and cymbal). This suggests that the perception of musical instrument sound and speech in noise share similar perceptual and attention-related processing. Since young prelingually implanted children enjoy singing (Trehub, Vongpaisal, & Nakata, 2009), our results suggest that singing by parents and children themselves might be a motivating way to enhance CI children's perception of speech in noise. Further, musical instrument training might lead to improvement of this important skill. However, randomized and controlled intervention studies in children with CIs are necessary to confirm the role of singing and music instrument playing for improvement of perception of speech in noise.

Author Note

Correspondence concerning this article should be addressed to Ritva Torppa, University of Helsinki, P. O. Box 21 (Haartmaninkatu 3), FIN-00014 University of Helsinki, Finland. E-mail: ritva.torppa@helsinki.fi

References

Adank, P., Davis, M. H., & Hagoort, P. (2012). Neural dissociation in processing noise and accent in spoken language comprehension. *Neuropsychologia, 50*, 77–84.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723.

Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jääskeläinen I. P., et al. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: Magnetoencephalographic recordings. *Psychophysiology, 35*, 211–224.

Anderson, C. A, Wiggins, I. M., Kitterick, P. T., & Hartley, D. E. H. (2017). Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. *Proceedings of the National Academy of Sciences of the United States of America, 114*, 10256–10261.

Asp, F., Mäki-torkko, E., Karltorp, E., Harder, H., Hergils, L., Eskilsson, G., & Stenfelt, S. (2012). Bilateral versus unilateral cochlear implants in children: Speech recognition, sound localization, and parental reports. *International Journal of Audiology, 51*, 817–832.

Baker, M., Buss, E., Jacks, A., Taylor, C., & Leibold, L. J. (2014). Children's perception of speech produced in a two-talker background. *Journal of Speech Language and Hearing Research, 57*, 327–337.

Beer, J., Kronenberger, W. G., & Pisoni, D. B. (2011). Executive function in everyday life: Implications for young cochlear implant users. *Cochlear Implants International, 12*, S89–S91.

Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mother's songs to infants. *Psychological Science, 13*, 72–75.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology, 2*, 94.

Bolger, D., Coull, J. T., & Schön, D. (2014). Metrical rhythm implicitly orients attention in time as indexed by improved target detection and left inferior parietal activation. *Journal of Cognitive Neuroscience, 26*, 593–605

Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *Journal of the Acoustical Society of America, 123*, 2078–2086.

Bryk, S. W., & Raudenbush, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Caldwell, A., & Nittrouer, S. (2013). Speech perception in noise by children with cochlear implants. *Journal of Speech Language and Hearing Research, 56*, 13–30.

Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia, 50*, 2652–2658.

Coffey, E. B. J., Mogilever, N. B., & Zatorre, R. J. (2017). Speech-in-noise perception in musicians: A review. *Hearing Research, 352,* 49–69.

Delorme, A., & Makeig, S. (2004). EEGlab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*, 9–21.

Drennan, W. R., & Rubinstein, J. T. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of Rehabilitation Research and Development, 45*, 779–789.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., & Westermann, S. (2001). ICRA Noises: Artificial noise signals with speech-like spectral and temporal properties for hearing aid assessment. *Audiology, 40*, 148–157.

Du, Y., & Zatorre, R. J. (2017). Musical training sharpens and bonds ears and tongue to hear speech better. *Proceedings of the National Academy of Sciences of the United States of America, 114,* 13579–13584.

Escera, C., & Corral, M. J. (2007). Role of mismatch negativity and novelty-P3 in involuntary auditory attention. *Journal of Psychophysiology, 21*, 251–264.

Estabrooks, W. (1994). *Auditory-verbal therapy for parents and professionals*. Washington DC: Alexander Graham Bell Association for the Deaf.

Falk, S., & Dalla Bella, S. (2016). It is better when expected: Aligning speech and motor rhythms enhances verbal processing. *Language, Cognition and Neuroscience, 31,* 699–708.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *Journal of the Acoustical Society of America, 108*, 3023–3029.

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *Plos One, 10*.

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex, 23*, 2038–2043.

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *Journal of the Acoustical Society of America, 110*, 1150–1163.

Fu, Q.-J., & Galvin, J. J., III (2008). Maximizing cochlear implant patients' performance with advanced speech training procedures. *Hearing Research, 242*, 198–208.

Fu, Q. J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *Jaro-Journal of the Association for Research in Otolaryngology, 6*, 19–27.

Geers, A. E., Davidson, L. S., Uchanski, R. M., & Nicholas, J. G. (2013). Interdependence of linguistic and indexical speech perception skills in school-age children with early cochlear implantation. *Ear and Hearing, 34*, 562–574.

Gfeller, K. (2016). Music-based training for pediatric CI recipients: A systematic analysis of published studies. *European Annals of Otorhinolaryngology, Head and Neck Diseases, 133,* 50–56.

Gordon, R. L., Magne, C. L., & Large, E. W. (2011). EEG correlates of song prosody: A new look at the relationship between linguistic and musical rhythm. *Frontiers in Psychology, 2*, 352.

Hall, J. W., Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing, 23*, 159–165.

Hausen, M., Torppa, R., Salmela, V. R., Vainio, M., & Särkämö, T. (2013). Music and speech prosody: A common rhythm. *Frontiers in Psychology, 4*, 566.

Horvath, J., Winkler, I., & Bendixen, A. (2008). Do N1/Mmn, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology, 79*, 139–147.

Houston, D. M., & Bergeson, T. R. (2014). Hearing versus listening: Attention to speech and its role in language acquisition in deaf infants with cochlear implants. *Lingua, 139*, 10–25.

Ibrahim, J. G., & Molenberghs, G. (2009). Missing data methods in longitudinal studies: a review. *Test, 18*, 1–43.

Jung, K. H., Won, J. H., Drennan, W. R., Jameyson, E., Miyasaki, G., Norton, S. J., & Rubinstein, J. T. (2012). Psychoacoustic performance and music and speech perception in prelingually deafened children with cochlear implants. *Audiology and Neuro-Otology, 17*, 189–197.

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences, 3*, 323–328.

Kelly, A. S., Purdy, S. C., & Thorne, P. R. (2005). Electrophysiological and speech perception measures of auditory processing in experienced adult cochlear implant users. *Clinical Neurophysiology, 116*, 1235–1246.

Kileny, P. R., Boerst, A., & Zwolan, T. (1997). Cognitive evoked potentials to speech and tonal stimuli in children with implants. *Otolaryngology-Head and Neck Surgery, 117*, 161–169.

Klatt, D. H. (1980). Software for a cascade-parallel formant synthesizer. *Journal of the Acoustical Society of America, 67*, 971–995.

Kong, Y.-Y., Mullangi, A., Marozeau, J., & Epstein, M. (2011). Temporal and spectral cues for musical timbre perception in electric hearing. *Journal of Speech Language and Hearing Research, 54*, 981–994.

Kral, A., & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in Neurosciences, 35*, 111–122.

Kujala, T., & Näätänen, R. (2010). The adaptive brain: A neurophysiological perspective. *Progress in Neurobiology, 91*, 55–67.

Lebedeva, G. C., & Kuhl, P. K. (2010). Sing that tune infants' perception of melody and lyrics and the facilitation of phonetic recognition in songs. *Infant Behavior and Development, 33*, 419–430.

Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2017). The temporal modulation structure of infant-directed speech. *Open Mind: Discoveries in Cognitive Science, 1,* 78–90.

Levitt, H. (1971). Transformed up-down methods in psycho-acoustics. *Journal of the Acoustical Society of America, 49*, 467–477.

Lo, C. Y., McMahon, C. M., Looi V., & Thompson, W. F. (2015). Melodic contour training and its effect on speech in noise, consonant discrimination, and prosody perception for cochlear implant recipients. *Behavioural Neurology, 1–10,* 352869.

Lonka, E., Kujala, T., Lehtokoski, A., Johansson, R., Rimmanen, S., Alho, K., & Näätänen, R. (2004). Mismatch negativity brain response as an index of speech perception recovery in cochlear-implant recipients. *Audiology and Neuro-Otology, 9*, 160–162.

Looi, V., & Radford, C. J. (2011). A comparison of the speech recognition and pitch ranking abilities of children using a unilateral cochlear implant, bimodal stimulation or bilateral hearing aids. *International Journal of Pediatric Otorhinolaryngology, 75*, 472–482.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences, 8*, 204–210.

Mattys, S. L. (1997). The use of time during lexical processing and segmentation: A review. *Psychonomic Bulletin and Review, 4,* 310–329.

Mishra, S. K., Boddypally, S. P., & Rayapati, D. (2015). Auditory learning in children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 58*, 1052–1060.

Moore, B. C. J. (2003). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology and Neurotology, 24*, 243–254.

Moore, J. K., & Linthicum, F. H., Jr. (2007). The human auditory system: A timeline of development. *International Journal of Audiology, 46*, 460–478.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*, 2544–2590.

Näätänen, R., Petersen, B., Torppa, R., Lonka, E., & Vuust, P. (2017). The MMN as a viable and objective marker of auditory development in CI users. *Hearing Research, 353*, 57–75.

Nager, W., Münte, T. F., Bohrer, I., Lenarz, T., Dengler, R., Moebes, J., et al. (2007). Automatic and attentive processing of sounds in cochlear implant patients - Electrophysiological evidence. *Restorative Neurology and Neuroscience, 25*, 391–396.

Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *Journal of the Acoustical Society of America, 87*, 2705–2715.

Oba, S. I., Fu, Q.-J., & Galvin, J. J. (2011). Digit training in noise can improve cochlear implant users' speech understanding in noise. *Ear and Hearing, 32*, 573–581.

O'Halpin, R. (2010). *The perception and production of stress and intonation by children with cochlear implants* (Unpublished doctoral dissertation). University College London, London, United Kingdom. http://eprints.ucl.ac.uk/20406/

Opolko, F., & Wapnick, J. (2006). *The McGill University master samples collection on DVD.* Quebec, Canada: McGill University.

Parbery-clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing, 30*, 653–661.

Parbery-clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia, 49*, 3338–3345.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology, 2*, 142.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research, 308*, 98–108.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research, 28*, 96–103.

Ponton, C. W., Eggermont, J. J., Don, M., Waring, M. D., Kwong, B., Cunningham, J., & Trautwein, P. (2000). Maturation of the mismatch negativity: Effects of profound deafness and cochlear implant use. *Audiology and Neuro-Otology, 5*, 167–185.

Putkinen, V., Tervaniemi, M., & Huotilainen, M. (2013). Informal musical activities are linked to auditory discrimination and attention in 2-3-year-old children: An event-related potential study. *European Journal of Neuroscience, 37*, 654–661.

Rock, A. M. L., Trainor, L. J., & Addison, T. (1999). Distinctive messages in infant-directed lullabies and play songs. *Developmental Psychology, 35*, 527–534.

Rosen, S. (1992). Temporal information in speech – Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 336*, 367–373.

Ruffin, C. V., Kronenberger, W. G., Colson, B. G., Henning, S. C., & Pisoni, D. B. (2013). Long-term speech and language outcomes in prelingually deaf children, adolescents and young adults who received cochlear implants in childhood. *Audiology and Neuro-Otology, 18*, 289–296.

Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *Plos One, 9*, e86980.

Sandmann, P., Kegel, A., Eichele, T., Dillier, N., Lai, W., Bendixen, A., et al. (2010). Neurophysiological evidence of impaired musical sound Perception in cochlear-implant users. *Clinical Neurophysiology, 121*, 2070–2082.

Schön, D., & Tillmann, B. (2015). Short- and long-term rhythmic interventions: Perspectives for language rehabilitation. *Neurosciences and Music V: Cognitive Stimulation and Rehabilitation, 1337*, 32–39.

Shahin, A. J. (2011). Neurophysiological influence of musical training on speech perception. *Frontiers in Psychology, 2*, 126.

Singer, J., & Wilett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York: Oxford University Press.

Slater, J., & Kraus, N. (2016). The role of rhythm in perceiving speech in noise: A comparison of percussionists, vocalists and non-musicians. *Cognitive Processing, 17*, 79–87.

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research, 291*, 244–252.

Speech Filing System Tools for Speech Research (n.d.). Retrieved from https://www.phon.ucl.ac.uk/resource/sfs/

Strait, D. L., Parbery-clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language, 123*, 191–201.

Strelnikov, K., Rouger, J., Barone, P., & Deguine, O. (2009). Role of speechreading in audiovisual interactions during the recovery of speech comprehension in deaf adults with cochlear implants. *Scandinavian Journal of Psychology, 50*, 437–444.

Stuart, A. (2005). Development of auditory temporal resolution in school-age children revealed by word recognition in continuous and interrupted noise. *Ear and Hearing, 26*, 78–88.

Torppa, R. (2015). *Pitch-related auditory skills in children with cochlear implants: The role of auditory working memory, attention and music* [Academic dissertation]. Studies in Psychology 113. Helsinki, Finland: University of Helsinki. https://helda.helsinki.fi/handle/10138/157046

Torppa, R., Faulkner, A., Huotilainen, M., Järvikivi, J., Lipsanen, J., Laasonen, M., & Vainio, M. (2014). The perception of prosody and associated auditory cues in early-implanted children: The role of auditory working memory and musical activities. *International Journal of Audiology, 53*, 182–191.

Torppa, R., Faulkner, A., Järvikivi, J., & Vainio, M. (2010). Acquisition of focus by normal hearing and Cochlear Implanted children: The role of musical experience. *Speech Prosody 2010.* Chicago, IL: International Speech Communication Association. https://www.isca-speech.org/archive/sp2010/sp10_977.html

Torppa, R., Huotilainen M., Leminen, M., Lipsanen, J., & Tervaniemi, M. (2014). Interplay between singing and cortical processing of music: A longitudinal study in children with cochlear implants. *Frontiers in Psychology, 5*, 1389.

Torppa, R., Salo, E., Makkonen, T., Loimo, H., Pykäläinen, J., Lipsanen, J., et al. (2012). Cortical processing of musical sounds in children with cochlear implants. *Clinical Neurophysiology, 123*, 1966–1979.

Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development, 20*, 383–396.

Trehub, S. E., Vongpaisal, T., & Nakata, T. (2009). Music in the lives of deaf children with cochlear implants. *Neurosciences and Music III: Disorders and Plasticity, 1169*, 534–542.

Tyler, R., & Holstad, B. (1987). *A closed set speech perception test for hearing-impaired children.* Iowa City, IA: University of Iowa.

Uther, M., Kujala, A., Huotilainen, M., Shtyrov, Y., & Näätänen, R. (2006). Training in Morse code enhances involuntary attentional switching to acoustic frequency: Evidence from ERPs. *Brain Research, 1073*, 417–424.

West, B. T. (2009). Analyzing longitudinal data with the linear mixed models procedure in spss. *Evaluation and the Health Professions, 32*, 207–228.

Wetzel, N., Widmann, A., Berti, S., & Schröger, E. (2006). The development of involuntary and voluntary attention from childhood to adulthood: A combined behavioral and event-related potential study. *Clinical Neurophysiology, 117*, 2191–2203.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *Journal of Neuroscience, 32*, 14010–14021.

WINKLER, I., DENHAM, S. L., & NELKEN, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences, 13*, 532–540.

WINKLER, I., TERVANIEMI, M., SCHRÖGER, E., WOLFF, C., & NÄÄTÄNEN, R. (1998). Preattentive processing of auditory spatial information in humans. *Neuroscience Letters, 242*, 49–52.

WON, J. H., DRENNAN, W. R., KANG, R. S., & RUBINSTEIN, J. T. (2010). Psychoacoustic abilities associated with music perception in cochlear implant users. *Ear and Hearing, 31*, 796–805.

WOODFIELD, A., & AKEROYD, M. A. (2010). The role of segmentation difficulties in speech-in-speech understanding in older and hearing-impaired adults. *Journal of the Acoustical Society of America, 128*, EL26–EL31.

YABE, H., SAITO, F., & FUKUSHIMA, Y. (1993). Median method for detecting endogenous event-related brain potentials. *Electroencephalography and Clinical Neurophysiology, 87*, 403–407.