Pitch-related auditory skills in children with cochlear implants: The role of auditory working memory, attention and music

Ritva Torppa

Cognitive Brain Research Unit, Cognitive Science,
Institute of Behavioural Sciences
University of Helsinki
Finland

Academic dissertation to be publicly discussed,
by due permission of the Faculty of Behavioural Sciences
at the University of Helsinki in Auditorium 107 at the Athena building,
Siltavuorenpenger 3 A, on the 6th of November, 2015, at 12 o'clock

University of Helsinki
Institute of Behavioural Sciences
Studies in Psychology 113: 2015

Supervisors:

Professor Minna Huotilainen, PhD
Cognitive Brain Research Unit
Cicero learning
Institute of Behavioural Sciences
University of Helsinki, Finland *and*
Brain Work Research Centre
Finnish Institute of Occupational Health
Helsinki, Finland

Professor Andrew Faulkner, D. Phil
Research Department of Speech, Hearing and Phonetic Sciences
University College London
London, United Kingdom

Professor Martti Vainio, PhD
Institute of Behavioural Sciences
University of Helsinki, Finland

Reviewers:

Emeritus Professor Brian C. J. Moore, PhD
Department of Experimental Psychology
University of Cambridge
Cambridge, United Kingdom

Jyrki Tuomainen, PhD, Senior Lecturer
Speech, Hearing and Phonetic Sciences
University College London
London, United Kingdom

Opponent:

Dr., Res. Dir. Curtis Ponton, PhD
House Research Institute
Los Angeles, CA, United States *and*
Chief Scientist, Vice President
Compumedics Neuroscan
Charlotte, NC, United States

Contents

Abstract

The cochlear implant (CI) provides a sensation of hearing and the opportunity to develop spoken language for deaf-born children. However, many CI children show poor language outcomes, which may be related to the deficiency of CIs in delivering pitch. The present thesis studies the development of those neural processes and behavioural skills linked to the perception of pitch which may play a role in language acquisition. We measured with event-related brain potentials (ERPs) the neural discrimination of and attention shift to changes in music, the perception of word and sentence stress and related acoustic cues, and the auditory working memory (forward digit span) in 4–13-year-old normally hearing (NH) and early-implanted children. We studied how the development of these aspects is related to musical activities known to advance brain development and perceptual skills in the NH population, and whether the perception of music (pitch or rhythm) is connected to word stress or visuospatial perception in NH adults. With regard to the development of neural responses, we found for the CI children usually well-formed ERP waveforms resembling those found for the NH children. However, some brain responses implied impoverished processing for the CI children, especially for timbre and pitch. The CI children who sang regularly at home were advantaged over the other CI children for the development of attention shift, which was linked to improved auditory working memory, implying better neural discrimination, an advantaged development of neural networks for attention and better updating of auditory working memory for the CI singers. We found that for the CI children perception of word and sentence stress improved with improving discrimination of pitch (f_0) and intensity and auditory working memory. For the perception of stress and related aspects, including pitch and auditory working memory, only the CI children participating in supervised musical activities performed and developed similarly to the NH children. Moreover, the perception of musical rhythm improved with improving word stress and visuospatial perception for the NH adults. Thus, the results indicate that (i) perception of music and speech are connected not only via pitch and timbre, but also via rhythm, and (ii) the combination of singing at home and taking part in supervised musical activities, using also rhythmic exercises and visual cues, might be the best way to optimize pitch-related abilities, underlying cognitive functions, spoken language skills and quality of life for early-implanted children.

Tiivistelmä

Sisäkorvaistute (SI) mahdollistaa syntymäkuurojen lasten puhekielen kehityksen. SI-lasten puhekielen taidot vaihtelevat kuitenkin paljon ja ovat usein heikompia kuin kuulevilla lapsilla. Tämä saattaa liittyä SI:n heikkoon kykyyn välittää äänten korkeuksia. Tässä väitöskirjassa tarkasteltiin puhekielen kehitystä tukevien, äänen korkeuksien havaitsemiseen liittyvien hermostollisten mekanismien, kuulomuistin ja kuuntelutaitojen kehitystä 4–13-vuotiailla kuulevilla lapsilla sekä lapsilla, joiden SI oli aktivoitu varhaisessa iässä. Musiikillisten äänten hermostollista erottelua ja kuulotarkkaavuutta tarkasteltiin mittaamalla kuuloherätevasteita aivosähkökäyrällä (EEG). Sana- ja lausepainon havaitsemista ja näihin liittyvien akustisten vihjeiden erottelukykyä tutkittiin kuuntelukokeilla ja kuulomuistia numerotoistotestillä. Musiikin harrastamisen tiedetään parantavan kuulevien lasten kuulohavaintotaitoja ja kuulomuistia, ehkä myös kuulotarkkaavuutta. Siksi tarkastelimme SI-lasten ryhmässä musiikkiaktiviteettien yhteyksiä mittaustuloksiimme. Tutkimme myös, onko musiikin havaitseminen yhteydessä sanapainon tai suuntien havaitsemiseen kuulevilla aikuisilla. Havaitsimme, että SI-lasten kuuloherätevasteet olivat hyvin samantapaisia kuin kuulevien lasten. Kuitenkin vasteet erityisesti muutoksille soittimesta toiseen (äänen laatu) ja äänen korkeudessa heijastivat SI-lasten heikkoa kuuloerottelukykyä ja -tarkkaavuutta. Säännöllisesti kotona laulavien SI-lasten kuulotarkkaavuusvasteet kehittyivät voimakkaammiksi ja nopeammiksi kuin muiden SI-lasten. Nämä vasteet olivat nopeampia paremman kuulomuistin myötä. Tulokset viittaavat laulavien SI-lasten hyvään hermostolliseen erottelukykyyn, kuulotarkkaavuuden aivoverkostojen kehitykseen ja kuulomuistin päivitykseen. Sana- ja lausepainon havaitseminen parantui hyvän äänen korkeuden (f_0) ja voimakkuuden erottelun sekä kuulomuistin myötä, joissa vain ohjattuihin musiikkiaktiviteetteihin osallistuneet SI-lapset kehittyivät yhtä hyvin kuin kuulevat lapset. Musiikin rytmien havaitseminen parantui hyvän sanapainon ja suuntien havaitsemisen myötä. Tulokset korostavat musiikin havaitsemisen olevan yhteydessä musiikin rytmien havaitsemiseen, ei pelkästään äänen korkeuksien ja laatujen havaitsemiseen. Ne korostavat kotona laulamisen sekä äänen korkeuden, myös rytmin havaitsemisen harjoituksia ja suuntavihjeitä (kuten laululeikkejä), sisältävän ohjatun musiikkitoiminnan tärkeyttä SI lasten puhekielen ja elämänlaadun parantamisessa.

Acknowledgements

This thesis has been made with a dream team of supervisors. First, I am deeply grateful to my supervisor Professor Andrew Faulkner. When I met him for the first time, I had an idea in my mind, but only very little scientific knowledge. The first plan for this thesis was the fruit of his excellent knowledge on cochlear implants and perception of speech. Without him this thesis would not exist, and I will never forget our discussions in the wonderful atmosphere of London and UCL!

I want to thank as much my supervisor Professor Minna Huotilainen. I had to sleep over one night after the first discussion with her to find out if I really want to be a brain researcher. However, her excellent knowledge on brain responses and music research, and her inspiring way to talk, convinced me. The journey with event-related potentials and music has been long but thanks for Minna, worth to go through!

The third member of this dream team is Professor Martti Vainio. He has been utmost important for me from the perspective of phonetics and prosody. He also helped me concretely with recordings and figures and supported me when I needed that. Thank you!

I am utterly grateful to Professors Mari Tervaniemi and Petri Toiviainen for accepting me into Finnish Centre of Excellence in Interdisciplinary Music Research. This team was a window to music research, covered the costs of our measurements, and guaranteed the peace to work until year 2013. Special thanks for Professor Mari Tervaniemi for being my co-author and for her kind support!

I want to express my gratitude for my co-authors: university lecturer Jari Lipsanen, Johannes Pykäläinen, Hannu Loimo, Emma Salo, Professor Juhani Järvikivi, Maija Hausen, Dr. Viljami Salmela, Dr. Marja Laasonen, laboratory engineers Miika Leminen and Tommi Makkonen (your technical support was really important!), Doc. Teppo Särkämö and Dr. Eino Partanen. The Brain and Music Team has provided an atmosphere where everyone supports everyone. Thank you for all of you, including Dr. Vesa Putkinen, Dr. Paula Virtala, Tanja Linnavalli, Katri Saarikivi, Caitlin Dawson and Dr. Elvira Brattico. I am grateful to CBRU, especially for Academy Professor Risto Näätänen for his unique impact in the ERP research field, and for Doc. Sari Ylinen for help with new research plans. Thanks for Professors Kimmo Alho, Andrej Kral and Istwan Winkler for scientific advices, for Marko Rönkä (MED-EL), Mika Teivainen (Cochlear), Ville

List of original publications

This thesis is based on the following original publications, referred to in the text by Roman numerals (I–IV).

I Torppa, R., Salo, E., Makkonen, T., Loimo, H., Pykäläinen, J., Lipsanen, J., Faulkner, A., & Huotilainen, M. (2012). Cortical processing of musical sounds in children with Cochlear Implants. *Clinical Neurophysiology,* 123, 1966–1979.

II 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.

III 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, 1821–91.

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

The articles are reprinted with the kind permission of the copyright holders.

Abbreviations

CI	Cochlear implant
CI child	Child with a cochlear implant
CIm	CI child who participated in supervised musical training
CIn	CI child who did not participate in supervised musical training
CIs	CI child who sang at home regularly
CIns	CI child who did not sing at home regularly
DAT	Dynamic attending theory
EEG	Electroencephalography
ERP	Event-related potential
f_0	Fundamental frequency
MBEA	Montreal Battery of Evaluation of Amusia
MMN	Mismatch negativity
NH	Normal hearing
NH child	Child with normal hearing
PT	Planum temporale

1 Introduction

Approximately one or two of every 1000 newborns has profound congenital hearing loss (Nikolopoulos & Vlastarakos, 2010). As of 2013, the cochlear implant (CI) provides a sensation of hearing for 80 000 individuals born with hearing loss (Boons et al., 2013a). Despite the positive effect of CIs, the language and speech perception outcomes of children with CIs (CI children) vary extensively, many of them showing lower language skills than normal hearing (NH) children (Boons et al., 2013a, 2013b; Geers et al., 2003; Niparko et al., 2010). This thesis investigates issues linked to the idea that a poor ability to perceive prosody, assessed here by perception of word and sentence stress, may contribute to poor speech and language outcomes. CI children have variable and often poor ability to perceive word and sentence stress (Meister et al., 2011; O'Halpin, 2010), both of which are relevant for segmentation of continuous speech and spoken language development (Friedrich et al., 2009; Jusczyk et al., 1999; Thiessen et al., 2005). Prosodic perception can be expected to be degraded due to the limitations of CIs in delivering pitch (Ciocca et al., 2002; Green et al., 2004; Laneau & Wouters, 2004), leading also to difficulties in perception of music (Hsiao & Gfeller, 2012; McDermott, 2004; Limb & Roy, 2014). It has been suggested that improving perception of pitch and music can lead to improved perception of speech, especially in noisy situations where CI listeners typically have severe difficulties (Drennan & Rubinstein, 2008). Therefore, this thesis addresses the development of speech prosody and music, and the possible associated factors: discrimination of acoustic cues, auditory working memory, auditory attention, visuospatial perception, and most importantly, musical activities in early-implanted children whose CI had been activated prior to the age of three years one month. Early-implanted children are now beginning to form a majority of CI children, and little was known on the issues under investigation in this child population.

1.1 Cochlear implants and perception of acoustic cues for prosody and music

When the variations of air pressure that constitute sound reach the ear, they produce corresponding movement of the round window in the interface of the middle and the inner ear. This leads to the movement of the basilar membrane in the cochlea. The inner hair

cell bodies are attached to the basilar membrane, and their cilia are in contact with the tectorial membrane. Movement of the basilar membrane relative to the tectorial membrane causes the deflection of the cilia of the inner hair cells, leading to the generation of action potentials in the neurons of the auditory nerve (Moore, 2003a, 2003b). Deafness is a consequence of the damage to or total loss of sensory inner hair cells due to genetic cause, infectious diseases like meningitis or rubella, or other factors (Wilson & Dorman, 2008).

The CI bypasses these damaged or missing hair cells and all other structures of the auditory system that precede them, and stimulates directly the auditory nerve through electrodes inserted in the inner ear. A microphone placed above or within the pinna receives sounds. The input sounds, over a frequency range approximately from 200 Hz to 8500 Hz, are filtered in a speech processor into bands of frequencies. Within each of these frequency bands, the amplitude envelope is extracted, encoding time-varying sound level at rates up to a few hundred Hz (Limb & Roy, 2014; Wilson & Dorman, 2008; for CI coding strategies, CIS, Wilson et al., 1991; ACE, Kiefer et al., 2001). Pulse levels representing these envelopes are directed to electrodes along the electrode array so as to encode the time-varying spectrum of sound as time-varying pulse levels distributed spatially along the array. The outputs of low frequency bands are directed to apical electrodes, and the outputs of high frequency bands are directed to basal electrodes. Thus the auditory nerves are stimulated in the order of frequency mapping in the normal cochlea, in so-called tonotopic order. The electric current pulses normally stimulate the auditory nerves at a fixed pulse rate, which is in CIS and ACE processors at least 700 pulses per second and sometimes higher (Wilson & Dorman, 2008). An exception to these coding strategies is the fine structure processing speech coding strategy (FSP), where additionally the temporal fine structure of sounds is encoded by pulses of varying rate synchronized to the temporal fine structure, which are directed to up to four of the most apical electrodes (Riss et al., 2014).

Pitch. The natural sounds that convey a sense of pitch are quasi-periodic tones. The sound pressure waveform of these tones repeats at a constant or relatively slowly changing rate. Such tones are composed of a series of sinusoidal waves (harmonics), whose frequencies are integer multiples of the fundamental frequency (f_0), which is the repetition frequency of the complex wave (Moore, 2003a, 2003b). It is not yet completely clear how pitch is

derived from these complex tones even in the normal auditory system. However, from the perspective of CIs, the concepts of place and temporal cues for frequencies, and together with this, for pitch, are the most relevant ones because CIs cannot deliver optimally these cues to the auditory nerve.

The place cue for pitch refers to the perceptual mechanism related to the auditory filters of the basilar membrane. In NH, the basilar membrane acts like a bank of bandpass filters, each filter responding most strongly to a narrow range of frequencies and located at a specific point along the length of the cochlea (in the so-called tonotopic order, described above). Any single sinusoidal tone, having only one frequency component, gives rise to maximum vibration at a specific place along the basilar membrane (Moore, 2003a). However, the bandwidths of the filters on the basilar membrane increase with increasing center frequency (Moore, 2003a). For low frequency harmonics of complex tones, the bandwidths are sufficiently narrow that each harmonic gives rise to a specific peak on the basilar membrane, i. e., these harmonics are resolved. In areas responding to higher frequencies, the filter bandwidth spans several harmonics, so that each place (filter) responds to several harmonics. Thus, the higher harmonics do not give rise to specific peaks, and they are unresolved on the basilar membrane. The series of local peaks for resolved harmonics on the basilar membrane, and the harmonic relationship between these peaks, provides place cues for pitch (f_0) calculation (e.g. Moore, 2003a, 2014). This calculation is possible even though the f_0 may be missing, allowing identification of the pitch of sounds over the telephone or other sound environments where low frequencies are attenuated or missing (He & Trainor, 2009).

The nerve spikes induced by the resolved harmonics tend to be phase locked or synchronized to the stimulating waveform, i, e., when spikes do occur, they occur at approximately the same phase of the waveform. For a single sinusoid, the timing of the phase-locked responses encodes the period of the tone. This phase-locking provides also a temporal fine structure code for the frequency of each resolved harmonic of complex tones. For resolved lower harmonics, the frequency of each is encoded by the phase-locked firing, and the harmonic structure, and hence f_0, is encoded in the ensemble of fine timing information across these harmonics. However, the temporal information carried by the pattern of firing becomes increasingly imprecise above approximately 2 kHz (Moore, 2008, 2014). For higher, non-resolved harmonics, the movement on the basilar

membrane reflects the sum of several harmonics, and thus shows the same periodicity as the input sound waveform (f_0). Phase-locked responses to peaks in this complex basilar membrane vibration will thus also reflect f_0. Hence, when only unresolved high harmonics are present in a tone, and there are no place cues to pitch, the temporal envelope of the basilar membrane response to the summed unresolved harmonics is the only available cue to pitch. When the temporal envelope code is the only peripheral cue to pitch, discrimination of changes in f_0 is rather poor (Moore, 2003a, 2014).

Each peripheral model (place or temporal) may explain some, but not all, aspects of pitch perception. For example, in the periphery of normal auditory system, the pitch of complex tones may sometimes also be derived from combined place and temporal cues (Luo et al., 2012).

The effective number of electrodes of CIs is often less than the actual number of electrodes (12 to 22 in current devices) due to the spread of electric current from active electrode to adjacent places (Abbas et al., 2004; Chatterjee & Shannon, 1998). Even if there was minimal current spread and all electrodes conveyed independent information, the level of detail of the representation of the sound spectrum would be much less than that provided by the number of filters in the normal inner ear. Therefore, not even the lower harmonics of complex tones are resolved with CIs (Drennan & Rubinstein, 2008; Moore, 2003a), and the peripheral coding of cues for pitch of complex, periodic tones is highly limited with CIs. Except for the special case of isolated low frequency sinusoidal tones, CIs do not allow phase-locked auditory nerve responses to individual harmonics. Further, most CIs (like those using CIS or ACE coding strategies) filter out fine temporal structure above few hundred Hz in the envelope extraction process. Since all harmonics are normally unresolved with CIs, the envelopes extracted by the CI speech processor from pitch-bearing sounds will reflect the sum of several harmonics, and thus will tend to reflect f_0. Thus, the peripheral temporal coding of the cues for the pitch of complex sounds for the CI listener depends entirely on a temporal cue comparable to that for normal listeners when a complex sound contains only high (non-resolved) harmonics (Geurts & Wouters, 2001; Laneau & Wouters, 2004: Moore, 2003a; Ping et al., 2012). Unfortunately, this cue is difficult to detect for f_0:s above 300 Hz (Green et al., 2002; Laneau et al., 2004), and CI users also seem to have difficulties in binding the temporal cue to the place cue (Chatterjee & Oberzut, 2011; Limb & Roy, 2014).

These limitations of CIs lead to consistent difficulties for CI listeners in the perception of pitch even in a single tone (monophonic) musical context (CI adults and adolescents, Leal et al., 2003; McDermott & McKay, 1997; Petersen et al., 2015; Pijl, 1997; Sucher & McDermott, 2007; Timm et al., 2014; Vandali et al., 2005; CI children, Mitani et al., 2007; Nakata et al., 2005; Olszewski et al., 2005; Stordahl et al., 2002) and in speech (Ciocca et al., 2002; Green et al., 2004; Laneau & Wouters, 2004). Even NH listeners, especially those without musical training, sometimes confuse changes in pitch with changes in loudness or timbre (Melara & Marks, 1990a, 1990b; Sucher & McDermott, 1997). This may be common also within CI users (Sucher & McDermott, 1997). Further, changes of the harmonics with changes in pitch (f_0) can cause unusual changes in loudness if the loudness has not been well balanced between the CI channels (for techniques to prevent this in psychophysical studies, like roving or loudness balancing, see for example Chatterjee & Peng, 2008). With CIs, perception of two simultaneous pitches and of melody in polyphonic music is even more challenging than perception of pitch or melody in single tones (Donelly et al., 2009; Galvin et al., 2008, 2009).

Music instrument timbre and speech sounds. As with the perception of pitch, the perception of musical timbre is degraded for CI listeners (adults and adolescents, Gfeller et al., 2002; Nimmons et al., 2008, Petersen et al., 2015; children, Stabej et al., 2012; for a review, McDermott, 2004; Limb & Roy, 2014). For NH listeners, the acoustic cues for perception of differences of timbre between musical instruments involve the spectral envelope, spectral fine structure and intensity envelope (attack time; Caclin et al., 2005). In addition, NH listeners can use these temporal and spectral cues both independently and in combination (Caclin et al., 2005). CI users perceive musical instrument timbre mainly from the intensity envelope (attack time; Kong et al., 2011; McDermott, 2004; Timm et al., 2012). However, some adult CI users can learn to weight the acoustic cues for musical timbre similarly to NH listeners, at least with training (Macherey & Delpierre, 2013).

CI listeners also have difficulties in the perception of differences in spectral shape that distinguish different speech sounds. In speech, the positions of the tongue and other structures (like lips and jaw) during vocalization induce peaks in the sound spectrum at specific frequencies, called formants, and these define largely the vowel quality and vowel identity (Stevens, 1998). The restrictions of the CI in delivering the spectral shape

(Moore, 2003a) lead to difficulties in determining the phoneme quality from the formant structure (Välimaa et al., 2002a; see also Geers et al., 2003). CI users also have difficulties in the perception of consonants pronounced at different articulation places, cued by transitions of formants (Donaldson & Kreft, 2006; Välimaa et al., 2002b).

Loudness. The peripheral mechanisms underlying perception of loudness are not fully understood. In NH, loudness may depend however on a summation of neural activity across frequency channels, and depends largely on the rate of neural firing in the auditory periphery (neural firing rate) (Moore, 2003a). Above a certain sound level, any individual neuron will cease to respond to an increase in sound level with an increase in firing rate; the neuron is saturated. The range of sound levels between threshold and the level at which saturation occurs is called the dynamic range. There are three types of auditory neurons encoding loudness in the auditory system. Each of them has different dynamic ranges. The neurons with high spontaneous firing rates have a narrow dynamic range. The neurons with medium spontaneous rates have slightly higher thresholds and wider dynamic range than those with high spontaneous firing rates, and the neurons with low firing rates have the lowest thresholds and so-called sloping saturation, where the increase in firing rate is at first rapid but slows down at higher levels. The variation in these rate vs. level functions is related to the type of the synapse of the neurons with the inner hair cell. Moreover, the neurons with wide dynamic ranges probably play a crucial role at high sound levels. The wide dynamic range of these neurons is probably dependent on the compression that happens on the basilar membrane, related in turn to the functioning of the outer hair cells (Moore, 2003a).

In CIs, the sound level is coded by pulse magnitude or duration, or by analog current. Increase in any of these leads to increases in neural spike rates. The increase is very rapid as a consequence of the bypass of the compression of the basilar membrane, and the absence of delay due to the lack of neurotransmitter release (Moore, 2003a). Moreover, the auditory nerve fibres stimulated by a given electrode all tend to show the same firing pattern, and when the neurons start firing, they continue firing at a similar rate. Consistent with the findings on firing rates, a small change of pulse level leads to a large change in loudness. Therefore, typically the range of current between the detection threshold and an uncomfortable sensation is very small, in the range 3 to 20 dB. This is much less than

the dynamic range in acoustic hearing (approximately 120 dB). For these reasons, two-stage compression is used in CIs (an automatic gain control system followed by instantaneous compression) (Moore, 2003a; Zeng, 2004).

Duration and gaps. Current CI processing strategies are based mostly on extraction and representation of the temporal envelopes of sounds from the filtered stimulus (McDermott, 2004), making the slow-varying changes in level and spectral shape easy to discriminate. In line with this, discrimination of syllabic duration (Meister et al., 2011; O'Halpin, 2010) and gap detection thresholds (Busby & Clark, 1999; Drennan & Rubinstein, 2008) are typically comparable in CI users and NH listeners. It also seems that the perception of rhythm in music is fairly good, even though not "perfect", in CI listeners (Drennan & Rubinstein, 2008).

1.2 Processing of acoustic cues in the brain

It can be assumed that the cues for music and prosody, although they are different for NH and electric hearing as explained above, are analysed in the brain in similar networks in CI and NH listeners. Evidently, the cortical development of these networks has to be sufficient to enable accurate perception for CI children. In NH, initial pitch analysis is carried out in the medial primary auditory cortex in two mirror-symmetric tonotopic maps (Formisano et al., 2003; Griffiths & Hall, 2012). Further, invariant representations of pitch (independent of musical instruments, voices etc.) seem to be processed in posterior regions of auditory cortex, in planum temporale (PT) (Garcia et al., 2010; Plack et al., 2014). Even for NH listeners the efficient cortical representations (neural networks) for pitch may only emerge during development with exposure to the appropriate sounds (Oxenham et al., 2011).

The basic acoustic features of musical instrument timbres and human speech are processed in core and belt (middle) regions of the auditory cortex (Kumar et al., 2007; Leaver & Rauschecker, 2010; Warren et al., 2005). The spectral envelopes of different sounds are probably encoded in the PT (Kumar et al., 2007). Category-selective subregions for both speech sounds and musical instruments have been identified in anterior superior auditory cortex (Leaver & Rauschecker, 2010). It seems that information flows from primary auditory cortex to PT, which projects to the anterior parts of the

temporal gyrus (Kumar et al., 2007). There is some evidence that the anterior parts of the superior temporal gyrus respond particularly to changes in phoneme categories (vowels, Obleser et al., 2006; consonants, Obleser et al., 2007).

Changes in loudness are probably coded in auditory cortex by neuronal populations that are non-randomly distributed in the isofrequency dimension orthogonal to the primary tonotopic axis (Woods et al., 2009). Medial auditory cortical fields may be more responsive to stimuli with higher intensities than more lateral ones (Brechmann et al., 2002; Woods et al., 2009).

Perception of time-related changes seems to rely on widely distributed neural networks, including motoric areas. For example, discrimination of vowel duration activates not only the auditory cortex but also the inferior frontal gyrus and insula (Steinbrink et al., 2012), and the cerebellum is involved in duration interval discrimination (Grube et al., 2010). Moreover, increasing sound duration increases activity in the left anterior insula, right inferior frontal, right middle temporal, and right post-central gyri in addition to bilateral supra-temporal gyri (Ross et al., 2009). PT seems to be important for sensory-motor integration at least in relation to speech and other vocal tract behaviors (Hickock et al., 2009). Perception is often multisensory, as indicated by, for example, the effect of visual (lip-reading) cues on the perception of speech sounds (McGurk & MacDonald, 1976). Activation of the PT can be seen during lip-reading, reading written language, piano score reading and observation of finger movements on a piano keyboard (key-touch reading), the latter only for highly skilled musicians. Thus it seems that the PT is involved in the multisensory integration of well-learned auditory-visual couplings in general (Hasegawa et al., 2004).

1.3 Effects of early deafness: Cortical reorganization after sound onset and attention

After the 27th fetal week, the ear can transmit sounds to the cortex, and exposure to sounds can lead to long term memory representations of them. This has been found for exposure to both speech and music (Partanen et al., 2013a, 2013b). During this period, myelination, essential for rapid synchronized conduction, occurs through the brainstem up to auditory thalamus (Moore & Guan, 2001), and sound deprivation can affect this process (Moore & Linthicum, 2007). Furthermore, the dendritic tufts and axons in the cortical marginal

layer (later layer 1) develop during this period (Moore & Guan, 2001). Sound deprivation during this period can thus lead to deficiencies in the development of layer 1 (McMullen & Glaser, 1988; McMullen et al., 1988). Importantly, the layer 1 axons seem to run across the cortical surface, carrying stimulation to other cortical areas. Moreover, the activating influences of layer 1 on deeper cortical layers probably last until adulthood (Moore & Guan, 2001). Clinical evidence suggests a deficit in attention to auditory stimulation in congenitally deaf CI children (Houston et al., 2003), which may be partially related to a deficit in early development of the marginal layer (layer 1) (Moore & Linthicum, 2007).

Sound deprivation from birth to the switch-on of the CI can also have consequences for the development of the auditory system. Towards the age of six months after birth, the multilayered structure of the auditory cortex begins to develop (Moore & Guan, 2001). According to animal studies, myelination, essential for this process, is sensitive to activity levels (Barres & Raff, 1993). Therefore, deafness during this period can result to subnormal myelination, affecting further the early construction of cortical columns. Moreover, after birth, development of the cortical networks of deaf infants relies on visual, tactual and proprioceptive stimuli, the latter also from the speech apparatus since deaf infants cry aloud, vary their pitch to some extent, and even produce speech-like sounds (Oller & Eilers, 1988). For CI children, the auditory cortex is sometimes abnormally activated by visual or tactile stimulation, implying cross-modal reorganization due to deafness, and harming auditory performance (Sharma et al., 2015). Deafness can lead to decoupling of the auditory system from other senses and poor sensory integration even though it seems that early implantation (before approximately 2:5 years) allows integration of visual and auditory cues together (Schorr et al., 2005). Further, the increase in white-matter in association cortices, important for the maturation of auditory orienting, is already strong before the age of 8–12 months in normal-hearing children (Kushnerenko et al., 2013, for a review). Therefore, missing auditory input even within the first years of life may harm the neural basis of attention to sounds.

Electrophysiological measurements have shown that the brain of newborn NH babies responds to changes in prosody (Sambeth et al., 2008) and to changes in rhythmic aspects of sound sequences (in beat patterns) implemented through omission of sounds (Winkler et al., 2009). Further, the brain of 4 month old NH infants responds to changes in pitch of tones with a missing fundamental (He & Trainor, 2009). In NH infants less than one year

old, behavioural experiments conducted with a head-turn procedure have shown that these infants respond to changes in melodic contour (Trehub et al., 1987), can categorize auditory sequences on the basis of rhythm or tempo (Trehub & Thorpe, 1989), and can infer meter from patterns of rhythms (Hannon & Johnson, 2005). Also a listening preference study has given evidence on that by seven months of age infants learn to distinguish the rhythmic patterns of music (strong and weak beats inducing meter) implemented through changes in intensity (Phillips-Silver & Trainor, 2005). So, early-implanted children begin building up the neural networks for all of these auditory aspects, including the acoustic cues for music, much later than NH children, and the building up may be affected by changes in the auditory system due to deafness and degraded input from CI.

It is however clear that the auditory system reorganizes dramatically after the activation of the CI, especially if the child has been implanted within the first 3.5–4.0 years of life (Ponton et al., 2000, 2001; Sharma et al., 2002, 2009; for a review, Kral & Sharma, 2012). For the reorganization of networks for processing acoustic cues, early-implanted children with CIs may need to focus their attention specifically towards them. Auditory cortex is affected especially by behaviourally relevant stimuli under focused attention. For example, if ferrets are trained to detect a pure tone within a series of sounds, the cortical responses specific for the behaviourally relevant target tones are rapidly facilitated in the primary auditory cortex (Fritz et al., 2003). Conversely, Norena et al. (2006) found that if the enriched acoustic environment was not informative for the animals, the information led to habituation of the primary auditory cortex responses. Attention towards sounds (or lack of it) also modulates activation in auditory cortical areas in humans (Fritz, 2007; Woods & Alain, 2009; Woods et al., 2009). In the rehabilitation of hearing-impaired and CI children it has been emphasized that the child's awareness of sounds is the first step towards auditory learning (Cole & Flexer, 2011, p. 189), and that the missed parts of spoken language should be brought directly to their attention (Cole & Flexer, 2011, p. 91). The role of attention has been noticed and may play a crucial role in the cortical reorganization of CI children, and the deficits in the neural networks for attention, if such exist, may play a crucial role here.

1.4 Perception of word and sentence stress

The perception of prosody plays an important role in language acquisition. English-speaking infants aged 7.5 months rely on stress-based cues in the segmentation of words from fluent speech (Houston et al., 2004; Jusczyk et al., 1999[1]; Mattys et al., 2005, for a review), and at later stages their segmentation performance is assisted by the exaggerated prosody of infant-directed speech, where the parents mark the important words by using sentence stress (Thiessen et al., 2005). Further, better processing of word stress in infancy leads to better spoken language skills at later ages (Friedrich et al., 2009). Even in adulthood, NH listeners use prosodic word stress patterns in word segmentation (Vroomen et al., 1998). Word segmentation and word learning is also supported by phonotactic, acoustic-phonetic information (like coarticulation or vowel disharmony) and lexical information (Kuhl, 2004; Mattys et al., 2005; Vroomen et al., 2008). However, if the listener has difficulties in hearing the phonotactic or acoustic-phonetic cues, or if the language skills are only emerging or restricted, the stress cues override the other cues in segmentation of words (Mattys et al., 2005). The CI children have difficulties in recognition of phonemes, discrimination of detailed acoustic-phonetic cues and, like all children or even more, restricted language skills. Therefore, stress cues, if accessible, are likely to remain important for their language learning throughout their childhood[2].

Later-implanted children show deficiencies and great individual variability in the perception of sentence stress (O'Halpin, 2010) and of word stress (Lyxell et al., 2009; O'Halpin, 2010), although they seem to develop stress perception (O'Halpin, 2010) on a similar but delayed trajectory to typically developing children (Vogel & Raimy, 2002; Wells et al., 2004). Their difficulties are evidently partially a consequence of their difficulties in perception of pitch (f_0). However, stress patterns are also signaled by changes in duration and intensity (e.g., Kochanski et al., 2005; Lieberman, 1960; Meister et al., 2011; Vainio & Järvikivi, 2007). CI listeners are also disadvantaged over NH

---

[1]In these studies metrical stress i. e weak-strong vs. strong–weak stress patterns, was used in the experiments. This can be signaled with vowel reduction together with pitch, duration and intensity cues,
[2]Word stress is usually in the beginning of the word in languages like Finnish, English and Dutch, and therefore plays in these languages an important role in word segmentation. However, in languages like French, where word stress is not in the beginning of the word, other ques play more important role (Vroomen et al., 1998; Mattys et al., 2005).

listeners in the perception of intensity changes, as reviewed above. Variations in the ability to detect changes of pitch (f_0) and intensity may thus affect the prosodic perception of CI users (Meister et al., 2011; O'Halpin, 2010). It is not known how accurately early-implanted children can perceive stress or the abovementioned acoustic cues. More studies are needed on these aspects and into the links to abilities to perceive the acoustic cues to stress in early-implanted children.

1.4.1 Auditory working memory

The speech perception, language and reading skills of CI children are strongly associated with performance in the forward digit span task where the child has to repeat numbers (Harris et al., 2013; Pisoni & Cleary, 2003; Pisoni et al., 2011). For CI children, the performance in this task is more strongly connected to the language skills than the performance in backward digit span task. Compared to NH children, they also show poorer performance in forward digit span task than in backward digit span task (Pisoni et al., 2003). This makes it important to study the development of the CI children especially in the forward digit span, which is traditionally thought to measure the so-called phonological loop subcomponent of working memory. The term working memory refers to the temporary storage and manipulation of information, and the functions involved in the integration of incoming information with information in existing memory stores (e.g., Baddeley, 1992). The phonological loop subcomponent is thought to be a verbal storage system composed of a short-term phonological store plus a subvocal rehearsal processes (Baddeley, 1996; Baddeley et al., 2003). However, a good performance in forward digit span correlates with good discrimination of pitch (Seppänen et al., 2012) and larger and earlier event-related responses (P300) to pitch changes, thought to reflect updating of auditory working memory (George & Coch, 2011). Performance in forward digit span task is thus related not only to phonological processing but also to the functioning of the central executive component of working memory (Alloway et al., 2004; Engle et al., 1999; George & Coch, 2011). It is not known how performance in the digit span task is related to stress perception or discrimination of acoustic cues by CI children. It is also not known how performance in the digit span task, and auditory working memory components related to that, develop in early-implanted children, although performance in

digit span task is typically poorer in later-implanted children than in NH children (Harris et al., 2013; Pisoni et al., 2011).

1.5 Music

Musical activities seem to be a powerful tool for enhancing auditory perception from the level of the brain to the behavioural level (Wan & Schlaug, 2010). Self-production may play a key role in this effect: the plastic changes in the brain related to pitch or other sound encoding are induced more efficiently with active exposure to music than only by listening to sounds (Pantev & Herholz, 2011). For instance, Hyde and colleagues (2009) showed that compared to control children, 15 months of musical training (keyboard lessons) of 6-year-old children led to enlargement of the corpus callosum, auditory and motor cortices. Similarly, compared to non-musicians, in adult musicians several sensory, motor, and higher-order cortical areas as well as regions in the hippocampus, cerebellum, and corpus callosum are enlarged (Herholz & Zatorre, 2012; Jäncke, 2009; Pantev & Herholz, 2011). Adult musicians also show enhancements in the architecture of various white matter tracts, important for cortico-cortical connections (Bengtsson et al., 2005; Halwani et al., 2011; Imfeld et al., 2009). Musical training early in life seems to be particularly effective, inducing stronger plastic changes in the brain than musical activities beginning later in the life (Herholz & Zatorre, 2012).

In line with these neural changes, cross-sectional studies show that compared to musically non-trained NH listeners, musically trained NH listeners have enhanced behavioural perception of pitch for both speech and music (adults: Deguchi et al., 2012; Parbery-Clark et al., 2009; Schön et al., 2004; Tervaniemi et al., 2005; children, Magne et al., 2006; Marques et al., 2007) and of pitch when timbre is varied (i.e., invariant perception of pitch) (Pitt, 1994). Musicians also show enhanced perception of the timbre of musical instruments and human voices (Chartrand & Belin, 2006), of speech syllable duration (adults: Marie et al., 2012), of musical rhythm and meter (adults: Geiser et al., 2009), and of emotional prosody (adults: Lima & Castro, 2011). Moreover, they show enhanced auditory working memory (adults: George & Coch, 2011; Parbery-Clark et al., 2009; children, Strait et al., 2012) as well as visual and auditory attention skills (children, Kraus et al., 2012; Strait et al., 2012). Results from longitudinal intervention studies show

that musical training improves NH children's perception of sentence intonation (Moreno et al., 2009), emotional prosody (Thompson et al., 2004), verbal memory (Ho et al., 2003; Roden et al., 2012) and auditory working memory (Fujioka et al., 2006). These experimental studies appear to show that enhancements are attributable to musical training rather than to genetic or environmental factors (Besson et al., 2011). The findings that the younger the age at which musical training begins, the larger is the extent of the specific anatomical differences between musically trained and non-musically trained listeners, further support the view that musical training enhances cortical development and through this, auditory perception (for a review, Münte et al., 2002).

For adult CI listeners and CI children, musical training seems to benefit the perception of musical pitch (Chen et al., 2010), melodic contour, musical timbre, and general music perception (Petersen et al., 2012; Yucel et al., 2009). However, it is not known how early-implanted children benefit from musical activities.

Parental singing is known to be an important way of regulating the emotions and state of arousal of infants and young children (Rock et al., 1999). Consistent with this, singing arouses the attention of children with CIs and is used in speech therapy sessions (Ronkainen, 2011). It is also recommended for rehabilitation of music perception of children with CIs (Rocca, 2012). Singing could play a special role in CI children's auditory attention and through this, in neural plasticity related to music perception (section 1.3).

It is also important to address the question of why the CI children sing. It is possible that parental singing at an early age plays a role here. For example, the experiences from the Lindfors Foundation speech-music groups (lindforsinsaatio.net/lindfors-foundation-speech-music-groups/) imply that CI children begin to sing at home if the parents are encouraged to sing at home with them right after implantation. However, there is no scientific evidence on this so far.

1.5.1 Are music and speech perception connected via rhythm?

Traditionally music and speech have been thought to be processed in different areas in the brain, music in the right hemisphere and speech in the left hemisphere (Tervaniemi & Hughdahl, 2003). However, in adults, music and speech activate overlapping neural

regions in superior, anterior and posterior temporal areas, temporoparietal areas, and inferior frontal areas (Abrams et al., 2011; Koelsch et al., 2002; Rauschecker & Scott, 2009; Rogalsky et al., 2011; Schön et al., 2010; Tillmann et al., 2003), including also Broca's and Wernicke's areas in the left hemisphere that were previously thought to be language-specific. Moreover, newborns show overlapping neural activity in response to infant-directed speech and to instrumental music (Kotilahti et al., 2010). These findings indicate that processing of music and speech are connected in the brain.

Previously, it has been found for NH listeners that perception of pitch and lexical tones in speech is connected to perception of pitch and melody in music, and musical training advances perception of pitch and intonation in speech (Jiang et al., 2010; Liu et al., 2010; Magne et al., 2006; Marques et al., 2007; Moreno et al., 2009; Nan et al., 2010; Patel et al., 2005, 2008; Schön et al., 2004). These findings imply that perception of music and speech is linked in the domain of pitch. Rhythm also has important functions in both music and speech. Both are systems which are dependent on how acoustic events unfold over time (Cason & Schön, 2012). Moreover, some findings already support an association between the perception of musical rhythm and speech. For instance, Marie et al. (2011) found that musicians process the lengthening of the final syllable of sentence more accurately than non-musicians. Further, priming with musical meter improves phonological processing of speech (Cason & Schön, 2012), and synchronizing musical meter and linguistic stress in songs enhances processing of both lyrics and musical meter (Gordon et al., 2011).

It has already been shown that, for CI listeners, good perception of music, especially of timbre, melody and pitch, is related to good perception of speech (Drennan & Rubinstein, 2008; Wang et al., 2012). If perception of word stress were associated with better perception of musical rhythm, this would open up new perspectives for further studies on CI children and their rehabilitation.

1.5.2 Music and visuospatial perception

Importantly for children with CIs, visuospatial processing has been recently linked to music perception. A stimulus-response compatibility effect has been found between the pitch (high/low) of auditory stimuli and the location (up/down) of the answer button (Rusconi et al., 2006), and musicians' abilities in visuospatial perception have been

shown to be better than average (Brochard et al., 2004; Patston et al., 2006). Thus perception of musical pitch may be spatial in nature (Rusconi et al., 2006). However, further studies are needed. If visuospatial perception were correlated with music perception, this would have implications for rehabilitation of music perception of CI children.

1.6. Event-related potentials

The neurocognitive functions and neural plasticity related to music perception can be measured with event-related potentials (ERPs). ERPs are gathered with electro-encephalography (EEG), measuring the dynamics of electric field potentials generated by neuronal activity in the brain. EEG reflects the post-synaptic potentials of neurons which are oriented in parallel and activated synchronously (Luck, 2005). Auditory event related potentials are brain responses to sounds, formed by averaging the EEG segments, resulting in attenuation of the activity that is not temporally synchronous and preservation of the time-locked activity (Picton, 2010). The adult auditory ERP waveform in response to a sound onsets consists of a series of peaks. They are labelled based on the polarity of the peak (P for positive, N for negative) and temporal order as P1 (around 50 ms from stimulus onset), N1 (100 ms), P2 (180 ms), and N2 (250 ms) (Luck, 2005; Picton, 2010). These ERPs reflect processing in the auditory cortex (N1, Näätänen & Picton, 1987; P2, Crowley & Colrain, 2004; N2, Näätänen & Picton, 1986). Each peak of the ERP waveform reflects a contribution from several functions or neural processes, which are also called subcomponents (Näätänen & Picton, 1987).

The latencies and amplitudes of auditory ERPs can provide temporally fine-grained information about sound-evoked neuronal activity. This information can be linked to the stages of sound processing, from the early encoding of sound properties in the auditory brainstem to later, higher-order processes such as attention and memory at the cortical level (Luck, 2005; Picton, 2010). The later, more cognitive components like mismatch-negativity (MMN) and following positive P3a are usually recorded using the so-called oddball paradigm. In this paradigm, an occasional deviant stimulus is inserted into a repeating sequence of standard sounds. MMN and P3a can be extracted from the ERP difference signal, which is formed by subtracting the ERP signal for the deviating

auditory events from the ERP signal for the repeating, standard sounds (MMN: Näätänen et al., 2007; P3a: Alho et al., 1998; described in more detail later in this section).

ERPs can be measured in passive listening situations where the subject is not required to pay attention to sounds (as in the present thesis). This makes the technique well suited to young children (Kujala & Näätänen, 2010). ERPs can give information about brain plasticity. The enlargement of the response is probably based on the involvement of new neurons due to learning (Kujala & Näätänen, 2010; Kujala et al., 2007). So far, ERPs are the best way to directly measure neural plasticity of neural networks in individuals with CIs, since the metal in the inner parts of CI makes the use of other brain imaging methods very demanding and even dangerous.

P1. According to Ponton and Eggermont (2001), positivity of the P1 response is consistent with a relatively deep sink (in cortical layers IV and lower III) and a superficial current return, and the generators may include thalamo-cortical loops and primary and secondary auditory areas (Sharma et al., 2007). For NH subjects, the latency of P1 becomes shorter with increasing age as also is the case for CI children (Alvarenga et al., 2013; Sharma et al., 1997, 2002a, 2002b). For recently implanted children, the P1 responses are prolonged (Ponton et al., 1996a, 1996b; Sharma et al., 2002a, 2002b), which is consistent with hypomyelination in their auditory system (Moore & Linthicum, 2007). The P1 latency of early-implanted (before 3:5 years) children seems to reach the normal range between 3 and 6 months after implantation (Sharma et al., 2002a, 2002c, 2005). This rapid shortening of P1 latency may reflect a resumption of myelin formation driven by axonal activity (Moore & Linthicum, 2007).

The P1 amplitudes for CI subjects vary with stimulus parameters, making the comparison between CI and NH listeners' P1 responses hard. For example, Kelly and colleagues (2005) found that the P1 amplitude of CI users reduced with increasing pure tone frequency, and P1 amplitude was smaller at 4 kHz for the CI group than for the NH group, while it was similar between groups at 1 kHz. Further, in the previous studies, the stimulus has usually been electric for CI children and acoustic for NH listeners, again making it difficult to compare and interpret the development of P1 for CI and NH children. For example, Ponton and Eggermont (2001) used acoustic clicks for NH children and electric pulses delivered directly to the electrodes for CI children, bypassing

the speech processor. They found that P1 was larger for the CI group than for the NH group. However, using speech presented in free field, via the CI processor, and electric pulse trains delivered directly to the electrodes, bypassing the CI processor, the P1 amplitude has been found to decrease over time for CI children (speech: Alvarenga et al., 2013; electric pulses: Jiwani et al., 2013), and the "abnormally" large P1 amplitude for electric pulses seem to decrease to similar values as P1 for acoustic stimulus in NH children after 10 years of CI use (Jiwani et al., 2013). There are no studies on the P1 for music instrument sounds in early-implanted children.

MMN. The mismatch negativity (MMN) reflects how the listener can predict the regularities in the auditory environment and how sound changes violate these perceived and remembered regularities (Kujala & Näätänen, 2010; Kujala et al., 2007; Näätänen et al., 2007; Winkler et al., 2009). This theory holds that the MMN is a result of a comparison between the features of the incoming sounds and the sound features predicted from a memory model of the invariant aspects of the auditory environment. Some theories assume that a neuronal correlate of the memory trace for the standards is a simple stimulus-specific adaptation of auditory cortical neurons to repeated stimuli (Nelken & Ulanovsky, 2007; May & Tiitinen, 2010). These theories are however controversial (Näätänen et al., 2005, 2011). Recent computational models suggest that more complex prediction and comparison processes as well as adaptation are necessary to explain the MMN (Garrido et al., 2009). The latency and amplitude of the MMN also reflect the perceptual difference between the deviant and standard and discrimination accuracy (Näätänen et al., 2007).

The MMN has been elicited in CI recipients, reflecting discrimination ability and cortical plasticity after implantation (Lonka et al., 2004; Ponton et al., 2000; Sandmann et al., 2010; Timm et al., 2014; for a review, Johnson, 2009). Even though the MMN increases and becomes earlier with better behavioural performance, it can sometimes reflect only soon-to-appear behavioural skill, i. e., it can be recorded prior to behavioural discrimination ability becoming apparent (for a review, Kujala et al., 2007). Therefore the MMN is not directly comparable to behavioural discrimination. Importantly for studies of children, MMN elicitation does not require motivation, and concentration skills play a less important role in MMN elicitation than in behavioural tasks.

The main cortical generators of the MMN are located in the auditory cortical areas (Alho et al., 1996; Kropotov et al., 1995; Levänen et al., 1996; Opitz et al., 2002; Tervaniemi et al., 2000). An additional contribution from the frontal cortex (Alho et al., 1994; Giard et al., 1990; Rinne et al., 2000; Schönwiesner et al., 2007) and parietal areas (Takahashi et al., 2013) has been shown, implying a wide neural network for MMN elicitation. It has been assumed that the auditory cortex generators reflect memory trace formation and comparison processes while the frontal source is involved in triggering involuntary attention to sound changes (Näätänen et al., 2007).

Musically trained NH children show enhanced MMN for pitch (f_0) changes in violin tones (Meyer et al., 2011), for changes from major to minor chords (Virtala et al., 2012), and for pitch and voice onset time (VOT) changes in speech (Chobert et al., 2011). Compared to musically non-trained children, longitudinal studies show more MMN enhancement in musically active children for melodic and rhythmic modulations, mistuning and timbre (Putkinen et al., 2014), and for syllable duration and voice onset time changes (Chobert et al., 2014). There are no studies on MMN for changes in musical tones or effects of musical activities of MMN in early-implanted children.

P3a. The MMN for deviant tones can be followed by a P3a response, which reflects an involuntary attention switch towards a salient change in the auditory environment (Alho et al., 1998; Escera & Corral, 2007; Escera et al., 1998; Wetzel et al., 2006; in CI recipients, Kelly et al., 2005; Kileny et al., 1997; Nager et al., 2007). Shifting of attention brings potentially important information into focus, allowing re-evaluation of the entire situation (Horváth et al., 2008). This is in contrast to the pre-attentive detection of deviant events reflected by the MMN (Friedman et al., 2001; Tremblay et al., 1998; van Zuijen et al., 2006). P3a responses may be also related to updating auditory working memory (Barcelo et al., 2006), i. e., a central executive component related to updating the items held in working memory by replacing old information with new, more appropriate information (Miyake et al., 2000). P300 (P3b) responses to target sounds become larger and earlier with increasing forward digit span, which suggests that P3b reflects updating of working memory (George & Coch, 2011; Polich et al., 1983). Interestingly, Barcelo and colleagues (2006) found that familiar sounds that signaled the need to change the rule in a task and occasional task-irrelevant novel sounds activated a similar neural (P3a)

network and disrupted behavior in a similar way (see also Barcelo et al., 2002). They also concluded that novelty P3a may reflect updating of working memory, and proposed a similar function for P3a to deviant events. However, this proposal has so far not been assessed in the context of changes in musical tones. Importantly, very little is known about the attention functions of early-implanted children, even though these might be affected by early deafness (section 1.3) and are proposed to be highly important for perception and learning of degraded auditory stimuli with CIs (Beer et al., 2011; Houston et al., 2014; Wild et al., 2012).

Several brain areas seem to underlie the P3a: frontal areas (Løvstad et al., 2012; Schröger et al., 2000; Takahashi et al., 2013; Volpe et al., 2007), auditory cortical areas (Alho et al., 1998; Opitz et al., 1999, Takahashi et al., 2013), temporo-parietal junction (Knight & Scabini, 1998), parietal areas (Takahashi et al., 2013), and hippocampus (Knight, 1996). It is worth noting that the frontal component of MMN seems to be separable from the frontal component of P3a, peaking earlier for MMN than for P3a (Schönwiesner et al., 2007). Evidently, the neural networks for MMN and P3a are separable functionally and statistically (Takahashi et al., 2013).

Like the MMN, P3a becomes larger with increasing physical difference between the deviant and standard (Wetzel et al., 2006; Winkler et al., 1998), and for CI children P3a becomes larger and earlier with improving speech recognition (Kileny et al., 1997). P3a has been used in several studies to assess whether musical training enhances attention functions. Augmented P3a has been found for adult musicians (Brattico et al., 2013; Trainor et al., 1999, Vuust et al., 2009) and for children with high amounts of informal musical activities, including singing, at home (Putkinen et al., 2013). Similarly, P3a has been shown to occur earlier for musically trained participants (Nikjeh et al., 2009).

2 Aims and hypotheses

The main aim of the present thesis was to investigate the differences and similarities between early-implanted children and NH children in the perceptual and cognitive skills or processes underlying perception of music and of word and sentence stress. Another aim was to assess whether and how musical activities might assist CI children in achieving better perception, auditory working memory and attention functions.

More specifically, Study I investigated how CI children differ from NH children in the neurocognitive processing of changes in musical tones (in P1, MMN or P3a). We tested hypothesis: (I) CI children have smaller and/or later P1, MMN and P3a than NH children, especially for the MMN and P3a, for changes in timbre and pitch.

Study II assessed the interplay between the development of neurocognitive processing of music and hearing status and singing of CI children during a time period of between 14 to 17 months. Singing at home was chosen to be the criterion for dividing the CI children into musical activity groups for several reasons. The musical activities of the CI children themselves comprised mainly singing, and we expected that cortical development was affected more by regular motoric training than by pure listening (Pantev & Herholz, 2011). We also expected that singing has a specific role in the development of auditory attention shift reflected in P3a responses. It is also evident that the early onset of musical activities is essential for strong effects in the brain (Herholz & Zatorre, 2012). Therefore, the CI singing groups (see section 3.1.1) were formed on the basis of the regularity of musical activity (singing) in the home setting and the time they had sung before the study began. Here we tested two hypotheses: (I) CI children have smaller and/or later MMN and P3a than NH children for changes in timbre and pitch: the differences between groups become smaller over time. (II) The MMN and P3a is/becomes larger and/or earlier in CI children who sing regularly at home compared to other CI children. We had an additional hypothesis III (not presented in the publications included in the thesis): Larger and/or earlier P3a responses are associated with longer digit spans. An additional hypothesis IV (not presented in the publications included in the thesis) was: Singing of CI children is related to the singing of the parents in early years of the hearing life of the CI children.

Study III compared development of the perception of word and sentence stress and associated auditory cues as well as auditory working memory for CI and NH children

(also during 14 to 17 months), and assessed the role of auditory discrimination of pitch (f_0), intensity and duration as well as auditory working memory and supervised music group activities in perception of stress within CI children. Feedback, challenging situation (like the presence of simultaneous sounds) requiring good concentration skills as well as tasks provided by the group leaders were expected to be important for the development of performance in the behavioural tasks. With regard to the development of auditory working memory, training leading to improved digit span performance typically involves visuospatial cues, is designed to become more demanding during the course of training, and includes feedback (in NH children: Klingberg et al., 2005; in CI children, Kronenberger et al., 2011). These aspects are typical of supervised group activities but not for singing by oneself. Therefore, in Study III the CI children were divided into those who attended supervised musical activities outside of the home and those who did not. Within CI children, we tested three hypotheses: (I) Prosodic perception is related to auditory discrimination abilities; (II) Prosodic perception is related to auditory working memory; (III) Prosodic perception is associated with musical activities. We also hypothesized that auditory working memory develops better in CI children attending supervised musical activities than in other CI children.

Study IV investigated the associations between perception of music and word stress and between visuospatial perception and music perception in NH adults. We hypothesized: (I) Perception of music, particularly perception of rhythm, improves with improving perception of word stress; (II) Perception of music improves with improving visuospatial perception.

3 Methods

3.1 Participants

In Studies I–III, the participants were 4–13-year-old Finnish-speaking unilaterally implanted CI children and NH children (Table 1). Inclusion criteria for the CI children were: CI activation prior to three years one month; no diagnosed developmental or linguistic problems; more than 6 CI electrodes in use; no re-implantation between measurements in the case of longitudinal Studies (II and III). All of them had been using their implants for at least 22 months prior to the first measurements, had full insertion of the electrode array, attended mainstream school or day care, and communicated with spoken language. They did not benefit from residual hearing in the unimplanted ear.

The NH children were healthy and without linguistic or hearing problems. Their hearing had been screened at child welfare clinics and according to the parents reports the hearing of the children was normal. In all studies the NH groups were matched to the CI groups at the group level by age, gender, and handedness as well as social and musical background using questionnaires filled in by parents and personnel at schools or day care concerning the children's musical and other hobbies and musical activities at home, school and daycare centres (the questionnaires are presented here: http://www.cbru.helsinki.fi/music/RitvaTorppa/).

Parents of all participating children gave written informed consent prior to testing and the participants gave consent orally after the study was explained to them. All studies were carried out in accordance with the Declaration of Helsinki, and the procedures for Studies II–II were approved by the local ethical committees of the participating hospitals. In Study I, 24 CI children filled the initial inclusion criteria. Only 22 CI children were included in the final analysis because data recorded from two CI children had to be excluded due to problems in the quality of ERP responses. Twenty two NH children were matched to this CI group (Table 1).

In Study II, 21 CI children fulfilled the initial inclusion criteria. The same 22 NH children as in Study I served as a control group. In Study III, 21 CI children fulfilled the inclusion criteria, and 21 NH children were matched to this CI group (Table 1).

Table 1. The details of the participating children used for statistical analyses.

ID[1]	Age at T1	Hand[2]	Music[3]	SE[4]	Aetiology[5]	Age at CI switch-on (months)	CI use prior T1 (months)	CI processor type[6]
CIs/m 01	5y 11m	R	20(betw)	R	U	18	53	NF
CIs/m 03	9y 2m	R	12(betw)	R	U	32	77	MT
CIs/m 04	7y 10m	R	24(betw)	R	U	25	69	MT
CIns/CIm 09	7y 4m	R	0(betw)	R	C	19	69	MO
CI 10*	12y 6m	R	12	R	U	32	130	MO
CI 12*	4y 1m	R	24	R	C	15	34	NF
CIs/m 13	5y 5m	R	22(betw)	R	U	18	47	NE
CIs/m 14	4y 4m	R	0(betw)	R	U	18	34	NF
CIs/n 15	5y 1m	R	0	R	C	17	44	NE
CIns/n 16	7y 2m	R	0	R	C	25	61	NF
CIns/n 17	9y 4m	L	0	R	U	19	93	NF
CIns/n 18	12y 1m	R	0	R	U	27	118	NF
CIns/n 19**	7y 5m	R	0	R	U	29	60	NE
CIs/n 20	5y 8m	R	0	R	U	20	48	NF
CIs/n 21	5y 7m	L	0	L	C	19	48	NF
CIs/n 22	7y 1m	R	0	R	U	21	48	NE
CIns/n 23	7y 10m	L	0	R	U	18	76	MT
CIns/m 24	4y 2m	R	23(betw)	R	C	14	36	NF
CIs/m 26	4y 2m	R	23(betw)	R	C	20	30	NF
CIns/n 27	4y 2m	R	0	R	C	13	37	NF
CIs/n 28	6y 2m	R	24	R	U	22	52	NF
CIns/n 29	8y 7m	R	0	L	C	37	66	NF
CIs/n 30	6y 7m	R	0	R	C	25	54	NF
N CI = 23 N CIs = 12 N CIns = 9 N CIm = 8 N CIn = 13		N R+L = 20+3	N attend: before = 9 betw = 8	N R+L = 21+2	N U = 13 N C = 10			N NF = 14 N NE = 4 N MO = 2 N MT = 3
NH 02	7y 11m	R	36(betw)					
NH 03	4y 6m	R	0					
NH 04	8y 2m	R	45(betw)					
NH 05	10y 0m	R	0(betw)					
NH 06	5y 8m	R	0(betw)					
NH 07	6y 9m	R	0					
NH 08	5y 7m	R	0(betw)					
NH 09***	4y 6m	L	42(betw)					
NH 10	4y 0m	R	0(betw)					
NH 11	5y 6m	R	0					
NH 13	5y 0m	R	35(betw)					
NH 14	4y 6m	R	15(betw)					
NH 15***	12y 0m	R	0					
NH 16	8y 5m	R	0					
NH 17	9y 8m	R	0					
NH 18	6y 9m	R	0					
NH 19	7y 0m	R	0					
NH 20	4y 6m	R	12					
NH 21	6y 5m	R	15					
NH 22	6y 11m	R	0(betw)					
NH 23	5y 5m	R	12					
NH 24****	7y 0m	R	0					
NH 30	11y 2m	L	54(betw)					
N NH = 23		N R+L = 21+2	N attend before = 9 betw = 11					

*Included only in Study I.

** ERP data only from T1, excluded from Study I.

*** Excluded from Study III.

****Included only in Study III.

[1] Identification number, CI = CI child, NH = NH child

s = CI singer in Study II,

ns = CI non-singer in Study II,

m = in musically active CIm group in Study III,

n = in musically non-active CIn group in Study III.

[2] Hand = handedness.

[3] Music = amount of time attending to supervised musical hobbies outside of the home before T1 (months) (dancing excluded),

(betw) = child attended supervised musical hobbies outside of the home between measurements.

[4] SE = stimulated ear.

[5] U = unknown, C = Connexin 26.

[6] NF = Nucleus Freedom (coding strategy: ACE)

NE = Nucleus ESPrit 3G (coding strategy: ACE)

MT = Medel Tempo + (Coding strategy: CIS)

MO = Medel Opus 2 (Coding strategy: CIS).

N = number

For Study IV, sixty four 19-60-year-old Finnish-speaking, NH adults (without musical education at a professional level) were recruited. One participant was excluded because of a deaf ear, one because of weaker than first language level skills in Finnish, and one because of evident congenital amusia, and so 61 were selected for the final analysis. The ethical committee of the Faculty of Behavioural Sciences of the University of Helsinki approved the study and the participants gave their written informed consent.

3.1.1 Division of CI groups into musical activity groups

CI singing groups in Study II. The CI children were divided into two subgroups on the basis of the regularity of their singing in the home and the time they had sung before the Study began, using questionnaires (http://www.cbru.helsinki.fi/music/RitvaTorppa/). According to the answers, 12 CI children sang weekly at home one year before the study began and between T1 and T2 (*"CI singers")*. Nine CI children sang less than weekly or not at all (*"CI non-singers")* (Table 1). According to age-controlled ANOVA, these groups did not differ significantly from each other in the other aspects of home-related musical background as assessed by musical activity clusters (formed with cluster analysis based on the answers to the questionnaire, APPENDIX 1), amount of musical activities at day care or schools, supervised musical activities outside of the home, or factors related to their aided thresholds for hearing or CI devices, age, gender, socioeconomic background, or aetiology.

We also recorded samples of singing ("Tuiki tuiki tähtönen", in English, "Twinkle twinkle little star") of the CI children at T2 (the task was completed by nineteen CI children). A professional singing teacher scored blindly (without knowing whether the child was a CI singer or not) the rhythm, melody and lyrics they sang. It was concluded that the singing of CI children was recognisable and different from general speech. The comparisons between CI singers and CI non-singers showed that the accuracy of production of lyrics, melody and rhythm was better for CI singers than for CI non-singers. Age-controlled ANOVA confirmed that the CI singers were significantly better in production of rhythm ($F_{1,18} = 7.83$, $p = .013$) and in the overall accuracy of singing (the mean of production of lyrics, melody and rhythm) ($F_{1,18} = 5.28$, $p = .035$) than CI non-singers.

Musically active and non-active CI children in Study III. In order to divide the CI children into musically active and non-active groups for Study III, the same questionnaire as for Study II was used. The inclusion criterion for the musically active group (CIm) was participation in instruction of music or dance outside of the home during the course of the present study. Eight CI children met the inclusion criterion. Seven of them had participated in musical activities with an emphasis on singing, together with a parent at an early age. The CI children who did not meet the inclusion criterion were designated CIn (Table 1). Compared to the CIn group, the CIm group demonstrated more time engaged in musical activities and in dancing outside of the home prior to the study and significantly more musical activities in the home (Cluster A, see APPENDIX 1), implying that they also heard and saw others doing music (mainly singing but also some of them music instrument playing) at home more than CIn children. The groups did not differ significantly in the amount of singing by the child at home (Cluster D) or in factors related to their aided thresholds for hearing or CI devices, age, gender, or aetiology. However, the CIm group had a higher level of maternal education.

3.2 Stimuli and procedure for ERP experiments

Stimuli. We recorded ERPs with the multi-feature (MFP) paradigm over a relatively short period of time (Näätänen et al., 2004; Pakarinen et al., 2007). By using the MFP, it is possible to record responses to several types of changes in sounds during a single recording, which is important in order to gain a comprehensive view of auditory processing, which is beneficial in child measurements

Natural sounds were selected from the McGill University Master Samples DVD, edited to the desired duration and normalized in intensity. The standard was a piano tone with f_0 of 295 Hz (duration 200 ms). The deviant tones differed from the standards in pitch (f_0), timbre (Figure 1), duration, intensity increment, intensity decrement or by the presence of a silent gap in the middle of the tone. Each deviant differed from the standard in one of three degrees of change (small, medium and large), leading to 18 deviant tones (Table 2). The deviant tones were similar to the standard in all other features, except for those presented in Table 2, and for the changes in timbre (these contained changes in temporal intensity, spectral envelope and periodicity). In the stimulus sequence every

other tone was a standard and every other tone a deviant. The SOA was kept at 480 ms. The presentation order of the changes was randomized throughout the experiment. The probability of the standard tone was 0.5 and the probability of each deviant tone was 0.028 (Table 2). The standard tone was presented 2250 times and each deviant tone was presented 125 times. The total duration of the experiment was 36 min.
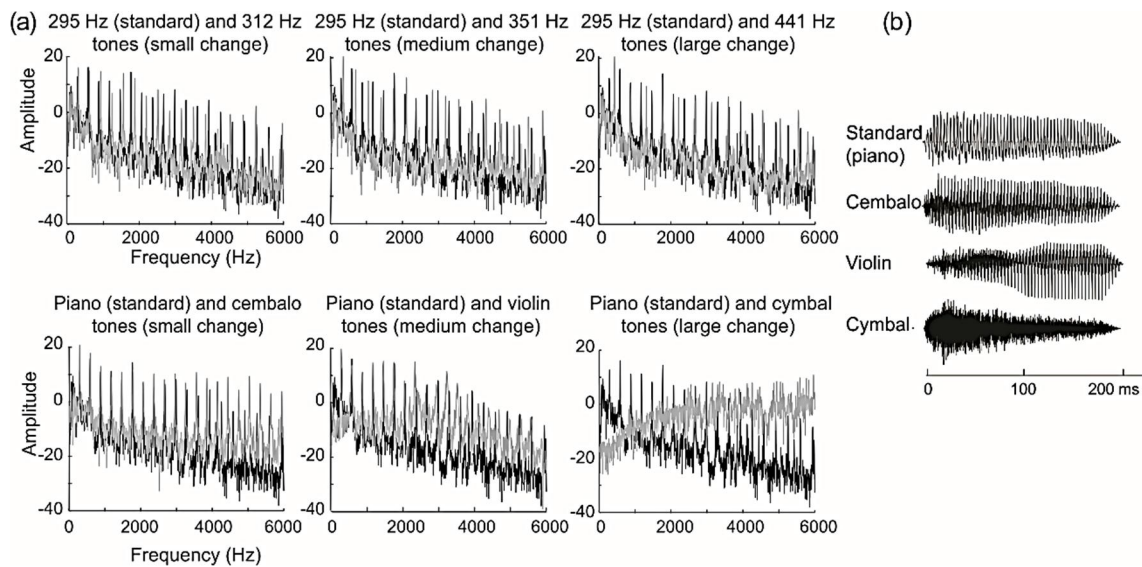
Figure 1. (a). Frequency spectra of the standard tone (black) in comparison to pitch and musical instrument deviants (gray). **(b)** Sound envelopes of the standard piano tone and the musical instrument deviants. The Figures have been reprinted with permission from Elsevier.

Table 2. Stimulus parameters in ERP experiment

Change type	Change amount	f_0 (HZ)	Intensity NH (dB)	Intensity CI (dB)	Duration (ms)	Musical instrument	Silent gap (ms)	Fall time (ms)	Silent interval[2] (ms)
None (std)	None	295	60	70	200	Piano	None	20	280
f_0	S	312	60	70	200	Piano	None	20	280
	M	351	60	70	200	Piano	None	20	280
	L	441	60	70	200	Piano	None	20	280
Intensity increment	S	295	63	73	200	Piano	None	20	280
	M	295	66	76	200	Piano	None	20	280
	L	295	69	79	200	Piano	None	20	280
Intensity decrement	S	295	57	67	200	Piano	None	20	280
	M	295	54	64	200	Piano	None	20	280
	L	295	51	61	200	Piano	None	20	280
Gap	S	295	60	70	200	Piano	5	20[1]	280
	M	295	60	70	200	Piano	40	20[1]	280
	L	295	60	70	200	Piano	100	20[1]	280
Musical instrument	S	295	60	70	200	Cembalo	None	20	280
	M	295	60	70	200	Violin	None	20	280
	L	295	60	70	200	Cymbal	None	20	280
Duration	S	295	60	70	175	Piano	None	20	305
	M	295	60	70	100	Piano	None	20	380
	L	295	60	70	50	Piano	None	10	430

Std = standard. S = small, M = medium, L = large. Probability of each deviant type: 3 x 0.028 = 0.084. Probability of deviants together: 0.5. Fall and rise time of the gap 5 ms. The Table has been reprinted with permission from Elsevier.

Procedure. During the experiment, subjects watched a silent video. All stimuli were presented in an acoustically insulated and dampened room through 2 loudspeakers placed at a 45º angle to each side of the subject, approximately 1 m in distance from the subject's ear, using the everyday settings of the CI. The stimuli were presented at a fixed (comfortable) level, at maximum of 60 dB(A) SPL for the NH group and 70 dB(A) SPL for the CI group. For one CI child the sound level had to be lowered to 65 dB(A) SPL at T1.

The EEG was recorded with Biosemi ActiveTwo amplifier and active electrodes (sampling rate of 512 Hz, low-pass filtering at 102.4 Hz) using a 64-channel electrode cap. On-line, the data were referenced to the CMS electrode. Off-line, the data were referenced to the electrode at the nose tip. To record eye movements and blinks, additional electrodes were placed at the left and right mastoid. The measurements were performed twice (T1 and T2), 14 to 17 months apart (in Study I, only data from T1 were included).

3.3 Stimuli and procedure for behavioural tests and experiments

An overview of the experimental tests and tasks of the participants is presented in Table 3. The table also defines the number of items, the Study where the test/experiment was used (I-IV) and how many times or when (Study III) that was conducted. The text below describes only the details of the stimuli in the experiments (when necessary), the questionnaires and the procedures.

Perception of stress. The stimuli for perception of stress were recorded from an adult male, an adult female, and two female children aged 7 years and 10 years. The stimulus in the word stress task was either a compound word or a phrase. In the sentence stress task, the child heard a sentence containing three content words, one of which bore prosodically marked narrow focus (the stimuli for the tasks are presented here: http://www.cbru.helsinki.fi/music/RitvaTorppa/) (Table 3)[3].

---

[3]In the word stress perception task, f_0, intensity and duration cues were available for the listeners (Hausen et al., 2013). In Finnish, sentence stress (also called prosodically marked narrow focus) is typically signaled with changes in f_0, intensity and duration (Vainio & Järvikivi, 2007).

Table 3. The behavioural experiments and tests.

Experiment/test	Auditory/visual stimulus	Task of the subject	Study (times repeated)
Perception of stress			
Perception of word stress[1,2]	Natural, recorded compound words and phrases + pictures representing the recorded objects.	Point at a picture representing "KISsankello" or "KISsan KELlo" (BLUebell " or "BLUe BEll"). 48 items for children aged > 6 years, 36 for aged < 6 years. 30 items for NH adults.	III (2x, at T1/T2), IV (1x)
Perception of sentence stress[2,3]	Natural, recorded sentences + pictures representing each word in the sentence.	Point at a picture representing the most important word in the sentence "POIKA maalaa veneen" ("The BOY paints the boat"). 48 items.	III (2x, at T1/T2)
Discrimination of acoustic cues			
Discrimination of intensity, duration and pitch (f_0), i, e., acoustic cues for stress [2,4]	Synthesized /tata/ syllable pairs + pictures representing same and different.	Judge if the /tata/ syllable pairs are same or different either by pointing at corresponding picture or orally. An adaptive procedure for 71% correct discrimination threshold, varying number of items.	III (2x, at T1/T2)
Pitch perception test by Hyde and Peretz (2004) (shortened adaptation)	Sine wave tones.	Judge if all five tones are similar or if there is a change in pitch. 80 trials (40 similar, 40 different).	IV (1x)
Auditory working memory			
Digit Span subtest of the ITPA	Natural speech (face to face).	Recall number sequences in the same order as in the original sequence. Varying number of items.	III (2x, at T1/T2)
Digit Span subtest of the WAIS-III	Natural speech.	Recall number sequences in the same/reverse order. Varying number of items.	IV (1x)
Nonverbal intelligence, PIQ			
Block design subtest of the WISC-IV	Red and white blocks.	Order the blocks based on the model you see. Varying number of items.	III (1x, at T2)
Music perception			
MBEA computer based scale subtest[5]	Melodies played with piano.	Judge if the two melodies are similar or different. 30 trials (15 same, 15 different).	IV (1x)
MBEA on-line Off-beat subtest[6]	Melodies played with varying instruments.	Judge if the melody contains an unusual delay. 24 trials (12 congruous, 12 incongruous).	IV (1x)
MBEA on-line Out-of-key subtest[6]	Melodies played with varying instruments.	Judge if the melody contains an out-of-tune tone. 24 trials (12 congruous, 12 incongruous).	IV (1x)
Visuospatial perception			
Discrimination of Gabor patches	Gabor patches proceeding from left to right.	Judge whether the two paths are similar or different. 30 trials (15 similar, 15 different).	IV (1x)

[1-4] Task based on: [1]Vogel & Raimy, 2002, [2]O'Halpin, 2010, [3] Wells et al., 2004, [4]Straatman et al., 2010.
PIQ: Performance intelligence quotient, WAIS-III: Wechsler Adult Intelligence Scale III (Wechsler, 1997), WISC-IV: Wechsler intelligence scale for children, 4th edition (Wechsler, 2010), ITPA: Illinois test of psycholinguistic abilities (Kirk et al., 1974), MBEA: Montreal Battery of Evaluation of Amusia, [5]Peretz et al., 2003, [6]Peretz et al., 2008.

Discrimination of acoustic dimensions. In the discrimination of acoustic cues for stress each trial comprised either two identical ("TAta"/"TAta") or two different ("TAta"/"taTA") patterns, created with the KLATTSYN-88 software synthesizer (Klatt, 1980) and the Speech Filing System (SFS) software (Huckvale, 2012; http://www.phon.ucl.ac.uk/resource/sfs/) (the stimuli for the tasks are presented here: http://www.cbru.helsinki.fi/music/RitvaTorppa/).

For testing intensity discrimination, the stimuli had intersyllable level differences ranging between 1 and 15 dB. All disyllables had an identical f_0 pattern and the syllable duration was fixed at 300 ms. For testing discrimination of syllable duration, the duration

of the two syllables varied, the total duration of each disyllable being always 600 ms. The duration ratio between syllables ranged from 1.02 to 2.38. The only variation in f_0 was the steady declination, as in the intensity series. In the two tasks measuring the discrimination of pitch (f_0), the f_0 pattern comprised two components: a rise-fall representing syllable stress and the same gradual declination as used in the series described above (Figure 2). The onset f_0 of the rise-fall was either 160 Hz (female f_0 range) or 295 Hz (child f_0 range). The peak in f_0 at the mid-point was higher than at onset according to 48 equally spaced multiplicative factors from 1.013 to 1.84. This rise-fall f_0 pattern was then summed with the declination component which, as above, had a linear fall in f_0 such that the f_0 at syllable offset was 94% of the f_0 at syllable onset. Because a preliminary analysis showed that pitch (f_0) discrimination thresholds did not differ between the two f0 ranges, the thresholds were averaged over the two f_0 ranges for further analyses.
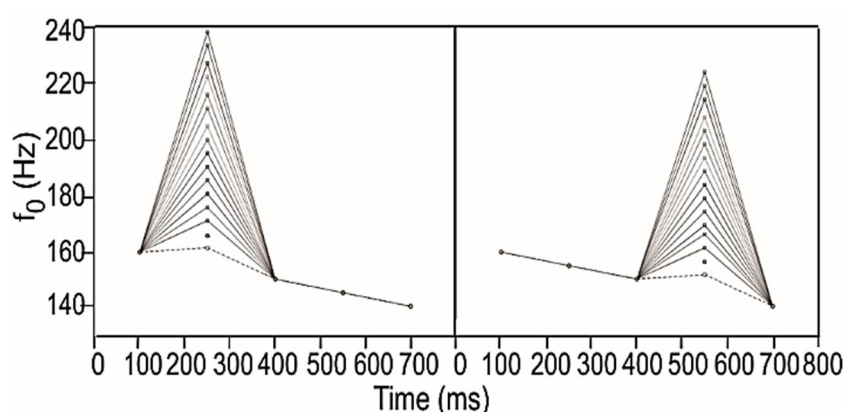
Figure 2. Example f_0 contours for the pitch (f_0) discrimination task (160-Hz baseline).

The pitch (f_0) discrimination ability of adult NH participants was assessed with a computer-based pitch perception test. The duration of each tone was 100 ms and the inter-tone interval was 350 ms. In the standard sequences the f_0 of all tones was C6 (1047 Hz) and in the other sequence types one (fourth) tone was altered by 1/16, 1/8, 1/4, 1/2, or 1 semitones (3, 7, 15, 30 or 62 Hz) upward or downward from C6 (Hausen et al., 2013, Supplementary audio files 3, 4 and 5).

Auditory working memory. Digit span tasks (Table 3) were used as a measure of auditory working memory.

Music perception. Music perception was tested with three online, computer-based subtests of the Montreal Battery for Evaluation of Amusia (MBEA; Peretz et al., 2003, 2008). In the Scale subtest the melodic difference was an out-of-scale tone (approximately 4.3 semitones apart from the original pitch). In the Off-beat subtest, in the incongruous trials there was a time delay, i. e., a silence of 5/7 of the note duration (i. e., 357 ms) in the melody (the tone began later than it was expected). In the Out-of-key sub-test, in the incongruous trials the melody had a tone that was outside of the key of the melody, sounding like a "wrong note" (http:// www.brams.umontreal.ca/amusiademo/).

Visuospatial perception. This task represented a visuospatial analog of the MBEA Scale subtest. The stimuli were created using Matlab and Psychophysics Toolbox extension (Brainard, 1997). In each trial the participants were presented with two series of Gabor patches (contrast 75%; spatial frequency ca. 0.8 c/°; size approximately 2°) proceeding from left to right. There was a 500-ms pause between the two paths. In the paths, a single Gabor was presented at a time (there was a 50 ms pause between two Gabors, the duration of each Gabor varied). The path was formed by simultaneously changing the position and the orientation of each Gabor relative to the preceding Gabor. The orientation of the Gabor followed the direction of the path. On half of the trials the two Gabor paths were identical. On the other half the second path was changed. In change trials the second series had one Gabor that deviated from the expected path. The task of the participant was to judge whether the two paths were similar or different. Each Gabor was analogous to a tone in the melody of the Scale subtest. Every semitone difference in the melody was equivalent to a 12° difference in the Gabor orientation/location, except for the deviant Gabor that had a 22° location change for each semitone.

Questionnaires. The parents of the participating children as well as the personnel in schools and daycare centres filled in questionnaires (section 3.1.1, http://www.cbru.helsinki.fi/music/RitvaTorppa/). The adult NH subjects filled a computerized questionnaire (Peretz et al., 2008) and a paper questionnaire (see Hausen et al., 2013, Data Sheet 1). In these, the participants were asked about their musical and educational background, cognitive problems, musical abilities and hobbies.

Procedures. For the CI group and part of the NH control group the perceptual tasks and forward digit span were performed in an acoustically isolated and dampened room. For part of the NH control group these tasks were performed in a quiet room in the participant's home. For both child groups nonverbal intelligence was measured in a quiet test room. In perceptual, recorded tasks, sounds were delivered for children with a laptop through two powered loudspeakers placed at a 45 ° angle to each side of the subject, and 70 cm distant from the subject's ear at a comfortable level (averaging 60 dBA for NH and 70 dBA for the CI group, measured at the pinna). All sounds were presented for CI children using the everyday settings of the CI.

The place of testing of NH adults was arranged individually for each participant: most assessments were done in a quiet workspace at a public library. The computer-based tests were conducted using laptops and headphones. The volume level was adjusted individually to a level that was clearly audibly to the subject.

3.4 ERP Data analysis

Basic analysis in Studies I and II. EEGLAB 8 (Delorme & Makeig, 2004) was used. Imported data were downsampled to 256 Hz, and high-pass filtered at 0.5 Hz. Because of the location of the CI device, some channels could not be used; data from these electrodes were interpolated. The analysis epoch was 550-ms long, starting 100 ms before the onset of the tones. The baseline level of the epochs was set to be zero during the 100 ms before the tone onsets.

Ocular and muscle artifacts were removed for both CI and NH groups using independent component analysis (ICA) with the Fastica algorithm (Makeig et al., 2004). In addition, ICA was used for the CI group to reduce the CI-related artifact. Data dimensionality was narrowed down by the number of interpolated channels and automatic epoch rejection at a threshold between ± 300 and ± 400 µV (individually adjusted to preserve at least 85% of original epochs for effective statistical analysis) was performed before ICA. After ICA, the epoch voltage rejection was done again with a threshold of ± 150 µV, followed by the analysis of the proportion of remaining epochs for each individual subject. The criteria of 75% (95) remaining epochs for each deviant was used

to include individual children in further analysis. One child with a CI did not reach the criterion, and was excluded from Study I and Study II at T1. The mean percentage of acceptance of epochs at T1 was 94% in the CI group (119 deviants, 2348 standards) and 93% in the NH group (116 deviants, 2330 standards), and at T2 was 93% in the CI group (116 deviants, 2330 standards) and 95% in the NH group (119 deviants, 2348 standards).

We calculated the median instead of average of ERP signals (Yabe et al., 1993), because the median method is optimal in cases where the data in general are of high quality, but some extreme values are expected due to liberal rejection criteria or other factors (Fox & Dalebout, 2002; Yabe et al., 1993). After this, we inspected again the individual ERP waveforms. Another child with a CI was excluded from analysis from Study I because of abnormally shaped responses (amplitudes exceeding in the range of MMN -20 µV) (this child was not included in Study II). The data were offline-filtered with a 25 Hz low-pass filter.

Further ERP data analysis for Study I. Data only from T1 was included. CI and NH groups were divided to two age groups: younger or older than 6 years 9 months. The baseline was set to be zero during 100 to 350 ms (whole period).

For ERP quantification, group-level peak latency of the response was determined at the Fz (P1 and MMN) or Cz (P3a) electrodes. P1 was identified as the maximum (most positive) peak occurring in a 70–140 ms time window. MMN was identified as the minimum (most negative) peak within the time window 90–250 ms after change onset, and P3a as the maximum peak within the time window 145–300 ms after change onset. The corresponding mean amplitudes were calculated for each subject from electrodes of interest (F3, Fz, F4, C3, Cz and C4) using a 60-ms (P1) or 40-ms (MMN and P3a) time window surrounding the peak latency of the age group. Because no clear differences in scalp distribution of the responses for electrodes of interest were found, amplitudes were then averaged over the aforementioned electrodes in order to reduce noise. Response amplitudes were subjected to one-sample, two-tailed t-tests in order to examine whether they differed significantly from zero for the CI and NH groups.

For ERP latency quantification, the individual peak latencies were calculated in a specified time window in relation to change onset, only for those responses that were found to be significant. The window was 85–250 ms for timbre and pitch (f_0) MMN, 100–

250 ms for gap and duration MMN, 100–300 ms for intensity decrement MMN and 145–350 ms for P3a. The latencies of responses for intensity increments were not analysed due to different processing between CI and NH groups.

Further ERP data analysis in Study II. The data from both T1 and T2 were used. The signals from F3, Fz, F4, C3, Cz and C4 channels were averaged to form a ROI (region of interest) channel. The baseline was set to be zero during the 50-ms period before the tone onsets.

The group-level peak latency for MMN and P3a was determined for the ROI difference signal (deviant minus standard) within the same time windows as for Study I for the entire CI and NH groups (age division was not performed). The mean amplitudes were calculated using a 30-ms time window surrounding the peak latency. For the NH group, the intensity increment MMN and P3a responses were not analysed due to different processing between CI and NH groups.

Similarly to Study I, ERP response amplitudes were subjected to one-sample, two-tailed t-tests. The individual peak latencies were calculated for the significant responses from the ROI-signal in a similar time windows as in for Study I except for the intensity increment and decrement MMN. For these, the window was set at 100–400 ms. In order to compare MMN and P3a between CI and NH groups or between CI singers and CI non-singers, we analyses the responses using the following principles. The response for the specific deviant type was included in the analyses if the MMN/P3a was significant at T1 and/or T2 for the both tested child groups.

3.5 Statistical analyses

In Study I, the mean amplitudes and peak latencies were compared between CI and NH groups and age groups by repeated-measures analysis of variance (ANOVA). A Greenhouse-Geisser correction was used when appropriate. The analyses were conducted separately for each change type.

For Studies II and III, the statistical analyses used linear mixed modeling (LMM: Singer & Wilett, 2003; West, 2009). Due to the large variability of age of the child participants, age was controlled for. In addition, for Study III maternal education was

controlled for because the CIn children had lower level of maternal education than the CIm children. We also tested the covariance structures and selected the best fitting ones based on Akaike's and Bayesian information criteria (AIC and BIC). For Studies I and II, the statistical analyses were conducted separately for each change type because the magnitudes of the changes were not equalized across change types.

For both Studies II and III, the LM models for testing hypotheses I and II included measurement time, age, and one or more hypothesized predictors of the dependent measure, as shown in the tables in the Results section. The additional hypothesis III for Study II was tested with LMM similar to that was used for testing hypothesis I, but with digit span as an additional independent variable. The additional hypothesis IV for Study II was tested with partial correlation analyses (age controlled). Because the responses to questions addressing parental singing were included in the cluster A (APPENDIX 1), we ran partial correlation analyses between the amount of singing of the CI child at home and the answers falling inside the cluster A.

For Study III, a set of small models was selected to test specific hypotheses. All non-significant interactions were omitted from the final results reported in the tables in the Results section. For Studies I-III post-hoc tests were conducted when necessary, and, for these, Bonferroni correction was used.

For Study IV, the associations between the MBEA scores and background variables possibly affecting the connections of music perception to word stress perception or visuospatial perception (age, pitch perception/discrimination, musical and general education as well as forward and backward digit span) were first examined using t-tests, ANOVAs, and Pearson correlation coefficients depending on the variable type. The variables that had significant associations with the music perception scores were then included in further analysis. Pitch discrimination thresholds calculated from the pitch perception test and auditory working memory were also controlled for when examining the associations of word stress and visuospatial perception with music perception. Linear step-wise regression analyses were then conducted to examine how much the different variables explained the variation of the music perception total score and subtest scores.

For all Studies I–IV, the level of significance was set at 0.05 and the analyses were performed using the current version of SPSS (also called PASW in Studies I and IV).

4 Results

4.1 Cortical processing of musical sounds for CI and NH children

The aim of Study I was to compare the CI and NH groups in the ERP responses (P1, MMN and P3a) to acoustical changes in musical sounds, reflecting the efficiency of the processing of piano tone onsets and the efficiency of the cortical networks for neural discrimination and auditory attention shift.
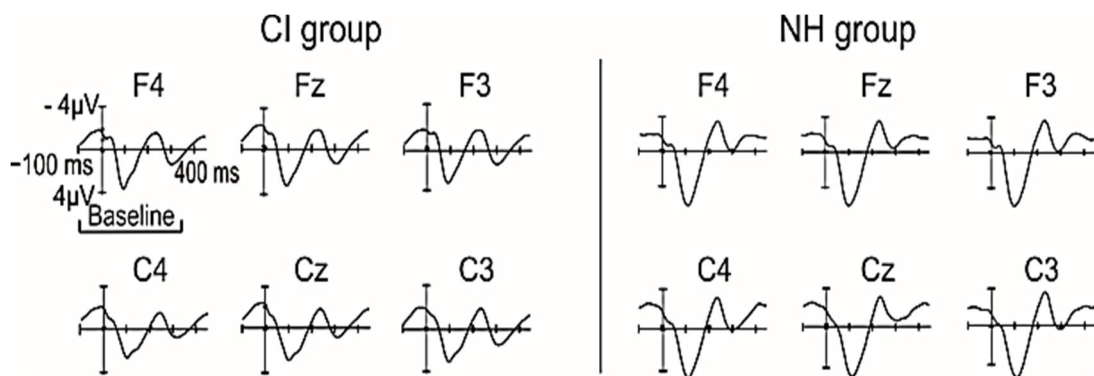
Figure 3. Standard waveforms over the frontocentral scalp regions of the CI and NH groups.

P1 with N2 and without N1 response was elicited for both CI and NH groups (Figure 3). Moreover, early MMN was followed by early P3a for the large change in timbre and for changes in pitch (f_0) in both groups (Figures 4a,b). Timbre MMN for small and medium change was non-existent for the NH group while the P3a for these changes was elicited for both groups (Figure 4a). The gap, duration and intensity decrement changes elicited MMN for both groups (Figure 4c,e,f). ERP responses for intensity increments differed between CI and NH groups. In NH group we observed a pattern of P3a followed by large reorienting negativity (RON) responses (Escera & Corral, 2007; Figure 4d). In CI group intensity increments did not elicit P3a or RON responses. Because of these substantial differences between groups, the group comparisons were not conducted.
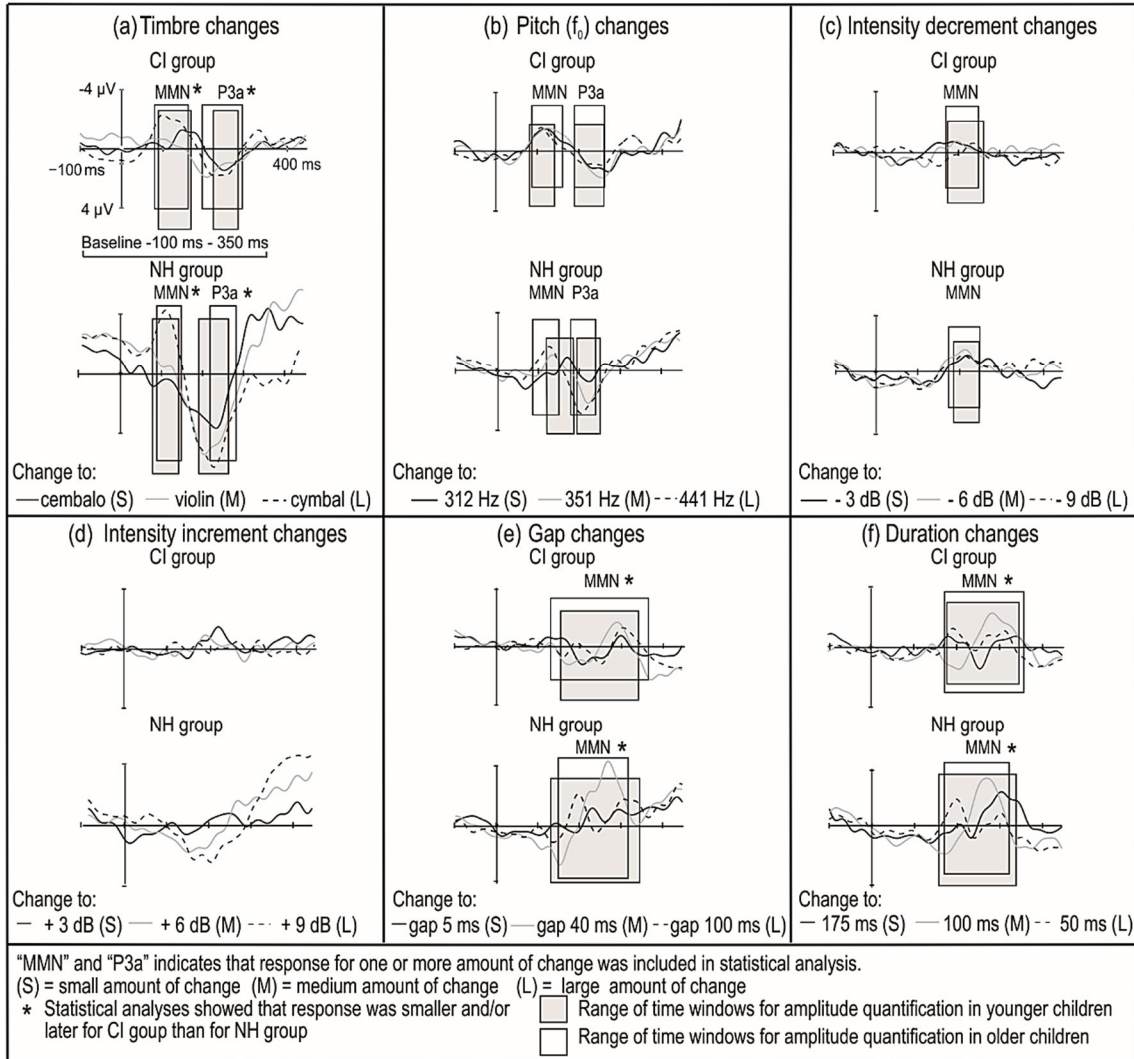
Figure 4. The subtraction (deviant - standard) waveforms at Fz electrode for CI and NH group for **(a)** timbre changes, **(b)** pitch (f_0) changes, **(c)** intensity decrements **(d)** intensity increments **(e)** gap changes and **(f)** duration changes.

Table 4. Significant results from CI vs. NH group comparisons.

	P1		Timbre MMN (L)		Timbre P3a (S,M,L)	Gap MMN (M,L2)		Duration MMN (S,M)	
	Amplitudes	Latencies	Amplitudes	Latencies	Amplitudes	Amplitudes	Latencies	Amplitudes	Latencies
	F	F	F	F	F	F	F	F	F
Group	28.00***	19.20***	10.36**	6.23*	14.81***	ns	ns	9.35**	8.25**
Age	ns	6.80*	4.32*	ns	ns	ns	ns	ns	ns
Amount	-	-	-	-	7.35**	6.82*	4.12*	10.88**	13.11***
Amount × group	-	-	-	-	4.34*	ns	13.50***	ns	ns
Amount × age	-	-	-	-	3.68*	ns	ns	ns	ns

Group = CI vs. NH group. Age = younger vs. older children. Amount = amount of change. Following the response type, in parentheses the amount of change included in analysis: S, M, L = small, medium, large amount of change. - = interaction or amount of change was not included in repeated-measures ANOVA. ns = result was not significant. (°p≤.1, *p≤.05, **p≤.01, ***p≤.001).

For group comparisons, P1 was smaller and earlier for the CI group than for the NH group and appeared earlier for older children than for younger children (Table 4, Figure 3). Compared to the NH group, the CI group had smaller and later timbre MMN (Table 4, Figure 4a), smaller timbre P3a (Table 4, Figure 4a), later MMN to the 40-ms (medium) gap (*amount × group,* Table 4, Figure 4e), and smaller and later duration MMN (Table 4, Figure 4f). Moreover, for timbre P3a, the differences between amount of changes were not significant for the CI group while for the NH group the P3a for the change from piano to cymbal was larger than the P3a for other timbre changes (*amount × group,* Table 4, Figure 4a). Also the main effect of amount was significant (Table 4). The pitch (f_0) MMN or P3a did not differ between groups (Figure 4b).

Further, timbre MMN was larger for older than for younger children, the MMN to the medium gap was larger and earlier than the MMN to the large gap, and the duration MMN was smaller and later for the small than for the medium duration change (Table 4, Figures 4a,e,f).

Summary of findings from Study I. The results from Study I indicate that the musical multi-feature paradigm is feasible for measuring ERP responses to changes in musical sounds for young children. Moreover, there are reliable neurocognitive responses similar to those seen for NH children to changes in most of the key acoustic features of musical sounds for CI children. Their MMN for several change types and their timbre P3a were smaller and/or later than for NH children, implying degraded neural discrimination and less efficient attention shift as a consequence of this. However, the results of Study II changed the picture and two subgroups of CI children were found.

4.2 Interplay between singing and cortical processing of music for CI children

The main aim of longitudinal Study II was to compare the development of ERP responses to changes in musical sounds for CI and NH children and to investigate whether the development (especially of P3a) was better with more singing of the CI children at home. Additionally, we investigated whether P3a response latencies or amplitudes were earlier/larger with better forward digit span (to find evidence indicating that P3a reflects

updating of auditory working memory), and whether singing of the CI children was related to singing of parents early in their hearing life.

Table 5. The MMN and P3a mean amplitudes and latencies in Study II.

Stimulus eliciting the response:		CI group				NH group			
		T1 μV	T2 μV	T1 ms	T2 ms	T1 μV	T2 μV	T1 ms	T2 ms
Timbre	cembalo(S)	-1.06(2.80)°	-0.82(2.46)°	-	-	1.22(2.60)	-0.16(2.59)	-	-
MMN	violin(M)	0.04(2.17)	-0.12(2.30)	-	-	1.56(2.51)	0.61(2.69)	-	-
	cymbal(L)	-2.44(2.82)***	-2.17(2.50)***	126(40)	133(31)	-1.98(2.39)***	-1.80(2.08)***	116(29)	113(23)
P3a	cembalo(S)	1.81(1.98)***	2.20(2.83)**	249(60)	276(54)	3.84(4.40)***	3.00(4.14)**	211(36)	206(39)
	violin(M)	2.82(2.57)***	3.31(2.88)***	218(45)	248(60)	6.29(4.77)***	5.19(4.33)***	215(29)	211(30)
	cymbal(L)	1.81(1.88)***	1.49(2.68)*	247(52)	242(60)	6.57(4.17)***	7.05(2.82)***	243(53)	231(50)
Pitch (f_0)	312 Hz(S)	-1.68(2.69)*	-0.72(1.67)°	147(48)	158(48)	-1.08(2.80)°	-1.52(2.31)**	156(41)	146(31)
MMN	351 Hz(M)	-1.47(1.55)***	-1.37(1.78)**	139(26)	148(41)	-1.75(2.91)**	-1.81(3.89)*	136(35)	135(34)
	441 Hz(L)	-1.46(2.81)*	-1.81(3.26)*	143(44)	135(46)	-1.33(2.55)*	-2.26(2.88)***	131(37)	128(30)
P3a	312 Hz(S)	0.94(1.53)*	1.03(2.58)°	265(72)	307(59)	0.55(3.31)	0.64(2.92)	-	-
	351 HzM	1.49(2.42)*	1.39(2.65)*	266(57)	283(65)	1.84(3.74)*	1.68(3.72)*	237(36)	229(48)
	441 Hz(L)	0.64(1.76)°	1.32(2.60)*	248(80)	274(54)	2.11(2.88)**	2.08(3.27)**	242(69)	237(57)
Intensity	3 dB(S)	-0.43(1.99)	-0.76(1.92)°	-	-	-1.20(2.55)*	-1.52(2.19)**	-	-
decrement	6 dB(M)	-0.82(1.86)*	-0.28(2.14)	255(83)	249(84)	-1.55(2.74)*	-1.64(2.93)*	248(78)	265(78)
MMN	9 dB(L)	-0.19(2.02)	-0.41(2.04)	-	-	-1.29(2.50)*	-2.46(2.65)***	-	-
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)**	-	-	-	-	-	-
Gap	5ms(S)	-0.14(2.23)	-0.68(2.66)	-	-	-0.98(2.89)	-2.34(3.19)**	-	-
MMN	40 ms(M)	-1.64(2.43)**	-1.10(2.95)	176(15)	167(25)	-4.00(4.20)***	-4.04(3.67)***	160(34)	149(27)
	100 ms(L)	-1.24(2.26)*	-0.33(2.42)	166(29)	165(27)	-1.58(3.55)*	-3.00(3.40)***	177(36)	179(34)
P3a	5 ms(S)	0.90(1.90)*	1.12(2.96)	250(65)	260(72)	n	n	-	-
	40 ms(M)	0.98(3.24)	1.04(2.33)*	261(34)	251(34)	n	n	-	-
	100 ms(L)	0.83(2.15)	0.71(2.54)	-	-	n	n	-	-
Duration	175 ms(S)	-0.77(2.06)	-1.29(1.39)***	168(44)	143(28)	-2.60(3.35)**	-3.54(1.93)***	154(45)	150(34)
MMN	100ms(M)	-2.27(2.10)***	-1.55(3.01)*	188(18)	188(34)	-3.07(3.68)***	-4.80(3.81)***	173(16)	181(39)
	50 ms(L)	-1.46(2.78)*	-.92(2.97)	167(38)	175(46)	-1.13(4.25)	-2.89(3.70)***	158(40)	147(32)
P3a	175 ms(S)	0.00(2.21)	0.63(1.32)*	209(29)	214(33)	n	n	-	-
	100 ms(M)	1.19(1.33)***	1.10(3.11)	289(43)	258(60)	0.22(1.56)	0.18(3.83)	-	-
	50 ms(L)	.90(2.19)°	.25(2.80)	-	-	1.51(3.14)*	.21(3.35)	-	-

S, M, L = small, medium and large amount of change. For both time points of the measurements (T1, T2), the mean amplitude (the standard deviation in parentheses) is followed by the significance of the responses (°p≤.1, *p≤.05, **p≤.01, ***p≤.001; two-tailed t-test against zero). Following these, the mean latencies (and standard deviation) of the responses are given. The columns marked with light gray present the amplitude and latency values included in statistical comparisons between CI singers and CI non-singers as well as between the entire CI group and the NH group. The columns marked with dark gray present the values included only in statistical comparisons between CI singers and CI non-singers. - = the mean amplitudes or individual latencies were not analysed. n = the responses were non-existent (wrong polarity in the time window of the response).
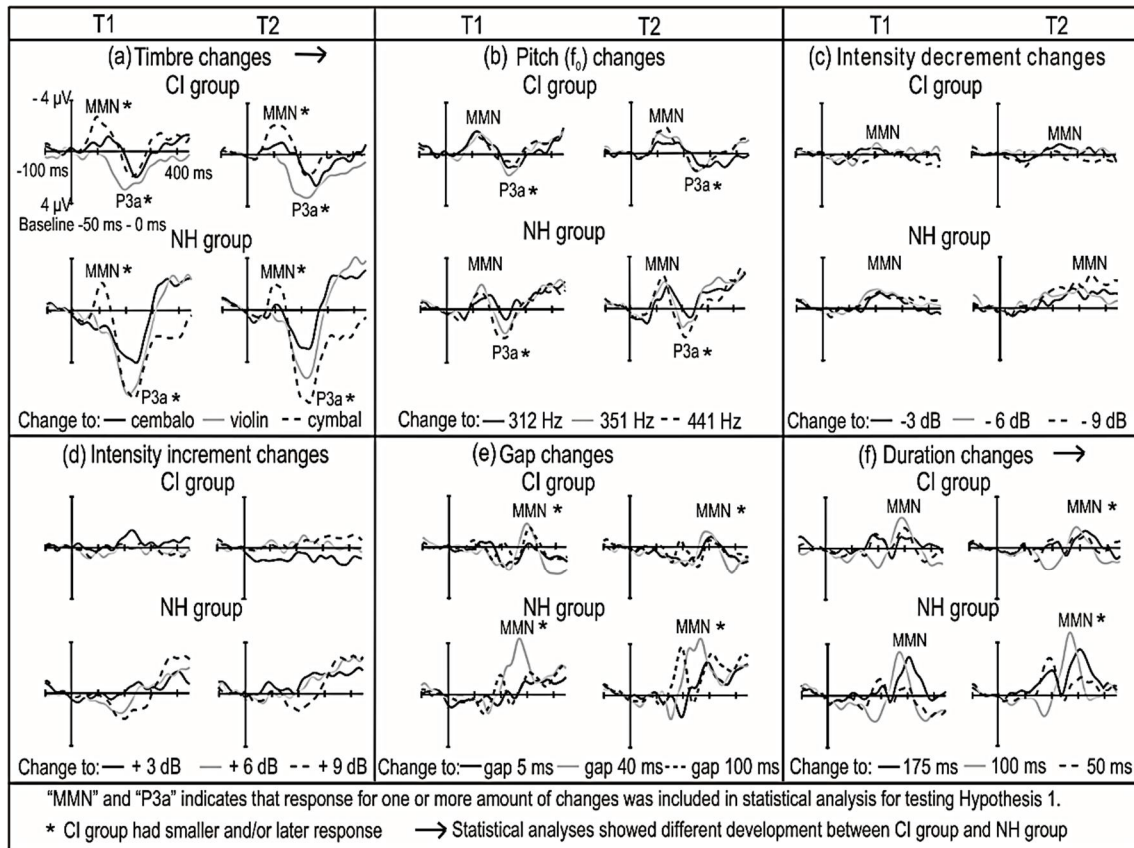
Figure 5. The subtraction (deviant - standard) ROI waveforms averaged across F3, Fz, F4, C3, Cz and C4 electrodes for CI and NH groups for **(a)** timbre changes, **(b)** pitch (f_0) changes, **(c)** intensity decrements **(d)** intensity increments **(e)** gap changes and **(f)** duration changes. These are given for both time points of the measurements (T1 and T2 on the left and right in each panel, respectively).

As found in Study I for the data from T1, the MMN was followed by P3a for the large change in timbre and changes in pitch (f_0) for CI and NH groups at T1 and T2 (Table 5, Figure 5a,b). The ERP responses for intensity increments differed between CI and NH groups at both T1 and at T2 to the extent that it was not possible to conduct the statistical group comparisons (Table 5, Figure 5d). There was more variation between T1 and T2 for the CI group than for the NH group in the MMN for intensity decrements, gaps and changes in duration (Table 5, Figure 5c,e,f), which seemed be a consequence of the variation of the ERPs of CI singers between T1 and T2 (Figure 6c,e,f).

Statistical analyses showed that, as for Study I, compared to the NH group the CI group had significantly smaller and/or later MMN/P3a responses for several change types: later timbre MMN, smaller and later timbre P3a, smaller and later duration MMN, smaller gap MMN (Table 6), and later MMN for the medium gap (*amount* x *group*, Table 6) (Figure 5a,e,f). We also found later pitch (f_0) P3a for the CI group than for the NH group (Table

6, Figure 5b). Timbre P3a became later over time only for the CI group while duration MMN became larger over time only for the NH group (*time x group*, Table 6, Figure 5).

Table 6. Results (unstandardized estimates for main effects) for testing Hypothesis I.

	Timbre MMN (L)		Timbre P3a (S, M, L)				Pitch (f₀) P3a (M, L)		Gap MMN (M, L)				Duration MMN (S, M, L)			
	Latencies		Amplitudes		Latencies		Latencies		Amplitudes		Latencies		Amplitudes		Latencies	
	B	F	B	F	B	F	B	F	B	F	B	F	B	F	B	F
Group	-14.82	4.92*	5.14	17.42***	-19.25	8.30**	-32.01	6.12*	-2.10	8.74**	12.61	.16	-2.49	6.66*	-10.61	4.87*
Time	-1.20	.03	.18	.28	-15.50	.69	-6.09	.55	.04	.01	4.71	1.25	2.40	10.10**	3.71	.74
Amount	-		-1.18[1] 3.64[2]	7.87***	18.20[1] -11.15[2]	.01*	3.21[2]	.70	-1.18[2]	8.68**	40.43[2]	3.34°	-.48[1] -1.35[2]	6.70[1]**	-8.03[1] 20.69[2]	15.80***
Time × group		ns		ns		4.99*		ns		ns		ns		7.65**		ns
Amount × group	-			10.72***		7.81***		ns		ns		12.09***		ns		ns

Group = CI vs. NH group. Amount = amount of change. Following the response type, in parentheses the amount of changes included in analysis: S, M, L = small, medium, large. *B* shows the direction/asterisks the strength of the connection (°p≤.1, *p≤.05, **p≤.01, ***p≤.001). Group: reference is the CI group. Time: reference is the second time point (T2). [1]*B* for small change, reference is the large change. [2]*B* for medium change, reference is the large change. - = interaction or amount of change was not included in LMM. ns = interaction was not significant. Age was always controlled.

In Study I we found very small or non-existent MMN preceding early P3a for small and medium changes in timbre. This suggested that the small MMN was a consequence of the overlap of the early P3a with the MMN. To test this possibility, if in the present Study the MMN preceding the P3a was unexpectedly small, we conducted partial correlation analysis (age controlled) between the amplitudes of the MMN, or the ERP responses in the expected time line of the MMN, and the amplitudes of the following P3a. If the correlation was positive, the MMN became smaller together with the enlargement of the P3a, and the overlap was evident.

For the NH group and the CI singers, the MMN was non-existent for the change to cembalo and to violin (Figures 5a and 6a). As figure 6a shows, in the group level, large MMN was followed by small P3a (for the CI non-singers), and vice versa, small or non-existent MMN was followed by large P3a (for the NH group and CI singers). Therefore, including all groups into correlation analysis was expected to give more information about the direction of the link and stronger correlations between MMN and P3a together with more participants in analysis, and all participants were included in correlation analysis. The MMN and P3a amplitudes were correlated positively (at T1, cembalo, $r_p =$ .48, *p* =.001; violin, $r_p =$ .65, *p* < .001; at T2 violin, $r_p =$ .49, *p* = .001), suggesting a co-dependence and a possibly overlapping MMN and P3a.

Gap P3a was elicited for the CI group only (Figure 5) and so we studied the possibility of overlap only for them. The MMN and P3a amplitudes were correlated positively (small gap, at T1, $r_p = .59$, $p = .008$, at T2, $r_p = .67$, $p = .001$; medium gap, at T1, $r_p = .54$, $p = .018$, at T2, $r_p = .55$, $p = .012$; large gap, at T2, $r_p = .58$, $p = .007$).

At T1, the duration MMN was followed by P3a for both groups while at T2, the P3a was elicited only for the CI group (Figure 5). Therefore, we conducted partial correlation analyses on the T2 data for the CI group. Again, the MMN and P3a amplitudes were correlated positively (small change, $r_p = .64$, $p = .001$; medium change, $r_p = .68$, $p = .001$; large change, $r_p = .79$, $p < .001$).
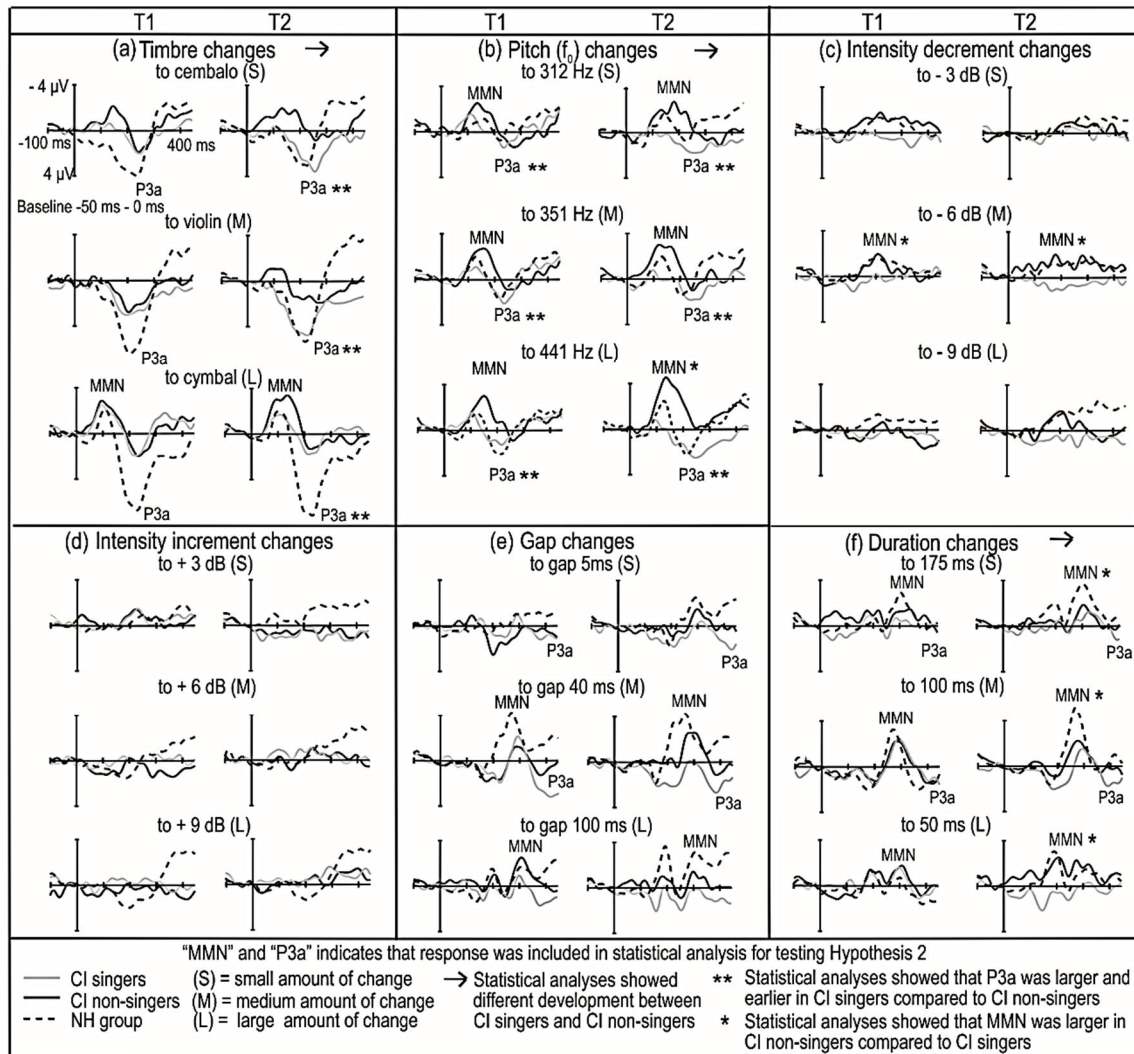
Figure 6. The subtraction (deviant - standard) ROI waveforms averaged across F3, Fz, F4, C3, Cz and C4 electrodes for the NH group, CI singers and CI non-singers for **(a)** timbre changes, **(b)** pitch (f₀) changes, **(c)** intensity decrement changes **(d)** intensity increment changes **(e)** gap changes and **(f)** duration changes. These are given for both time points of the measurements (T1 and T2 on the left and right in each panel, respectively).

P3a development was enhanced for the CI singers. The singing of the children divided the CI group into two subgroups having very different development of ERPs. Timbre MMN became smaller over time in the CI singers (*time × group*, Table 7; Figure 6a). In contrast, timbre P3a was earlier for the CI singers than for the CI non-singers; it became also larger over time for the CI singers but smaller and later over time for the CI non-singers and was larger at T2 for the CI singers than for the CI non-singers (*time × group*, Table 7; Figure 6a).

Table 7. Results (unstandardized estimates for main effects) for testing Hypothesis II

	Timbre MMN (L)		Timbre P3a (S,M,L)				Pitch (f_0) MMN (S, M, L)		Pitch (f_0) P3a (S, M, L)				Intensity decrement MMN (M)		Duration MMN (S, M, L)	
	Amplitudes		Amplitudes		Latencies		Amplitudes		Amplitudes		Latencies		Amplitudes		Amplitudes	
	B	*F*	*B*	*F*	*B*	*F*	*B*	*F*	*B*	*F*	*B*	*F*	*B*	*F*	*B*	*F*
Group	4.07	.19	-2.71	4.19°	64.98	7.07*	-4.12	3.52°	-1.52	5.36*	36.30	7.14*	-3.34	14.39**	-1.76	1.49
Time	11.21	10.82**	-1.02	.01	3.69	.62*	-1.28	.12	-.22	.40	-27.05	5.83*	-1.27	1.5	2.38	5.60*
Amount	-		.36[1] 1.43[2]	7.17***	18.20[1] -11.15[2]	5.04**	-.42[1] -.61[2]	1.02	-.01[1] .45[2]	.78	25.15[1] 13.62[2]	1.69	-		.14[1] -.72[2]	2.28
Time × group		13.21**		10.15**		8.81**		5.40*		ns		ns		ns		4.46*
Time × group × amount		-		ns		ns		2.42*		ns		ns		ns		ns
Time × group × age		9.80***		ns		ns		ns		ns		ns		ns		ns

Group = CI singers vs. CI non-singers. Amount = amount of change. Following the response type, in parentheses the amount of changes included in analysis: S, M, L = small, medium, large. *B* shows the direction/asterisks the strength of the connection (°p≤.1, *p≤.05, **p≤.01, ***p≤.001). Group: reference is the CI singing group. Time: reference is the second time point (T2). [1]*B* for small change, reference is the large change. [2]*B* for medium change, reference is the large change. - = interaction was not included in analysis. ns = interaction was not significant. Results for age are given only when that could not be controlled.

Pitch (f_0) P3a was larger and earlier for the CI singers than for the CI non-singers while, in contrast, pitch (f_0) MMN was larger for the CI non-singers than for the CI singers (Table 7, Figure 6b). Further, for the CI non-singers pitch (f_0) MMN became larger over time for the large change, and was significantly larger at T2 for them than for the CI singers (*time × group, time × group × amount*, Table 7; Figure 6b). The pitch (f_0) P3a of the CI non-singers, however, did not become larger over time with the pitch (f_0) MMN.

The CI singers had smaller 6 dB intensity decrement MMN than the CI non-singers (Table 7, Figure 6c). However, for the CI singers, the difference wave was already positive in the time line of MMN at T1 and T2 (Figure 6c), as were the difference waves for medium and large gaps (Figure 6e) and for the large duration change at T2 (Figure 6f). Evidently as a consequence of the early positivity (P3a), the CI singers also had smaller duration MMN at T2 than the CI non-singers (*time × group*, Table 7; Figure 6f).

P3a was earlier with longer digit span. We found that when the timbre P3a was earlier, then the forward digit span was longer (B = -6.15, p = .004) (Figure 7). For pitch (f_0) P3a latencies there was a significant interaction of *amount* and *digit span* (B = -2.18, p = .030): the P3a for medium change was significantly earlier with longer digit span (r_p (age controlled) between mean T1/T2 digit span and mean T1/T2 P3a latency for medium change = -.376, p = .015) (Figure 7). The other interactions with P3a latency (including those with CI vs. NH group) or connections to P3a amplitudes were not significant.
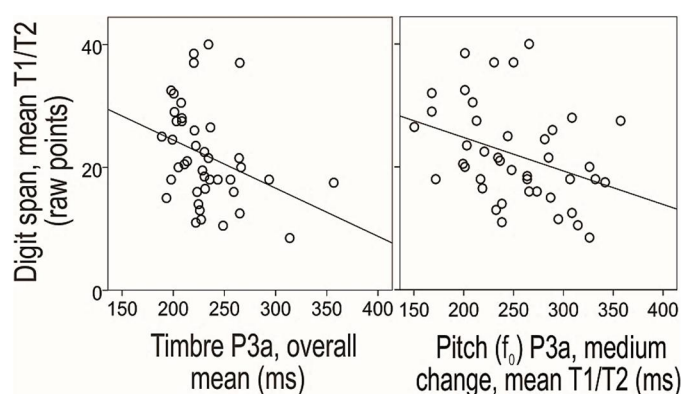
Figure 7. The relationship of digit span to the latency of timbre P3a and medium change in pitch (f_0).

Singing of the CI children was related to singing of the parents. It was found in correlation analysis for the answers falling inside the cluster A (APPENDIX 1) that singing of the CI children was connected only to the amount of singing of the parents to the child during the last year before measurements (r_p = .757, p = .010), one year before that (r_p = .627, p = .004) and during the first year after implantation (r_p = .618, p = .005).

Summary of findings from Study II. The development of timbre and gap P3a and duration MMN and P3a differed between CI and NH groups. Overlap of early P3a with MMN diminished P3a for CI and NH groups for changes in timbre, and for the CI group also for changes in duration and gaps at T2. The early P3a of CI singers evidently affected comparisons of MMN between the CI and NH groups as well as between CI singers and CI non-singers. Importantly, the development of P3a was enhanced for CI singers over all change types, especially for changes in pitch (f_0) and timbre. These P3a responses were positively correlated with auditory working memory, consistent with P3a reflecting updating of auditory working memory, not only distraction. The only background

variable correlated with the singing of the CI children at home was singing of the parents to the child before the measurements, beginning from the first year after implantation.

4.3 The development of perception of word and sentence stress of CI children: The role of auditory cues, auditory working memory and supervised musical activities

The main aim of the Study III was to investigate how CI children develop in perception of word and sentence stress and whether this development improves with improving discrimination of acoustic cues, improving auditory working memory and more supervised musical activities outside of the home (in the CIm group). Additionally, we were interested especially in the development of auditory working memory of CI children.

Table 8. Results (unstandardized estimates) for LMM analyses for contributors to word and sentence stress perception.

	8a) Word stress				8b) Sentence stress			
	H I	H II	H III	Composite	H I	H II	H III	Composite
	B	B	B	B	B	B	B	B
Pitch (F0)	-8.96	-	-	8.06	-49.18***	-	-	-37.87**
Intensity	2.38***	-	-	-1.10˚	-1.57*	-	-	-1.69˚
Duration	.50	-	-	-7.03	-9.73	-	-	-14.32
Digit span	-	1.12***	-	.38˚	-	1.22*	-	-.17
PIQ	-	.26	-	.07	-	.98	-	.56
Group	-	-	-17.54***	-13.24***	-	-	-23.24**	-8.08

Group = CIm vs. CIn group. H = hypothesis. Composite = composite model.
- = the independent variable was not included in LMM. B shows the direction/ asterisks the strength of the correlation (˚p≤.1, *p≤.05, **p≤.01, ***p≤.001). All models: controlled for time, age and education of mother.

Higher levels of word stress perception were associated with lower thresholds for (better) discrimination of intensity (Table 8a, H I) while higher levels of sentence stress perception were associated with lower thresholds for discrimination of pitch (f_0) and intensity (Table 8b, H I). The correlations with discrimination of duration were not significant (Table 8). Word and sentence stress perception were unrelated to PIQ (Table 8a, 8b, H II). However, higher values of stress perception were associated with longer forward digit span and with more musical activity: the CIm group outperformed the CIn group (Table 8a, 8b, H III, Figure 8). The composite models including all of the hypothesized predictors showed that for word stress, the only significant, and hence the

strongest, factor was musical activity (Table 8a, Composite, Figure 8), and for sentence stress, the only significant factor was pitch (f_0) discrimination (Table 8b, Composite).
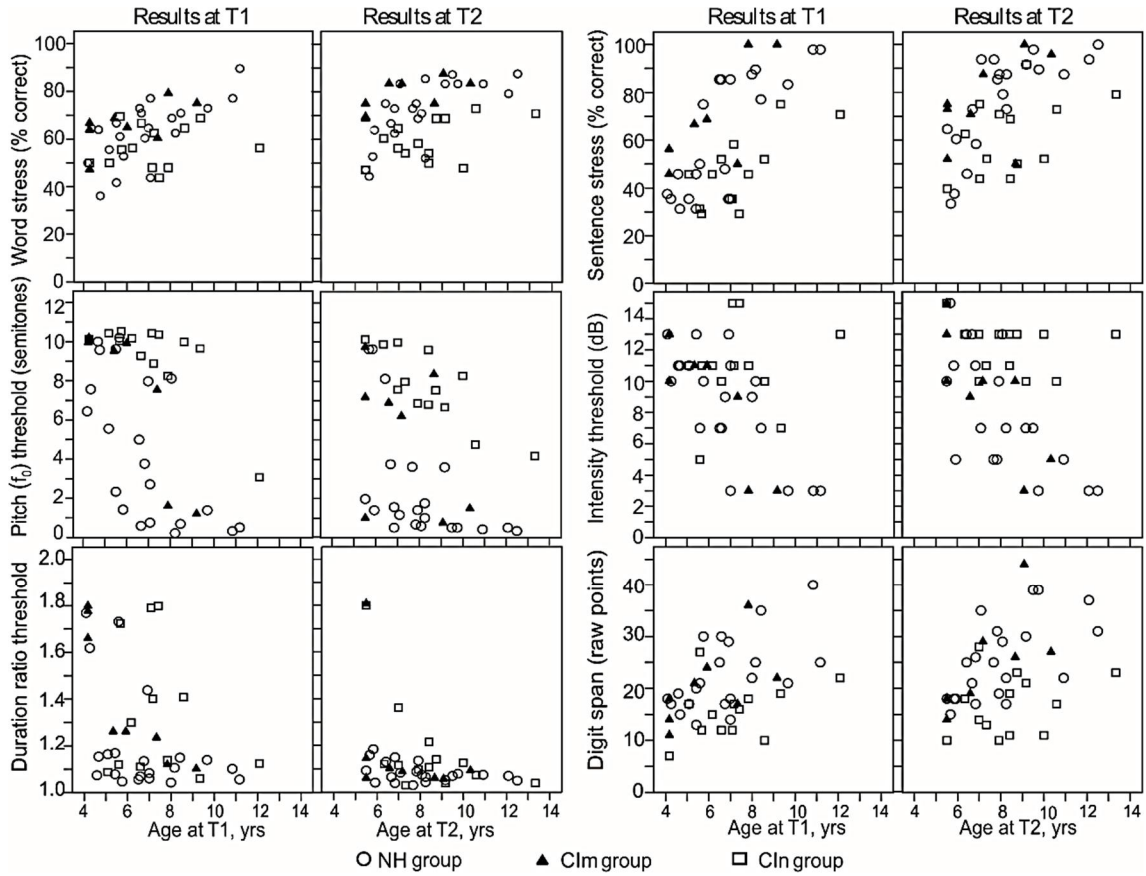
Figure 8. Comparisons of results for CI and NH children as a function of age and musical activity for CI children.

Table 9. Results (unstandardized estimates) for differences between CIm/CIn and NH group.

	Word stress		Sentence stress		Pitch (f_0)		Intensity[1]		Digit span	
	CIm/NH	CIn/NH	CIm/NH	CIn/NH	CIm/NH	CIn/NH	CIm/NH	CIn/NH	CIm/NH	CIn/NH
	B	B	B	B	B	B	B	B	B	B
Time	-9.44***	-6.35**	-12.49***	-15.04***	.16*	.16***	.38	.00	-3.28***	-2.27***
Age	4.03***	1.13***	8.27***	7.27***	-.08***	-.07***	-1.40***	.01**	2.43***	1.66**
Group	-7.85*	-11.36	-9.01	16.30**	-.13 °	-.43***	.22	5.39	-.09	8.23***
Age × group	ns	3.25*	ns	ns	ns	ns	ns	-1.27**	ns	ns

Time: reference is T2. Group: reference is the CIm or CIn group. [1] Thresholds: more negative value = better performance. ns = interaction was not significant. B shows the direction/asterisks the strength of the connection (°p≤.1, *p≤.05, **p≤.01, ***p≤.001). Education of mother was always controlled.

Next, we investigated how CI musical activity groups differed from each other and from the NH group in the development of perception of word and sentence stress. For the CIn vs. NH comparison, for word stress perception, there was a significant interaction of group with age (Table 9): the CIn group did not develop over age while the NH group

did. Surprisingly, the CIm group performed better than the NH group (Table 9, Figure 8). For sentence stress perception (Table 9), the CIm group performed as well as the NH group, while the CIn group performed more poorly than the NH group (Figure 8). The development with time and age was similar across groups (Table 9).

We also investigated whether CI musical activity groups differed from each other and from the NH group in the significant predictors of word and sentence stress. For intensity discrimination, both group comparisons revealed significant interactions of age and group (Tables 9 and 10). The CIn group did not develop over age while the CIm and NH groups did (Figure 8). For pitch (f_0) discrimination, the CIm group performed better than the CIn group (Table 10, Figure 8), and the CIm group did not differ from the NH group while the CIn group performed less well than the NH group (Table 9, Figure 8). For forward digit span, the CIm group outperformed the CIn group at T2, and only the CIm group developed between T1 and T2 (*time × music group*, Table 10). Moreover, the CIn group performed less well than the NH group while the CIm group performed similarly to the NH group (Table 9).

Table 10. Results (unstandardized estimates) for differences between CIm and CIn groups in the factors predicting prosodic perception.

	Pitch (f_0)[1]	Intensity[1]	Digit span
	B	B	B
Time	.17*	-.73	-4.00**
Age	-.08***	-1.62*	1.65 °
Music group[2]	.30***	-7.50 °	-8.98*
Time x music group	ns	ns	-3.00*
Age x music group	ns	1.70**	ns

[1]Thresholds: more negative value = better performance.
[2]Reference is the CIm children. ns = interaction was not significant. *B* shows the direction/asterisks the strength of the connection (°p≤.1, *p≤.05, **p≤.01, ***p≤.001). Education of mother was always controlled.

We also tested within the CI group whether musical activity contributed to discrimination after controlling for digit span. With this control in place, the correlation of musical activity with intensity discrimination was no longer significant ($B = 1.88$, $p = .142$), but the connection to pitch (f_0) discrimination remained strong ($B = .26$, $p = .002$), indicating that the relationship of musical activity to discrimination of pitch (f_0) was not purely due to variation in digit span.

Overview of interrelations. Because of the small sample size, LMM analysis cannot give an interpretable picture of the overall interrelations of these measures. For example, while both auditory discrimination and digit span are linked to prosodic perception, those two predictors might be highly intercorrelated. Therefore, partial correlation analyses were performed for CI children on average measures across the two measurement points. The first partial correlation analysis (age controlled) included only the hypothesized predictors of prosodic perception. As a result, digit span was connected to discrimination of intensity (r_p = -.717, p = .001) and of pitch (f_0) (r_p = -.630, p = .004). In addition, discrimination of pitch (f_0) and intensity were interconnected (r_p = .650, p = .003), as were discrimination of duration and intensity (r_p = .535, p = .018). Because digit span was found to be correlated with auditory discrimination, we examined links of auditory discrimination to word and sentence stress perception after partialling out digit span and age. This showed that the correlation of pitch (f_0) and intensity discrimination with sentence stress remained significant (pitch (f_0): r_p = -.680, p = .002; intensity: r_p = -.487, p = .040), as did the correlation of discrimination of intensity with word stress (r_p = -.559, p = .016).

Does singing by CI children at home play a role? In Study II it was found that the CI singers had enhanced P3a responses, and these responses were earlier with better digit span which in turn in Study III was connected to perception of prosodic stress. Therefore it was assumed that the perception of stress or auditory working memory would also be better with more singing by the CI child at home. To test this, we conducted additional analyses with similar procedures as for testing hypothesis III (for the LMM, see Table 8) and for testing the differences between CI musical activity groups in digit span (for the LMM, see Table 10) (see also section "Statistical analyses"). However, we added the CI singing group as an additional independent variable in the LMM. These analyses and results are not provided in the publications of the thesis.

For sentence stress, the CI singers performed better than the CI non-singers (B = -13.81, p = .038) and the main effect of musical activity group remained significant (B = -18.71, p = .011), implying that sentence stress perception was better with both more singing at home and more supervised musical activities. For word stress, the correlation with CI singing group was not significant and the correlation with musical activity group

remained strong (B = -16.41, p < .001). For digit span, the correlations with CI singing group was not significant. However, the interaction of time and musical activity group remained significant (B = 3.00, p =.048) while the main effect of musical activity group did not (B = -7.76, p = .099), implying that the singing of the CI children may have mediated the performance in digit span at T1, but not the development of digit span between measurements (for mediation, Baron & Kenny, 1986).

Summary of findings from Study III. The main result was that the CIm group performed at least equivalently to the NH group in stress and pitch (f_0) perception and in digit span, while the CIn group performed more poorly than both the NH group and the CIm group. Moreover, only the CIm group improved with age in word stress perception, intensity discrimination and improved over time in forward digit span. The higher values of word stress perception of the CI group were associated with longer forward digit span and better intensity discrimination: higher values of sentence stress perception were additionally associated with better pitch (f_0) discrimination. Further, more singing by the CI children was associated with improved sentence stress perception and might have mediated the improved performance in digit span at T1.

4.4 Connections of music perception to word stress and visuospatial perception for NH adults

The main aim of Study IV was to investigate whether music perception improved with improving word stress perception or with improving visuospatial perception for NH adults. We expected that especially the perception of musical rhythm would improve with improving word stress perception.

As a first step, the connections between the variables that could play a role in the connections of music perception to word stress or visuospatial perception were investigated (Table 11). Age was not linearly correlated with the music perception total score, but when the age groups were compared to each other using ANOVA, a significant difference was found (F = 6.21, p = .001). A post-hoc test (Tukey HSD) showed that the 40–49 age group had significantly higher music perception total scores than the 19–29 (p = .004) and 50–59 age groups (p = .002). The music perception total score was higher

with more musical education (Table 11) and better pitch discrimination thresholds (Table 11), the latter calculated as the size of the pitch change that the participant detected with 75% probability. Moreover, word stress perception was not correlated with pitch discrimination while it was positively correlated with music perception Total score (r = .34, p = .007), with Off-beat subtest score (r = .39, p = .002), and with forward digit span. It was not correlated with backward digit span (Table 11).

Table 11. Correlations between word stress, visuospatial and music perception and the variables possibly affecting the connections between these.

	Word stress	Visuospatial	Music perception (total)
Pitch perception: Change trials	.01	.05	.31*
No change trials	-.06	.07	-.15
All trials	-.06	.14	.09
Pitch discrimination threshold	-.13	-.03	-.32*
Auditory working memory	.26*	.10	.10
Digit span forward	.26*	.07	.07
Digit span backward	.13	.11	.06
Music education (years)	.12	.02	.32*
General education (years)	.18	-.11	.10

*p < .05, **p < .01.

Table 12. Results for four regression models.

Model	Total score				Scale subtest				Out-of-key subtest				Off-beat subtest			
	B	F	R^2	R^2 ch.	B	F	R^2	R^2 ch.	B	F	R^2	R^2 ch.	B	F	R^2	R^2 ch.
1		5.21**	.15	.15		3.35**	.10	.10		3.35*	.10	.10		1.67	.05	.05
Music ed.	.28*				.15				.31*				.14			
Age group	-.20				-.26*				-.04				-.15			
2		3.77**	.21	.06		1.78	.11	.00		2.77*	.17	.06		2.83*	.17	.11
Music ed.	.28*				.15				.32*				.12			
Age group	-.14				-.25°				-.01				-.04			
AWM	-.01				-.01				-.12				.13			
Pitch	-.25**				-.04				-.23				-.32*			
3		4.43**	.29	.08		2.05°	.16	.05		2.55*	.19	.02		3.33*	.23	.06
Music ed.	.28*				.15				.32*				.11			
Age group	-.08				-.20				-.03				.01			
AWM	-.02				.00				-.13				.12			
Pitch	-.26*				-.05				-.24°				-.33*			
VSP	.28*				.22°				.15				.26*			
4		5.27**	.37	.08		1.85	.17	.01		2.87*	.24	.05		4.34**	.33	.09
Music ed.	.28*				.15				.32*				.12			
Age group	-.09				-.20				-.00				.00			
AWM	-.10				-.03				-.20				.04			
Pitch	-.23*				-.03				-.21°				-.29*			
VSP	.27*				.21				.14				.25*			
Word stress	.30*				.12				.24°				.32**			

Musical ed. = music education (yes/no). Age group = under/over 50 years. AWM = auditory working memory (digit span forward + backward). Pitch = pitch discrimination threshold. VSP = visuospatial perception. R^2 ch. = R^2 change. °p≤.1, *p≤.05, **p<.01, ***p<.001.

Based on the findings above, step-wise regression analyses were performed to see how much the variables found to be related to word stress, musical or visuospatial perception could explain the variation of the music perception total score and subtests. Four different

predictor models were examined, as shown in Table 12 (models 1–4). For the total music perception score, the R^2 change showed that both visuospatial perception and word stress perception explained about 8% of the variance (Table 12, Figure 10). Musical education and pitch discrimination threshold were also significant predictors. Auditory working memory was not a significant predictor (Table 12).
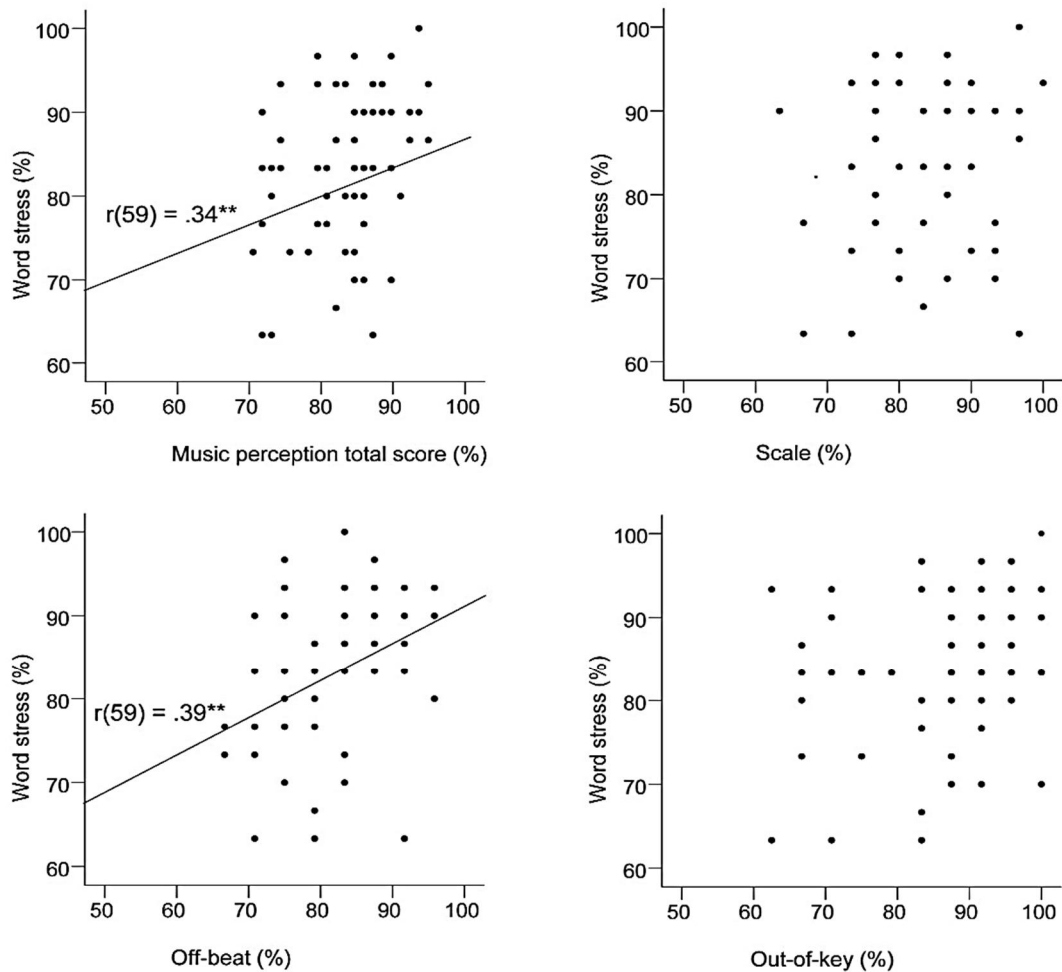
Figure 9. Scatterplots of word stress scores and music perception task scores.

For the Scale subtest, the only significant predictor in the first, and only, significant regression model was age group (Table 12). Visuospatial perception had only a marginally significant association with the Scale subtest with which it had been designed to be analogous. For the Out-of-key subtest, the final regression model was significant and explained 24 % of the variance, and the only significant predictor was music

education (Table 12). For the Off-beat subtest, the final model was significant and explained 33% of the variance, and the most significant predictor was word stress perception, which alone explained 9% of the variance (Table 12, Figure 9). Also pitch discrimination and visuospatial perception were significant predictors.

Summary of findings from Study IV. The main result from Study IV was that especially perception of musical rhythm (measured with the Off-beat subtest of the MBEA) was positively correlated with word stress perception. Also MBEA music perception total score and performance for the Out-of-scale subtest were better with more musical education. Moreover, pitch discrimination was connected to music perception but not to perception of word stress, which in turn was connected to forward digit span (auditory working memory), repeating the results from Study III for CI children. Visuospatial perception was a significant predictor of the MBEA total score and Off-beat subtest score. Visuospatial perception was not connected to the MBEA Scale subtest with which the task was analogous, implying that the association with music perception could be driven by some other variable, like attention.

5 Discussion

This thesis investigated the development of cortical processing of music and perception of prosodic stress of CI children. In addition, we studied the development of auditory working memory, auditory attention shift, and discrimination of acoustic cues for stress at the behavioural level and music at the neural level. Most importantly, we investigated the interplay between these and musical background (including singing at home and supervised musical activities outside of the home) for CI children. We also studied the connections of perception of music to word stress and visuospatial perception for NH adults.

More specifically, the auditory ERPs for piano tones and six acoustic change types, behavioural perception and forward digit span were measured twice (at T1 and T2) for 4–13-year old CI and NH children (Studies I-III). The CI children were divided into CI singers and CI non-singers based on the amount of singing of the CI children at home (Study II) and to musically active and non-active (CIm and CIn) groups based on the amount of supervised musical activities of the CI children outside of the home (Study III). In addition, music perception, word stress and visuospatial perception, pitch perception and digit span forward and backward were assessed once for 19–60-year old NH adults (Study IV).

The main findings were that for the CI children, the development of cortical processing of music, especially attention shift towards sound changes (P3a), was more advanced with more informal singing of the CI children at home (Study II), and the perception of prosodic stress was better for CI children with supervised musical activities outside of the home, the advantages of these musical activities extending to acoustic discrimination linked to prosodic perception as well as to auditory working memory (Study III). The results from NH adults (Study IV) resembled the findings on word stress in CI children (Study III). For both NH adults and CI children, perception of word stress was positively correlated with performance in the forward digit span task (auditory working memory) while the connection of word stress to pitch discrimination was not significant. Additionally, for NH adults, perception of musical rhythm improved with improving word stress and visuospatial perception (Study IV).

The implications of these results for CI children's perception of music and auditory attention are discussed further in section 5.1, for their prosodic perception and auditory working memory in section 5.2, and for their more general development in section 5.3.

5.1 The neural basis of music perception of CI children: The role of singing and attention

The CI children had well-formed ERP waveforms with P1, MMN and/or P3a resembling those recorded for the NH group (Study I), in line with the previous findings for CI children in non-musical contexts (Kileny et al., 1997; Ponton & Eggermont, 2001; Ponton et al., 2000, among others) and for musical context in adults and adolescents using CIs (Koelsch et al., 2004; Petersen et al., 2015; Sandmann et al., 2010; Timm et al., 2014). This implies that early-implanted children have neural abilities for discrimination of all measured change types, and that the neural networks for acoustic cues and MMN (see Introduction, sections 1.2 and 1.6) have developed rather well. Surprisingly, the MMN and/or P3a was clearly visible even for one semitone pitch (f_0) changes for the CI group. This may be explained by the rather low baseline frequency (295 Hz), allowing some CI children to follow the temporal cue for pitch (Green et al., 2002; Laneau & Wouters, 2004). Good pitch processing may be also related to the early age at implantation of the CI children, allowing their neural networks for pitch to develop well.

5.1.1 Differences between CI and NH groups

The P1 responses were smaller and earlier for the CI group than for the NH group (Study I), which may reflect impoverished processing specifically of natural piano tones for the following reasons. First, with simpler or familiar speech stimuli, P1 latencies and amplitudes typically develop to be similar for early implanted CI children and NH children (amplitudes: Jiwani et al., 2013; latencies: Sharma et al., 2002; among others). Second, because decreased response amplitudes reflect reduced synaptic density and efficiency (Picton & Taylor, 2007), the small P1 responses probably reflect poor neural representations of piano tone onsets. Third, stimulus differences affect PI amplitudes for CI listeners (Kelly et al., 2005). Moreover, the early P1 could be a consequence of

electrical stimulation per se, which in post-lingually deafened adults seems to reach auditory cortex faster than for acoustic stimuli in NH adults (Picton, 2010). Alternatively, it could be a consequence of plastic neural changes in the CI children. Congenital deafness can lead to hypersynchronization of peak latencies of local field potentials over distant cortical regions on the primary auditory cortex (Kral et al., 2009, their Fig. 9). The cortical networks for P1 could be affected by such a hypersynchronization, leading to early P1 responses.

The small and late timbre MMN and P3a (Studies I and II) and late pitch (f_0) P3a (Study II) for the CI group echo previous behavioural findings showing difficulties in discrimination of pitch and timbre by CI recipients (Limb & Roy, 2014; McDermott, 2004). For intensity increments, the ERP responses differed between CI and NH groups. We found a pattern of P3a followed by negative RON responses especially for large changes for the NH group (Studies I and II), indicating that these changes were clearly detectable for them. This pattern of P3a followed by negative RON responses was invisible for the CI group. This suggests that their unusual processing of intensity increments is probably a consequence of the activation of the automatic gain control of the speech processor above the 70 dB reference (Stöbich et al., 1999; see also Introduction, section 1.1), which made the present intensity increment changes difficult to detect for the CI group. Additionally, the activation of the gain control system might induce variation between subjects in the time-sensitive ERP responses, and cancel out the responses in the group level. However, the similar processing between CI and NH group for intensity decrements indicates that early-implanted children can follow intensity cues until the gain control system is activated, until the ceiling effect has been reached. In line with our results, Timm et al. (2014) found no differences between adult CI users and NH counterparts in intensity decrement MMN.

Another novel finding was the different development between CI and NH groups in MMN and P3a for gaps and changes in duration (Study II). Evidently, when the gap and duration changes elicited P3a only for the CI group, then their gap or duration MMN was smaller than for the NH group as a consequence of overlap of early P3a (shown with correlation analyses), especially for the CI singers. Conversely, the duration MMN of the NH group increased between measurements partially because of the lack of the overlap of P3a with MMN. Therefore, the CI vs. NH group comparisons of MMN were of little

value. Probably the changes in duration or gaps become less distracting for the NH children over time (see Wetzel et al., 2006, for the development of distraction over age in another context) while not for the CI children. This could be related to the reliance of CI users on sound envelopes, leading further to their reliance on surface rhythm (implemented in sound durations and gaps) in music perception (Gfeller & Lansing, 1991; Limb & Roy, 2014).

In summary, the results suggest that compared to NH children CI children have difficulties in processing of piano tone onsets and in neural discrimination of timbre and pitch, but not necessarily in discrimination of intensity decrements, gaps and changes in duration.

5.1.2 P3a without MMN: P3a reflects updating of auditory working memory?

Interestingly, as in previous studies (Horvath et al., 2008; Koistinen et al., 2012; Wetzel et al., 2006, among others), P3a without clear MMN was elicited, here especially for changes from piano to cembalo and violin for the NH group and for some of the CI group, for the CI singers (Study II). In the present thesis, the correlation analyses implied that the lack of MMN was a consequence of the partly overlapping MMN and P3a responses. The response to a change to cymbal differed from the responses to other timbre changes, eliciting MMN and P3a for all children. These results are consistent with the proposal that the attention shift can be a consequence of either a large physical difference or contextual novelty (Kushnerenko et al., 2013). Thus, for the CI singers and NH group, the change to cymbal might have been processed as a large physical change while the change to cembalo and violin was processed as a contextual difference, a change in musical instruments. The CI singers may have rather sophisticated neural networks for timbre, including the anterior temporal classification system (see Leaver & Rauschecker, 2010, see also Introduction, section 1.2), similarly to the NH children. Conversely, for the CI non-singers, the neural network for timbre might be less developed.

The connection of P3a latencies to digit span has implications for the interpretation of P3a for deviant tones in passive listening situations (see section 1.6). All ERP responses are sums of several components and each component can explain one part of the

manifestation of the response (Donchin & Coles, 1988, for a review). P300 responses to target sounds, consisting of P3a and P3b, are proposed to reflect not only discrimination, but also updating of working memory, i.e., a central executive component for monitoring and processing incoming information and then updating the items held in working memory by replacing information that is no longer relevant with new, more appropriate information (Donchin & Coles, 1988; Miyake et al., 2000). This assumption has been largely based on the finding that performance in digit span tasks is better with larger and earlier P300 responses (George & Coch, 2011; Polich et al., 1983). The present results suggest that the P3a responses to changes in musical sounds reflect updating of auditory working memory (see Barcelo et al., 2006), and with this, the functioning of the central executive component of auditory working memory (see section 1.4.1).

5.1.3 Advanced P3a responses with singing in the framework of discrimination, dynamic attending theory and neural networks for attention

The more advanced development of auditory attention shift (P3a) through all measured change types for CI singers (Study II) can be explained by better neural discrimination over all change types, which can also be related to better dynamic attending to the changes and better development of neural networks for attention as follows.

The better production of songs (evident in the production of rhythms and in the overall production of song elements in general, see section 3.1.1) for CI singers compared to CI non-singers at T2, and the earlier P3a with better auditory working memory strongly suggest that the P3a reflected better processing of acoustic changes by CI singers. Moreover, the CI non-singers had clearly visible MMN with degraded P3a which suggests that they did not link some of the changes, especially changes in pitch (f_0), to the behavioural level. Sometimes MMN can be recorded prior to behavioural discrimination ability becoming possible (for a review, Kujala et al., 2007). Also, the pitch (f_0) MMN became larger in CI non-singers without any evidence of increase in the pitch P3a, even though previous evidence shows that P3a increases with MMN (Draganova et al., 2009). The MMN without P3a for CI non-singers also shows similarities with NH subjects who suffer from tone-deafness, also called congenital amusia, who have near-to-normal

preattentive neural processing (MMN) of musical pitch incongruities even though they have highly limited behavioural accuracy in such a task (Peretz et al., 2009). The large MMN responses suggest that the neural networks for acoustic changes (see section 1.2) and for MMN (see section 1.6) have developed well in CI non-singers. At the behavioural level (Study III), pitch (f_0) discrimination was also enhanced for musically active CI children, and thus the present results are consistent with the proposal that multisensory musical training or singing is needed to enhance the discrimination of the acoustic cues for music, especially pitch, by CI children.

Singing by the child might be beneficial for auditory discrimination for several reasons. For example, because the ability of CI listeners to perceive pitch varies depending on the stimulus properties of the sound source (for example, Galvin et al., 2008), detecting predictable pitch (f_0) changes in the child's own familiar voice might be important, especially in the first years of hearing life. Further, the proprioceptive feedback from larynx in the context of predictable, well-learned children's songs might play a role in the perception of gross temporal changes and also in perception of pitch (f_0). Young children often produce a high pitch with high position and a low pitch with low position of the larynx (for a review, Trollinger, 2003), which might be easy to sense for deaf-born children, and even provide spatial cues for pitch. According to Welch (1985), reproducing an external song generates expectations of proprioceptive feedback which are then compared to the feedback received from the sensory receptors. This might be a reciprocal multisensory system which benefits perception of pitch by CI singers.

It is also possible that only CI singers interpreted the musical meter in the experiment, which can lead to dynamic variation in attention (Brochard et al., 2003; Potter et al., 2009). Meter is an aspect of relative rhythm, induced by accents (or beats) in the music, allowing the listener to synchronize to the rhythm of music (Brochard et al., 2003; Geiser et al., 2009; Hannon et al., 2004). Deafened, adult CI users have difficulties in deriving meter from piano music, perhaps due to the spectral and envelope properties of the piano tones (Phillips-Silver et al., 2015). In line with this, based on the present results on P1 responses (Study I), the CI children processed the piano tone onsets less accurately than the NH children. Because motor regions of the brain have been consistently shown to be involved in rhythm and meter perception (Chen et al., 2008; Overy & Turner, 2009), motoric training is essential for rhythm and meter perception (Cason et al., 2015; Phillips-

Silver et al., 2015), and since singing provides a rich content for motoric experiences of the regular meter in children's songs, singing may have advanced cortical networks for the perception of meter and may have led to better detection of regular meter. Dynamic attending theory (DAT: Jones, 1976; Jones & Boltz, 1989; Large & Jones, 1999) states that because of limited attentional resources, attention varies periodically according to internal dynamic oscillators. This determines the attending rhythm of an individual, and further, the times at which the prediction for and processing of external events are most effective. In line with DAT, it has been found that ERP responses in the P3a time range are more positive when the listeners hear the sound changes (deviants) in the on-beat position than in the off-beat position (Brochard et al., 2003; Potter et al., 2009). Moreover, Brochard and colleagues (2003) found that the ERP differences between the deviant in strong and weak accent (beat) positions arose earlier for subjects with musical training than in those without musical training. They interpreted this as indicating that musicians have stronger temporal expectancies, leading to the attention being deployed periodically more efficiently. This is in line with the consistently large and early P3a at T2 for the CI singers.

This is the first time that such a consistent difference for the development of P3a across all change types has been found in multi-feature paradigm studies. In the framework of DAT, it is possible that the attention system of early-implanted children relies on temporal regularities because of limited attentional resources. Dynamic temporal entrainment of attention in musical context could be beneficial for the cortical processing of acoustic changes, since attention reshapes receptive fields in the auditory cortex precisely and rapidly (Fritz et al., 2007). Further, the dynamic variation in attention induced by musical regularities could shape the attention networks.

The neural network for P3a is distributed across frontal, parietal and temporal cortical regions (Takahashi et al., 2013, among others), suggesting functional connectivity between them. In line with the connection found in this thesis between P3a and digit span, the neural networks for top-down and bottom-up auditory attention are highly overlapping (Alho et al., in press; Salmi et al., 2009), suggesting similar function for these networks. Congenital deafness can lead to deficiencies in neural networks for auditory attention (see introduction, chapter 1.3), and to degradation in white-matter volume in the auditory cortex and thus fewer afferent and efferent fibres (Emmorey et al., 2003).

Interestingly, it has also been found that people suffering from amusia have degraded connections between frontal and temporal regions in their right hemisphere (Loui et al., 2009). Thus, the lack of development of P3a responses for CI non-singers could be related to the consequences of early deafness for neural networks of auditory attention.

Conversely, musical activities, like singing, could cancel out these effects. NH musicians, especially singers, have enhanced white-matter (anatomical) connectivity between frontal and temporal cortical regions (Halwani et al., 2011), and singing-based aphasia therapy seems to lead to similar enhancement (Wan et al., 2014). Also faster plastic changes in auditory and frontal areas in 6 year old children participating in 15 months of musical training compared to other children have been found (Hyde et al., 2009). In line with this, musical activities at home, including singing, seem to enhance NH children's auditory attention functions, reflected in P3a responses for gap and duration changes (Putkinen et al., 2013). In conclusion, singing may well lead to enhancements in the neural networks for P3a, and this could lead to enhanced perception of music for CI singers. This conclusion is partially supported by the finding that the CI singers did not differ from the CI non-singers in factors related to hearing or CI devices, or other musical background than parental singing (see sections 3.1.1 and 4.2). Our results on the development of P3a responses between T1 and T2 indicate that singing can have effects on attention of CI children up until 13 years of age which was the age of our oldest participants at T2. This can be partially related to the late developmental trajectory of the prefrontal areas and neural circuits linked to them, essential for the neural networks for attention and working memory (Casey et al., 2000) while partially this may be a more general positive effect of singing on attentional capabilities, possibly observed also in NH children of the same age.

The more the parents had sung to CI children before measurements, the more the CI children sung by themselves, suggesting that parental singing encourages the CI children to sing. This indicates that parents should be encouraged to sing with their CI children starting right after implantation. The singing of the parents may play also a special role in the present results. It might be easy for the CI child to detect the acoustic changes, like changes in pitch and voice timbre, in the familiar voice of the parent, which could improve the perception of musical instrument pitch and timbre. Moreover, parental singing is known to arouse and regulate the attention of infants and young children (Rock et al.,

72

1999). This might be beneficial in the development of the neural networks for auditory attention.

5.1.4 Music perception and visuospatial perception are connected: Implications for CI children

Those NH adults who were better at visuospatial perception had better Total music perception and Off-beat subtest (measuring perception of rhythm) scores (Study IV). Importantly, because the expected association between the analogous test of music perception (the Scale subtest) and visuospatial perception was not significant, the link to music perception might be mediated not by pitch, but rather by perception of rhythm. The regular 500 ms pause between the two series and a regular 50 ms pause between two Gabors could be responsible for this finding.

As explained previously, DAT proposes that attention varies periodically, leading to enhanced performance in the task if it is performed at a moment when the attention is most effectively directed to that (Jones, 1976; Jones & Boltz, 1989; Large & Jones, 1999). In line with this, if the foreperiod (the time interval from the signal indicating that the task will soon appear to the beginning of the task) is predictable, the task performance becomes better or faster than for non-predictable foreperiods (Correa & Nobre, 2008). Intriguingly, the perception of auditory and visual rhythm seems to share similar neural bases. Escoffier and colleagues (2010) as well as Bolger and colleagues (2013) have shown that musical rhythm (meter) can affect the timing of the best performance in visual tasks. For example, if a visual task occurs on the on-beat position of the musical (rhythmic) sequence playing in the background, the performance of the subject in the visual task becomes faster (Escoffier et al., 2010). Moreover, after extensive short-term training, all rhythms, even those that are both trained and paced in visual modality, transform into auditory-motor representations and share similar neural networks (Karabanov et al., 2009). Therefore, those individuals who are more sensitive to rhythms in the auditory domain may also register better the time intervals between visual tasks. Thus the connection we found between visuospatial perception and performance at the Off-beat task could be related to variation in rhythm perception, leading to variation in detection of the regularity of the foreperiod and in attention towards the visual task.

Because the music perception Total score was improved also with improving visuospatial perception, it cannot be ruled out that visuospatial perception and music perception have a shared neural basis, and that visuospatial cues would enhance music perception, including pitch. In line with this possibility, the musical training procedure of Petersen and colleagues (2012) contained several exercises where the CI participants could benefit from visuospatial cues, and this kind of training led to enhancement of music perception in general. This suggests that in the rehabilitation of music perception for CI users visuospatial cues like movement in play songs would be beneficial.

5.2 Implications for stress perception and auditory working memory

5.2.1 The role of acoustic cues and auditory working memory in stress perception

For CI children, higher levels of word stress perception were associated with lower thresholds for (better) discrimination of intensity, and higher levels of sentence stress perception were associated with lower thresholds for pitch (f_0) and intensity discrimination, pitch (f_0) discrimination being the strongest contributor. The link of discrimination of pitch (f_0) to word stress perception was absent for NH adults. The links were not explained by variation in auditory working memory. The connections of discrimination to perception of sentence stress resemble findings for adults and children implanted later than those studied here (Meister et al., 2011; O'Halpin, 2010). The link of intensity discrimination to word stress perception was a novel finding, as was also the finding that for CI children perception of stress was positively correlated with performance in auditory working memory task, the link being supported by the similar connection to word stress for NH adults.

These results imply that in rehabilitation of stress perception, discrimination of pitch and intensity should be emphasized, and that CI devices should be developed towards better transmission of these acoustic cues. Evidently, auditory working memory plays an important role and should be controlled for when perception of stress is studied, and further, auditory working memory training should be addressed in the rehabilitation of

stress perception. Based on the present results, musical activities might enhance all abovementioned aspects.

5.2.2 The role of musical activities in stress, pitch and intensity perception and auditory working memory

Intriguingly, those CI children who had participated in supervised musical activities outside of the home (CIm children) performed at least equivalently to the NH group for stress perception, discrimination of acoustic cues for stress, and forward digit span, while other CI children (CIn children) performed consistently more poorly than both the NH children and the CIm children (Study III). It seems that musical activities before our first measurements were important for these skills, giving evidence on the positive role of the early onset of musical training. In addition, only CIm children developed from T1 to T2 for auditory working memory, implying that musical training at later ages (up until to 13 years of age) is important for the development of this cognitive skill of CI children. Only CIm children also developed with age for intensity discrimination, even though the latter result did not remain significant when auditory working memory was controlled for. Evidently, the better development of intensity discrimination in CIm children was connected to their better development of auditory working memory.

The superior perception of word and sentence stress and better development of discrimination of intensity for CI children attending supervised musical activities compared to other CI children were novel findings. However, longitudinal experimental studies of NH children show positive impact of musical training to skills closely associated to stress perception (perception of emotional prosody, Thompson et al., 2004; verbal memory, Ho et al., 2003; Roden et al., 2012). The advanced pitch (f_0) discrimination of musically active CI children is in line with the findings from NH listeners in Study IV and from previous studies, showing advanced pitch (f_0) perception for musically trained individuals (adults: Deguchi et al., 2012; Micheyl et al., 2006; Schön et al., 2004; Tervaniemi et al., 2005; children, Magne et al., 2006; Parbery-Clark et al., 2009). Importantly, the longitudinal study of Moreno and colleagues (2009) shows that musical training improves the perception of pitch in speech by NH children, and musical training also seems to improve pitch (f_0) perception for CI children (Chen et al., 2012) and adults using CIs (Petersen et al., 2012). Moreover, the CIm and CIn children did not

differ in factors related to CIs or thresholds for hearing, and maternal education and age were controlled for in statistical analyses. Beneficial effects of musical training have been expected in the CI population (Shahin, 2011). It is even possible that, due to the poorer baseline in auditory skills, the outcomes of musical training could be stronger for CI children than for NH children. Indeed, we found striking differences between CIm and CIn children. Therefore, the present findings support the interpretation that musical activities, including singing, enhance the perception of word and sentence stress as well as pitch and intensity perception.

The better perceptual skills found for CIm children might be partially related to the slower tempo and the predictable pitch and intensity changes in music and songs compared to speech (Patel, 2014), which both might be beneficial for CI children. The advanced pitch (f_0) perception of musically active CI children is also in line with the suggestion that musical training enhances the processing of rapid spectrotemporal changes (Tallal & Gaab, 2006). Notably, the low baseline frequency in the stimulus for pitch discrimination thresholds may have allowed musically trained CI children to follow the temporal cue for pitch (Green et al., 2002; Laneau & Wouters, 2004), or to follow a combination of temporal and place cues (Goldstein, 1973; Moore, 2003a, 2014).

As discussed earlier, the good pitch perception could also be partially related to the integration of proprioceptive cues with auditory cues for pitch (f_0) in those CI children who sing. This may be related to the present findings, because the emphasis in the supervised activities was on singing. However, the participation in supervised musical activities outside of the home may have additional benefits for pitch (f_0) perception of CI children. It is known that deafness since birth has effects on the development of peripheral visual neural system, leading to better attention towards and better perception of motion in the periphery of the visual field (Hauthau et al., 2013; Neville & Lawson, 1987). In the supervised musical activities, which were group activities (musical play schools, Lindfors Foundation speech-music groups), the CIm children had an opportunity to see the movements of others, and in these activities, the pitch movements were often visualized with hand cues or toys. Moreover, the CIm children were exposed to musical instrument playing by others and by themselves in their supervised musical activities and some of them at home (see section 3.1.1). Thus, they could see how the pitch (f_0) was produced with the keyboards or other instruments. Early-implanted children may be good

multisensory integrators (Schorr et al., 2005). They may be able to integrate the visuospatial cues provided by group musical activities with their auditory pitch (f_0) perception, and even with proprioceptive sensations related to pitch (f_0).

Hearing pitch (f_0) changes from several musical instruments and voices of varying timbre in the music groups may also have led to better perception of pitch independent of timbre by the CIm group than by the CIn group. In line with this proposal, Galvin et al. (2008) found that the CI adults who had participated in musical training were largely unaffected by instrument timbre in the perception of melodic contour, while the performance of the other CI listeners varied across instruments. Timbre-independent perception of pitch may have played a role especially in the present pitch discrimination task, where the stimulus was synthesized and unfamiliar to the children.

Auditory working memory. The striking findings on similar auditory working memory for CIm and NH children and better auditory working memory for CIm children than for CIn children are in line with the superior auditory working memory for musically trained adults and children (George & Coch, 2011; Lee et al., 2007; Strait et al., 2012; for a review, Besson et al., 2011). Moreover, the development over time (between our measurements) only for the CIm group is consistent with findings from longitudinal and intervention studies showing that music training enhances forward digit span of 4–6 years of old NH children and verbal memory of school-aged NH children (Fujioka et al., 2006; Roden et al., 2012). It has been found previously that CI children do not reach NH children's auditory working memory capacity (Pisoni et al., 2011), echoing the present finding for the CIn children. Therefore, the present results suggest that supervised musical activities with others can have a crucial impact on the auditory working memory development of CI children.

The statistical results suggested that within the CI group, singing at home and musical activities outside of the home had some influence on performance on the digit span task at T1 and in the development of auditory working memory before T1. However, singing at home was not connected to the development of digit span between T1 and T2. Therefore, singing by the child at home and participation in supervised musical activities outside of the home may tap different subcomponents of auditory working memory. The interlink between the performance on the digit span task, P3a responses and singing of

the CI children may indicate that singing is connected to the updating of auditory working memory. Singing is probably related to the central executive working memory subcomponent (Miyake et al., 2000). Musical activities outside of the home in turn might tap the short-term memory component of auditory working memory i, e, the capacity to hold items temporarily in memory (Baddeley, 2003).

Sentence stress perception was also better with more singing by the CI children themselves, which in turn was connected to parental singing. It has been found that infant-directed speech, where the parents naturally use exaggerated pitch contours and emphasize the important words with sentence stress, directs the infant's attention to the speech (Thiessen et al., 2005). In the present study, singing was also related to earlier attention shift towards sound changes, and these responses seem to reflect updating of auditory memory. It is well possible that both sentence stress perception and singing are related to updating of auditory working memory, which underlies especially the connection of perception of sentence stress to singing. Taken together, the results from this thesis thus suggest that both singing at home and musical activities outside of the home with others are needed to get the best benefits from musical activities for CI children.

Why would singing and supervised musical activities with others have different roles, singing advancing neural attention functions and updating of auditory working memory and supervised musical activities affecting processing in short-term memory? Singing at home by the child is an activity where the child usually sings alone, without competing sounds. The child repeats the same songs many times and singing is done by their own free choice, without feedback from others. Learning as a consequence of singing by oneself is probably largely based on iteration and trial and error, and the motivation to sing is based on the rewarding effects of singing. In contrast, in musical activities outside of the home, the auditory environment and singing tasks are more demanding. There are competing sounds, the child has to adapt his/her singing to the singing of others, and the child is expected to learn new songs, not only to repeat the already learnt ones. Moreover, there are lots of visual cues provided by others, which may lead to better learning related to short-term memory and to better behavioural performance than singing by oneself. Perhaps the visuospatial cues provided by others, the tasks becoming more demanding over time and feedback (Klingberg et al., 2005; in CI children, Kronenberger et al., 2011)

are all needed to enhance short-term memory, while singing without external guidance improves the updating of the auditory working memory component related to digit span.

5.2.3 Music perception and word stress perception are connected via rhythm: Implications for CI children

Better word stress perception was connected to better music perception for NH adults, especially in the Off-beat subtest of the MBEA (Study IV). The link between duration discrimination and word stress perception was absent for CI children implying that discrimination of simple tone duration would not drive the link with stress perception. However, the link could be driven by the perception of how the tones unfold over time in music and speech.

In the present word-stress task, there were two strong accents, or one strong accent and another weaker accent, in otherwise similar target words. The unfolding of these accent patterns over time indicated the auditory targets. It is possible that some listeners perceived the changes in accent patterns similarly to changes in musical meter, which is also implemented in the accentual patterns (beat) unfolding over time (Geiser et al., 2009; Hannon et al., 2004). So the link between word stress and performance of the Off-beat task for NH adults might reflect the fact that those who are better at perceiving musical meter are better at detecting word stress patterns.

Interestingly, the link of musical activities to word stress perception in CI children was extremely strong in the composite model (Study III). The ability to discriminate changes in intensity, which enables the perception of beat and meter, contributed to perception of word stress by CI children, and discrimination of intensity improved for CI children with more participation in supervised musical activities. The overall pattern of results leads to the question of whether supervised musical activities of CI children also led to enhancement of meter perception through enhanced intensity perception. This could contribute to the strong connection of supervised musical activities to word stress perception by CI children.

From the perspective of musical neuroscience, the present results suggest that music and speech share similar neural resources in the domain of rhythm. There are several other studies implying such a connection. For instance, musicians have been found to perceive

the metric structure of words more precisely than non-musicians (Marie et al., 2011). Further, the results of Cason and Schön (2012), Bolger and colleagues (2013) and Cason and colleagues (2015) show that musical rhythm and especially meter drives enhanced perception of speech. Most importantly, the link of rhythm perception to word stress perception suggests that, in the rehabilitation of CI children, improving rhythm perception in music might be a way to improve word stress perception. Evidently, rhythmic exercises should be not omitted from the rehabilitation of word stress perception by CI children.

5.3 Implications for speech, language and other development of CI children

The present findings on connections of musical activities and singing to auditory attention, to perception of word and sentence stress, to pitch perception and to auditory working memory as well as the connection found between rhythm perception and word stress perception, all have wider importance than discussed above.

The earlier and increased P3a responses for CI singers are highly important, suggesting that CI singers have better auditory attention functions in general. Attention towards sounds can enhance the representations of sounds, including speech, in the auditory cortex and brainstem (Fritz et al., 2007; Strait et al., 2014; Woods & Alain, 2009; Woods et al., 2009). The efficient functioning of auditory attention is also important for perception and learning of degraded auditory stimuli, including speech with CIs (Beer et al., 2011; Houston et al., 2014; Wild et al., 2012), and therefore also for language acquisition with CIs. Good attention functions are also necessary for any kind of learning and though this, for good academic success (Kronenberger et al., 2013, among others). Those CI children who sing regularly may thus benefit from their better attention functions for learning in general, from music perception to speech perception and language skills and beyond these. Even if the enhanced and early P3a responses were not related to general enhancement of auditory attention, they would nevertheless reflect good neural discrimination and efficient attention shift towards auditory changes. This is necessary in order to process rapidly changing auditory scenes like in traffic, or in schools, daycare centres and other places where attention should be directed quickly towards important sounds. The present results suggest a better quality of life for CI singers.

Improved perception of word and sentence stress can lead to better segmentation of words from continuous speech, and through this, to better language skills (Friedrich et al., 2009; Houston et al., 2004; Jusczyk et al., 1999; Mattys et al., 1999, 2005; Vroomen et al., 1998), especially for CI children (see section 1.4). Similarly, good perception of sentence stress, expressed mainly as changes in pitch, can enhance the language development of young children (Fernald & Mazzie, 1991; Thiessen et al., 2005). Detecting pitch variations in general may be important. Newborns can detect pitch (f_0) variations in speech and may begin to use these to aid language acquisition (Sambeth et al., 2008). Variation in pitch in infant-directed speech aids development of vowel categories (Trainor & Desjardins, 2002), pitch variations in songs improve infant's perception of the phonetic content of speech (Lebedeva & Kuhl, 2010), and even adults benefit from sentence stress, produced only by pitch variation, in learning of new words (Filippi et al., 2014). Because detailed phonetic cues are not available to CI children, it can be assumed that any enhancement of access to these prosodic cues with musical activities would have a strong impact on overall speech and language development.

Children with CIs typically show poor auditory working memory (Harris et al., 2013; Kronenberger et al., 2011; Kronenberger et al., 2014; Pisoni & Cleary, 2003, Pisoni et al., 2011). Deficits in working memory may also become a problem when the task carries a high cognitive load, like in hearing in background noise, in perception of spoken sentences, or in formulating sentences based on a picture (Beer et al., 2011). Auditory working memory for CI children is also strongly connected to their language learning and reading skills (Kronenberger et al., 2011; Ingvalson et al., 2014; Pisoni & Cleary, 2003; Pisoni et al., 2011). For NH children, auditory working memory plays a crucial role in language learning (Baddeley, 2003; Baddeley et al., 1998). Therefore, the similar digit span for CIm children and for NH children, and development over time only by the musically active children, are utmost important findings. The present results on enhancement of auditory working memory functions bode well for the language development and academic success of musically active children.

Last but not least, superior music perception with singing or other musical activities may enhance their quality of life through the entire life span. Music is highly attractive for young children, and it also attracts young CI children (Trehub et al., 2009). Even at later ages, it induces emotions and is a way to express them (Reybrouck & Brattico,

2015), it helps in regulation of emotions (Saarikallio, 2010), it gives us pleasure and rewards us (Zatorre & Salimpoor, 2013) and it aids in maintaining the healthy functioning of memory and other cognitive functions in old age (Särkämö et al., 2014). Importantly, good perception of music, including perception of rhythm and meter, as indicated by the results of this thesis, can also have positive effects on word stress and speech perception and language learning. Even though CI children do not achieve as good perception of music as NH children, this does not prevent them from enjoying music or singing (Trehub et al., 2009). There seems to be a reciprocal relationship between skills and interest and motivation, beginning in the preschool period (Aunola et al., 2006; Fisher et al., 2012), i. e., interest and motivation towards learning a particular skill leads to better learning and performance. Therefore, it is important for the development of CI children to give parents and professionals the message that supporting the music enjoyment of CI children might be beneficial for their music perception and, with this, for their quality of life.

5.4 Limitations of the study

The results of the present thesis show consistent advantages for those CI children who sing at home or take part musical activities outside of the home with emphasis on singing. The musical instrument playing of CI children in general was not regular. Only few of them had access to musical instruments at home, and so it was impossible to study specifically the advantages of instrument playing. Therefore, the present thesis cannot give interpretable results on whether musical instrument playing is beneficial for CI children.

Due to the young age of the participants, we could not have a good control over the focus of selective auditory attention. That is, the participants could not do another challenging task when they heard the to-be-ignored sound sequence (see for example, Alho et al., 1997; Alho et al., in press). Further studies should assess the attention functions of older CI children with more challenging experimental paradigms.

It is important to note that the study design cannot define the causality, and the differences found here could be a consequence of some predispositions which we could not find. To confirm causality, the CI children should have been randomly assigned to musical activity groups, like those attending musical activities outside of the home and

those who do not, or to those who sing a lot alone at home and those who do not. Unfortunately, this was not possible due to the small number of early-implanted children in Finland (less than 300, CI children living in areas distant from each other). Further, the rather small number of participants may restrict the generalization of the results. The small number of each type of CI device and processing strategy is also a weakness, and very little can be said about the role of these aspects in the results.

It cannot be completely ruled out that since no loudness-balancing between the standard and the deviants in pitch was done, due to the young age of the participants, the changes in pitch may have caused changes in loudness due to the functioning of the CI (see Introduction, section 1.1), partially leading to significant responses even for the smallest, one semitone changes. Moreover, we conducted many statistical analyses, but we corrected for multiple testing only for the post-hoc tests (Studies I, II and III). This might have sometimes led to type 1 errors, i.e., some connections could be significant by chance. As this was the first study of most of the aspects under investigation, we preferred to avoid type 2 errors. Therefore, we feel that the best solution was to use relatively liberal correction procedures.

6 Conclusions

This thesis investigated speech- and music-related brain processes and task performance for CI children and for NH children. With regard to the development of music-related brain processes, we found well-formed ERP waveforms for CI children, resembling those for the NH group. However, many times the ERP responses implied impoverished processing for the CI children, especially in the case of timbre and pitch. We also found different development of ERP responses between CI and NH groups. However, this was sometimes caused by the different development of these responses between CI singers and CI non-singers. With regard to the perception of word and sentence stress and related auditory cues as well as to development of auditory working memory, the CI children participating in supervised musical activities performed and developed similarly to the NH children while the other CI children performed or developed less well than NH children.

With regard to the quality of musical activities, we found that more singing of the CI children is related to clear advantages in the development of P3a, i.e., auditory attention shift towards sound changes, especially in pitch and timbre, and to perception of sentence stress. More supervised musical activities outside of the home were found to be related to advantages in the development of perception of word and sentence stress and related auditory cues (including pitch) and in auditory working memory. Therefore, both types of musical activities may have their own specific role in shaping the development of pitch-related auditory skills important for language development and quality of life of CI children. Advantages with musical activities were found already at T1 (especially for perception of pitch and prosody), but also between TI and T2 (for auditory attention shift and auditory working memory). This suggests that musical activities might have effects not only at an early age, but also later, up until age of 13 years.

The results of this thesis hopefully will help professionals to build up the rehabilitation of music and speech perception more efficiently, even if it is impossible to give every CI child an opportunity to take part in musical activities. In improving perception of stress it seems to be worth especially addressing perception of pitch (f_0), intensity and rhythm, as well as auditory working memory. Moreover, in improving perception of music, visuospatial cues seem to be beneficial. The results have implications for theories on the

connections between music and speech. They also give more evidence suggesting that speech and music processing are connected not only via pitch and timbre, but also via rhythm. For the ERP research field, the present results give new evidence indicating that P3a responses reflect updating of auditory working memory. Further, they imply that early P3a can affect MMN.

The novel findings here should be followed up, and hopefully, this thesis gives some guidelines as to how to do it. Furthermore, experimental studies are needed to confirm that musical activities enhance the skills under investigation in this study, and also speech, language and performance in everyday life. However, there is a high risk that while waiting these results, many CI children will miss an opportunity to take part in music. Therefore, meanwhile, parents should be encouraged to find ways to make CI children - as well as themselves - enjoy singing, because this can have no foreseeable negative effects. Professionals should search for ways to enable CI children to attend supervised musical activities outside of the home, independently of the parents' socioeconomic status, and spread the message that despite the difficulties of CI users in perceiving pitch, CI children can take part in and benefit from musical activities at home, school and daycare centres. The combination of singing at home and taking part in supervised musical activities outside of the home might be the best way to optimize the quality of life of early-implanted children.

7 References

Abbas, P. J., Hughes, M. L., Brown, C. J., Miller, C. A., & South, H. (2004). Channel interaction in cochlear implant users evaluated using the electrically evoked compound action potential. *Audiology and Neuro-Otology, 9,* 203–213.

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex, 21,* 1507–1518.

Alho, K., Escera, C. Díaz, R., Yago, E., & Serra, J. M. (1997). Effects of involuntary auditory attention on visual task performance and brain activity. *NeuroReport, 8,* 3233–3237.

Alho, K., Salmi, J., Koistinen, S., Salonen, O., & Rinne, T. (in press). Top-down controlled and bottom-up triggered orienting of auditory attention to pitch activate overlapping brain networks. *Brain Research.*

Alho, K., Tervaniemi, M., Huotilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R. J., et al. (1996). Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology, 33,* 369–375.

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

Alho, K., Woods, D. L., Algazi, A., Knight, R. T., & Näätänen, R. (1994). Lesions of frontal cortex diminish the auditory mismatch negativity. *Electroencephalography and Clinical Neurophysiology, 91,* 353–362.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology, 87,* 85–106.

Alvarenga, K. D. F., Vicente, L. C., Lopes, R. C. F., Ventura, L. M. P., Bevilacqua, M. C., & Moret, A. L. M. (2013). Development of P1 cortical auditory evoked potential in children presented with sensorineural hearing loss following cochlear implantation: a longitudinal study. *CoDAS, 25,* 521–526.

Arnoldner, C., Kaider, A., & Hamzavi, J. (2006). The role of intensity upon pitch perception in cochlear implant recipients. *Laryngoscope, 116,* 1760–1765.

Aunola, K., Leskinen, E., & Nurmi, J.-E. (2006). Developmental dynamics between mathematical performance, task motivation, and teachers' goals during the transition to primary school. *British Journal of Educational Psychology, 76,* 21–40.

Baddeley, A. (1992). Working memory. *Science, 255,* 556-559.

Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 49,* 5–28.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36,* 189–208.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105,* 158–173.

Barcelo, F., Escera, C., Corral, M. J., & Perianez, J. A. (2006). Task switching and novelty processing activate a common neural network for cognitive control. *Journal of Cognitive Neuroscience, 18,* 1734–1748.

Barcelo, F., Perianez, J. A., & Knight, R. T. (2002). Think differently: A brain orienting response to task novelty. *Neuroreport, 13,* 1887–1892.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Barres, B. A., & Raff, M. C. (1993). Proliferation of oligodendrocyte precursor cells depends on electrical activity in axons. *Nature, 361,* 258–260.

Baskent, D., & Shannon, R. V. (2003). Speech recognition under conditions of frequency-place compression and expansion. *Journal of the Acoustical Society of America, 113,* 2064–2076.

Beer, J., Kronenberger, W. G., & Pisoni, D. B. (2011). Executive function in everyday life: implications for young cochlear implant users. *Cochlear Implants International,* 12 Suppl 1, S89–91.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullen, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience, 8,*

1148–1150.

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

Bolger, D., Trost, W., & Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta Psychologica, 142,* 238–244.

Boons, T., De Raeve, L., Langereis, M., Peeraer, L., Wouters, J., & van Wieringen, A. (2013a). Expressive vocabulary, morphology, syntax and narrative skills in profoundly deaf children after early cochlear implantation. *Research in Developmental Disabilities, 34,* 2008–2022.

Boons, T., De Raeve, L., Langereis, M., Peeraer, L., Wouters, J., & van Wieringen, A. (2013b). Narrative spoken language skills in severely hearing impaired school-aged children with cochlear implants. *Research in Developmental Disabilities, 34,* 3833–3846.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433-436.

Brattico, E., Tupala, T., Glerean, E., & Tervaniemi, M. (2013). Modulated neural processing of Western harmony in folk musicians. *Psychophysiology, 50,* 653–663.

Brechmann, A., Baumgart, F., & Scheich, H. (2002). Sound-level-dependent representation of frequency modulations in human auditory cortex: A low-noise fMRI study. *Journal of Neurophysiology, 87,* 423–433.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science, 14,* 362–366.

Brochard, R., Dufour, A., & Despres, O. (2004). Effect of musical expertise on visuospatial abilities: Evidence from reaction times and mental imagery. *Brain and Cognition, 54,* 103–109.

Busby, P. A., & Clark, G. M. (1999). Gap detection by early-deafened cochlear-implant subjects. *Journal of the Acoustical Society of America, 105,* 1841–1852.

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America, 118,* 471–482.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology, 54,* 241–257.

Cason, N., Astesano, C., & Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio-motor training affect speech perception. *Acta Psychologica, 155,* 43–50.

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

Chartrand, J.-P., & Belin, P. (2006). Superior voice timbre processing in musicians. *Neuroscience Letters, 405,* 164–167.

Chatterjee, M., & Oberzut, C. (2011). Detection and rate discrimination of amplitude modulation in electrical hearing. *Journal of the Acoustical Society of America, 130,* 1567–1580.

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

Chatterjee, M., & Shannon, R. V. (1998). Forward masked excitation patterns in multielectrode electrical stimulation. *Journal of the Acoustical Society of America, 103,* 2565–2572.

Chen, J. K. C., Chuang, A. Y. C., McMahon, C., Hsieh, J. C., Tung, T. H., & Li, L. P. H. (2010). Music training improves pitch perception in prelingually deafened children with cochlear implants. *Pediatrics, 125,* E793–E800.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18,* 2844–2854.

Chobert, J., Francois, C., Velay, J.-L., & Besson, M. (2014). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex, 24,* 956–967.

Chobert, J., Marie, C., Francois, C., Schön, D., & Besson, M. (2011). Enhanced passive and active processing of syllables in musician children. *Journal of Cognitive Neuroscience, 23,* 3874–3887.

Ciocca, V., Francis, A. L., Aisha, R., & Wong, L. (2002). The perception of Cantonese lexical tones by early-deafened cochlear implantees. *Journal of the Acoustical Society of America, 111,* 2250–2256.

Cole, E. B., & Flexer, C. (2011). *Children with hearing loss: Developing listening and talking.* San Diego, Oxford, Brisbane: Plural Publishing.

Correa, A., & Nobre, A. C. (2008). Neural modulation by regularity and passage of time. *Journal of Neurophysiology, 100,* 1649–1655.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology,* 115, 732–744.

Deguchi, C., Boureux, M., Sarlo, M., Besson, M., Grassi, M., Schön, D., et al. (2012). Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: Behavioral, electrophysiological and psychoacoustic study. *Brain Research,* 1455, 75–89.

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.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *The Behavioral and Brain Sciences,* 11, 355–425.

Donaldson, G. S., & Kreft, H. A. (2006). Effects of vowel context on the recognition of initial and medial consonants by cochlear implant users. *Ear and Hearing,* 27, 658–677.

Draganova, R., Wollbrink, A., Schulz, M., Okamoto, H., & Pantev, C. (2009). Modulation of auditory evoked responses to spectral and temporal changes by behavioral discrimination training. *Bmc Neuroscience,* 10, 143.

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.

Emmorey, K., Allen, J. S., Bruss, J., Schenker, N., & Damasio, H. (2003). A morphometric analysis of auditory brain regions in congenitally deaf adults. *PNAS,* 100, 10049–10054.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology-General,* 128, 309–331.

Escera, C., Alho, K., Winkler, I., & Näätänen, R. N. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience,* 10, 590–604.

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

Escoffier, N., Sheng, D. Y. J., & Schirmer, A. (2010). Unattended musical beats enhance visual processing. *Acta Psychologica,* 135, 12–16.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology,* 27, 209–221.

Filippi, P., Gingras, B., & Fitch, W. T. (2014). Pitch enhancement faciliates word learning across visual contexts. *Frontiers in Psychology,* 5.

Formisano, E., Kim, D. S., Di Salle, F., van de Moortele, P. F., Ugurbil, K., & Goebel, R. (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron,* 40, 859–869.

Fox, L. G. & Dalebout, S. D. (2002). Use of the median method to enhance detection of the mismatch negativity in the responses of individual listeners. *Journal of the American Academy of Audiology,* 13, 83–92.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews,* 25, 355–373.

Friedrich, M., Herold, B., & Friederici, A. D. (2009). ERP correlates of processing native and non-native language word stress in infants with different language outcomes. *Cortex,* 45, 662–676.

Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience,* 6, 1216–1223.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in AI? *Hearing Research,* 229, 186–203.

Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain,* 129, 2593–2608.

Galvin, J. J., III, Fu, Q.-J., & Oba, S. (2008). Effect of instrument timbre on melodic contour identification by cochlear implant users. *Journal of the Acoustical Society of America,* 124, EL189–EL195.

Galvin, J. J., III, Fu, Q.-J., & Shannon, R. V. (2009). Melodic contour identification and music perception by cochlear implant users. *Neurosciences and Music III: Disorders and Plasticity,* 1169, 518–533.

Garcia, D., Hall, D. A., & Plack, C. J. (2010). The effect of stimulus context on pitch representations in the human auditory cortex. *Neuroimage,* 51, 808–816.

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology,* 120, 453–463.

Geers, A., Brenner, C., & Davidson, L. (2003). Factors associated with development of speech perception skills in children implanted by age five. *Ear and Hearing,* 24, 24S–35S.

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex,* 45, 93–102.

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia,* 49, 1083–1094.

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

Gfeller, K., & Lansing, C. R. (1991). Melodic, rhythmic, and timbral perception of adult cochlear implant users. *Journal of Speech and Hearing Research,* 34, 916–920.

Gfeller, K., Witt, S., Woodworth, G., Mehr, M. A., & Knutson, J. (2002). Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *The Annals of Otology, Rhinology, and Laryngology,* 111, 349–356.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance – a topographic event-related potential study. *Psychophysiology,* 27, 627–640.

Goldstein, J. L. (1973). An optimum processor theory for the central formation of the pitch of complex tones, *Journal of the Acoustical Society of America*, 54, 1496–1516.

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.

Green, T., Faulkner, A., & Rosen, S. (2002). Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants. *Journal of the Acoustical Society of America,* 112, 2155–2164.

Green, T., Faulkner, A., & Rosen, S. (2004). Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. *Journal of the Acoustical Society of America,* 116, 2298–2310.

Griffiths, T. D., & Hall, D. A. (2012). Mapping pitch representation in neural ensembles with fMRI. *Journal of Neuroscience,* 32, 13343–13347.

Grube, M., Cooper, F. E., Chinnery, P. F., & Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proceedings of the National Academy of Sciences of the United States of America,* 107, 11597–11601.

Halwani, G. F., Loui, P., Rueber, T., & Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. *Frontiers in Psychology,* 2.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology,* 50, 354–377.

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology-Human Perception and Performance,* 30, 956–974.

Harris, M. S., Kronenberger, W. G., Gao, S., Hoen, H. M., Miyamoto, R. T., & Pisoni, D. B. (2013). Verbal short-term memory development and spoken language outcomes in deaf children with cochlear implants. *Ear and Hearing,* 34, 179–192.

Hasegawa, T., Matsuki, K.-I., Ueno, T., Maeda, Y., Matsue, Y., Konishi, Y., et. al. (2004). Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. *Cognitive Brain Research,* 20, 510–518.

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

Hauthal, N., Sandmann, P., Debener, S., & Thorne, J. D. (2013). Visual movement perception in deaf and hearing individuals. *Advances in Cognitive Psychology,* 9, 53–61.

He, C., & Trainor, L. J. (2009). Finding the pitch of the missing fundamental in infants. *Journal of Neuroscience,* 29, 7718–7722.

Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron,* 76, 486–502.

Ho, Y. C., Cheung, M. C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology,* 17, 439–450.

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.

Houston, D. M., Pisoni, D. B., Kirk, K. I., Ying, E. A., & Miyamoto, R. T. (2003). Speech perception skills

of deaf infants following cochlear implantation: a first report. *International Journal of Pediatric Otorhinolaryngology, 67,* 479–495.

Houston, D. M., Santelmann, L. M., & Jusczyk, P. W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes, 19,* 97–136.

Hsiao, F., & Gfeller, K. (2012). Music perception of cochlear implant recipients with implications for music instruction: A review of literature. *Uppdate: Applications of Research in Music Education, 30,* 5–10.

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., et al. (2009). Musical training shapes structural brain development. *Journal of Neuroscience, 29,* 3019–3025.

Imfeld, A., Oechslin, M. S., Meyer, M., Loenneker, T., & Jancke, L. (2009). White matter plasticity in the corticospinal tract of musicians: A diffusion tensor imaging study. *Neuroimage, 46,* 600–607.

Ingvalson, E. M., Young, N. M., & Wong, P. C. M. (2014). Auditory-cognitive training improves language performance in prelingually deafened cochlear implant recipients. *International Journal of Pediatric Otorhinolaryngology, 78,* 1624–1631.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia, 48,* 2630–2639.

Jiwani, S., Papsin, B. C., & Gordon, K. A. (2013). Central auditory development after long-term cochlear implant use. *Clinical Neurophysiology, 124,* 1868–1880.

Johnson, J. M. (2009). Late auditory event-related potentials in children with cochlear implants: A review. *Developmental Neuropsychology, 34,* 701–720.

Jones, M. R. (1976). Time, our lost dimension – toward a new theory of perception, attention, and memory. *Psychological Review, 83,* 323–355.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39,* 159–207.

Jäncke, L. (2009). The plastic human brain. *Restorative Neurology and Neuroscience, 27,* 521–538.

Karabanov, A., Blom, O., Forsman, L., & Ullen, F. (2009). The dorsal auditory pathway is involved in performance of both visual and auditory rhythms. *Neuroimage, 44,* 480–488.

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.

Kiefer, J., Hohl, S., Sturzebecher, E., Pfennigdorff, T., & Gstoettner, W. (2001). Comparison of speech recognition with different speech coding strategies (SPEAK, CIS, and ACE) and their relationship to telemetric measures of compound action potentials in the nucleus CI 24M cochlear implant system. *Audiology, 40,* 32–42.

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.

Kirk, S. A., McCarthy, J. J, & Kirk, W. D. (1974). *Illinois test of psycholinguistic abilities ITPA - Revised edition: Examiner's Manual*. Illinois, USA: University of Illinois Press. Finnish version: Jyväskylä, Finland: Faculty of education, University of Jyväskylä.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., et al. (2005). Computerized training of working memory in children with ADHD - A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44,* 177–186.

Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature, 383,* 256–259.

Knight, R. T., & Scabini, D. (1998). Anatomic bases of event-related potentials and their relationship to novelty detection in humans. *Journal of Clinical Neurophysiology, 15,* 3–13.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America, 118,* 1038–1054.

Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *Neuroimage, 17,* 956–966.

Koelsch, S., Wittfoth, M., Wolf, A., Müller, J., & Hahne, A. (2004). Music perception in cochlear implant users: an event-related potential study. *Clinical Neurophysiology, 115,* 966–972.

Koistinen, S., Rinne, T., Cederström, S., & Alho, K. (2012). Effects of significance of auditory location changes on event related brain potentials and pitch discrimination performance. *Brain Research, 1427,* 44–53.

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.

Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., et al. (2010). Hemodynamic responses to speech and music in newborn infants. *Human Brain Mapping,* 31, 595–603.

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

Kral, A., Tillein, J., Hubka, P., Schiemann, D., Heid, S., Hartmann, R., et al. (2009). Spatiotemporal patterns of cortical activity with bilateral cochlear implants in congenital deafness. *Journal of Neuroscience,* 29, 811–827.

Kraus, N., Strait, D. L., & Parbery-Clark, A. (2012). Cognitive factors shape brain networks for auditory skills: Spotlight on auditory working memory. *Neurosciences and Music IV: Learning and Memory,* 1252, 100–107.

Kronenberger, W. G., Beer, J., Castellanos, I., Pisoni, D. B., & Miyamoto, R. T. (2014). Neurocognitive risk in children with cochlear implants. *Jama Otolaryngology-Head & Neck Surgery,* 140, 608–615.

Kronenberger, W. G., Pisoni, D. B., Henning, S. C., Colson, B. G., & Hazzard, L. M. (2011). Working memory training for children with cochlear implants: A pilot study. *Journal of Speech Language and Hearing Research,* 54, 1182–1196.

Kropotov, J. D., Näätänen, R., Sevostianov, A. V., Alho, K., Reinikainen, K., & Kropotova, O. V. (1995). Mismatch negativity to auditory stimulus change recorded directly from the human temporal cortex. *Psychophysiology,* 32, 418–422.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience,* 5, 831–843.

Kujala, T., Kuuluvainen, S., Saalasti, S., Jansson-Verkasalo, E., von Wendt, L., & Lepistö, T. (2010). Speech-feature discrimination in children with Asperger syndrome as determined with the multi-feature mismatch negativity paradigm. *Clinical Neurophysiology,* 121, 1410–1419.

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

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology,* 74, 1–19.

Kumar, S., Stephan, K. E., Warren, J. D., Friston, K. J., & Griffiths, T. D. (2007). Hierarchical processing of auditory objects in humans. *Plos Computational Biology,* 3, 977–985.

Kushnerenko, E. V., Van den Bergh, B. R. H., & Winkler, I. (2013). Separating acoustic deviance from novelty during the first year of life: a review of event-related potential evidence. *Frontiers in Psychology,* 4.

Kwon, B. J., & van den Honert, C. (2006). Dual-electrode pitch discrimination with sequential interleaved stimulation by cochlear implant users. *Journal of the Acoustical Society of America,* 120, EL1–EL6.

Laneau, J., & Wouters, J. (2004). Relative contributions of temporal and place pitch cues to fundamental frequency discrimination in cochlear implantees. *Journal of the Acoustical Society of America,* 116, 3606–3619.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review,* 106, 119–159.

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

Leaver, A. M., & Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: Effects of acoustic features and auditory object category. *Journal of Neuroscience,* 30, 7604–7612.

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 & Development,* 33, 419–430.

Lee, Y.-S., Lu, M.-J., & Ko, H.-P. (2007). Effects of skill training on working memory capacity. *Learning and Instruction,* 17, 336–344.

Levänen, S., Ahonen, A., Hari, R., McEvoy, L., & Sams, M. (1996). Deviant auditory stimuli activate human left and right auditory cortex differently. *Cerebral Cortex,* 6, 288–296.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America,* 32, 451–454.

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion,* 11, 1021–1031.

Limb, C. J., & Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hearing Research, 308,* 13–26.

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain, 133,* 1682–1693.

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

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: a new disconnection syndrome? *The Journal of Neuroscience, 29,* 10215–10220.

Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: The MIT Press.

Luo, X., Padilla, M., & Landsberger, D. M. (2012). Pitch contour identification with combined place and temporal cues using cochlear implants. *Journal of the Acoustical Society of America, 131,* 1325–1336.

Løvstad, M., Funderud, I., Lindgren, M., Endestad, T., Due-Tonnessen, P., Meling, T., et al. (2012). Contribution of subregions of human frontal cortex to novelty processing. *Journal of Cognitive Neuroscience, 24,* 378–395.

Lyxell, B., Wass, M., Sahlen, B., Samuelsson, C., Asker-Arnason, L., Ibertsson, T., et al. (2009). Cognitive development, reading and prosodic skills in children with cochlear implants. *Scandinavian Journal of Psychology, 50,* 463–474.

Macherey, O., & Delpierre, A. (2013). Perception of musical timbre by cochlear implant listeners: A multidimensional scaling study. *Ear and Hearing, 34,* 426–436.

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18,* 199–211.

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

Marie, C., Kujala, T., & Besson, M. (2012). Musical and linguistic expertise influence pre-attentive and attentive processing of non-speech sounds. *Cortex, 48,* 447–457.

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience, 23,* 294–305.

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19,* 1453–1463.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38,* 465–494.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General, 134,* 477–500.

May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology, 47,* 66–122.

McDermott, H. J. (2004). Music perception with cochlear implants: a review. *Trends in amplification, 8,* 49–82.

McDermott, H. J., & McKay, C. M. (1997). Musical pitch perception with electrical stimulation of the cochlea. *Journal of the Acoustical Society of America, 101,* 1622–1631.

McGurk, H. & MacDonald, J. (1976): Hearing lips and seeing voices. *Nature, 264,* 746–748.

McMullen, N. T., & Glaser, E. M. (1988). Auditory cortical responses to neonatal deafning – pyramidal neuron spine loss without changes in growth or orientation. *Experimental Brain Research, 72,* 195–200.

McMullen, N. T., Goldberger, B., Suter, C. M., & Glaser, E. M. (1988). Neonatal deafening alters nonpyramidal dendrite orientation in auditory cortex- a computer microscope study in the rabbit. *Journal of Comparative Neurology, 267,* 92–106.

Melara, R. D., & Marks, L. E., (1990a). Hard and soft interacting dimensions: differential effects of dual context on classification. *Perception & Psychophysics, 47,* 307–325.

Melara, R. D., & Marks, L. E. (1990b). Interaction among auditory dimensions: timbre, pitch, and loudness. *Perception & Psychophysics, 48,* 169–178.

Meister, H., Landwehr, M., Pyschny, V., Wagner, P., & Walger, M. (2011). The perception of sentence stress in cochlear implant recipients. *Ear and Hearing, 32,* 459–467.

Meyer, M., Elmer, S., Ringli, M., Oechslin, M. S., Baumann, S., & Jäncke, L. (2011). Long-term exposure

to music enhances the sensitivity of the auditory system in children. *European Journal of Neuroscience,* 34, 755–765.

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research,* 219, 36–47.

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

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology,* 41, 49–100.

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

Moore, B. C. J. (2003b). *An introduction to the psychology of hearing.* London, UK: Academic Press.

Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *JARO,* 9, 399–406.

Moore, B. C. J. (2014). Pitch: mechanisms underlying the pitch of pure and complex tones. In: A.N. Popper, A. N., & Fay, R. R. (eds.), *Perspectives on Auditory Research, Springer 379, Handbook of Auditory Research 50.* New York: Springer Science+Business Media.

Moore, J. K., & Guan, Y. L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *JARO,* 2, 297–311.

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

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

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience,* 3, 473–478.

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.

Nakata, T., Trehub, S. E., Mitani, C., Kanda, Y., Shibasaki, A., & Schellenberg, E. G. (2005). Music recognition by Japanese children with cochlear implants. *Journal of physiological anthropology and applied human science,* 24, 29–32.

Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: association with lexical tone agnosia. *Brain,* 133, 2635–2642.

Nelken, I., & Ulanovsky, N. (2007). Mismatch negativity and stimulus-specific adaptation in animal models. *Journal of Psychophysiology,* 21, 214.

Neville, H. J., & Lawson, D. (1987). Attention to central and peripheral visual space in a movement detection task: An event-related potential and behavioral study. II. Congenitally deaf adults. *Brain Research,* 405, 268–283.

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: Influence of music training. *Ear and Hearing,* 30, 432–446.

Nikolopoulos, T. P., & Vlastarakos, P. V. (2010). Treating options for deaf children. *Early Human Development,* 86, 669–674.

Nimmons, G. L., Kang, R. S., Drennan, W. R., Longnion, J., Ruffin, C., Worman, T., et al. (2008). Clinical assessment of music perception in cochlear implant listeners. *Otology & Neurotology,* 29, 149–155.

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

Nobbe, A., Schleich, P., Zierhofer, C., & Nopp, P. (2007). Frequency discrimination with sequential or simultaneous stimulation in MED-EL cochlear implants. *Acta Oto-Laryngologica,* 127, 1266–1272.

Norena, A. J., Gourevitch, B., Aizawa, N., & Eggermont, J. J. (2006). Spectrally enhanced acoustic environment disrupts frequency representation in cat auditory cortex. *Nature Neuroscience,* 9, 932–939.

Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology,* 42, 25–32.

Näätänen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: A unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology, 48,* 4–22.

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., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology, 115,* 140–144.

Näätänen, R., & Picton, T. W. (1986). N2 and automatic versus controlled processes. *Electroencephalography and Clinical Neurophysiology, Supplement, 38,* 169–186.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound – a review and an analysis of the component structure. *Psychophysiology, 24,* 375–425.

Obleser, J., Boecker, H., Drzezga, A, Haslinger, B., Hennenlotter, A., Roettinger, M., et al. (2006). Vowel sound extraction in anterior superior temporal cortex. *Human Brain Mapping, 27,* 562–571.

Obleser, J., Zimmermann, J., Van Meter, J., & Rauschecker, J. P. (2007). Multiple stages of auditory speech perception reflected in event-related fMRI. *Cerebral Cortex, 17,* 2251–2257.

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

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development, 59,* 441–449.

Olszewski, C., Gfeller, K., Froman, R., Stordahl, J., & Tomblin, B. (2005). Familiar melody recognition by children and adults using cochlear implants and normal hearing children. *Cochlear Implants International, 6,* 123–140.

Opitz, B., Mecklinger, A., von Cramon, D. Y., & Kruggel, F. (1999). Combining electrophysiological and hemodynamic measures of the auditory oddball. *Psychophysiology, 36,* 142–147.

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y., & Schröger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage, 15,* 167–174.

Overy, K., & Turner, R. (2009). The rhythmic brain. *Cortex, 45,* 1–3.

Oxenham, A. J., Micheyl, C., Keebler, M. V., Loper, A., & Santurette, S. (2011). Pitch perception beyond the traditional existence region of pitch. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 7629–7634.

Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., & Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) potential of the auditory event-related (ERP). *Clinical Neurophysiology, 118,* 177–185.

Pantev, C., & Herholz, S. C. (2011). Plasticity of the human auditory cortex related to musical training. *Neuroscience and Biobehavioral Reviews, 35,* 2140–2154.

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

Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., &, Huotilainen, M. (2013a). Learning-induced neural plasticity of speech processing before birth. *PNAS, 110,* 15145–15150.

Partanen, E., Kujala, T., Tervaniemi, M., & Huotilainen, M. (2013b). Prenatal music exposure induces long-term neural effects. *Plos One, 8.*

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

Patel, A. D., Foxton, J. M., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition, 59,* 310–313.

Patel, A. D., Wong, M., Foxton, J., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception, 25,* 357–368.

Patston, L. L. M., Corballis, M. C., Hogg, S. L., & Tippett, L. J. (2006). The neglect of musicians: Line bisection reveals an opposite bias. *Psychological Science, 17,* 1029–1031.

Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain, 132,* 1277–1286.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders - The Montreal battery of evaluation of amusia. *Neurosciences and Music, 999,* 58–75.

Peretz, I., Gosselin, N., Tillmann, B., Cuddy, L. L., Gagnon, B., Trimmer, C. G., et al. (2008). On-line identification of congenital amusia. *Music Perception, 25,* 331–343.

Petersen,B., Mortensen, M.V., Hansen, M., & Vuust, P. (2012). Singing in the key of life: A pilot study on effects of musical ear training after cochlear implantation. *Psychomusicology, 22*, 134–151.

Petersen, B., Weed, E., Sandmann, P., Brattico, E., Hansen, M., Sørensen, S. D., et al. (2015). Brain responses to musical feature changes in adolescent cochlear implant users. *Frontiers in Human Neuroscience, 9*.

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Turgeon, C., Lepore, F., & Peretz, I. (2015). Cochlear implant users move in time to the beat of drum music. *Hearing Research, 321*, 25–34.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308*, 1430–1430.

Picton, T. W. (2010). *Human auditory evoked potentials*. San Diego, CA: Plural Publishing Inc.

Picton, T. W., & Taylor, M. J. (2007). Electrophysiological evaluation of human brain development. *Developmental Neuropsychology, 3*, 249–278.

Pijl, S. (1997). Labeling of musical interval size by cochlear implant patients and normally hearing subjects. *Ear and Hearing, 18*, 364–372.

Ping, L., Yuan, M., & Feng, H. (2012). Musical pitch discrimination by cochlear implant users. *Annals of Otology Rhinology and Laryngology, 121*, 328–336.

Pisoni, D. B., & Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing, 24*, 106S–120S.

Pisoni, D. B., Kronenberger, W. G., Roman, A. S., & Geers, A. E. (2011). Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear and Hearing, 32*, 60S–74S.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology-Human Perception and Performance, 20*, 976–986.

Plack, C. J., Barker, D., & Hall, D. A. (2014). Pitch coding and pitch processing in the human brain. *Hearing Research, 307*, 53–64.

Polich, J., Howard, L., & Starr, A. (1983). P300 latency correlates with digit span. *Psychophysiology, 20*, 665–669.

Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *Journal of Neuroscience, 26*, 4970–4982.

Ponton, C. W., Don, M., Eggermont, J. J., Waring, M. D., Kwong, B., & Masuda, A. (1996b). Auditory system plasticity in children after long periods of complete deafness. *Neuroreport, 8*, 61–65.

Ponton, C. W., Don, M., Eggermont, J. J., Waring, M. D., & Masuda, A. (1996a). Maturation of human cortical auditory function: Differences between normal-hearing children and children with cochlear implants. *Ear and Hearing, 17*, 430–437.

Ponton, C. W., & Eggermont, J. J. (2001). Of kittens and kids: Altered cortical maturation following profound deafness and cochlear implant use. *Audiology and Neuro-Otology, 6*, 363–380.

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

Potter, D. D., Fenwick, M., Abecasis, D., & Brochard, R. (2009). Perceiving rhythm where none exists: Event-related potential (ERP) correlates of subjective accenting. *Cortex, 45*, 103–109.

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.

Putkinen, V., Tervaniemi, M., Saarikivi, K., Ojala, P., & Huotilainen, M. (2014). Enhanced development of auditory change detection in musically trained school-aged children: A longitudinal event-related potential study. *Developmental Science, 17*, 282–297.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience, 12*, 718–724.

Reybrouck, M., & Brattico, E. (2015). Neuroplasticity beyond sounds: Neural adaptations following long-term musical aesthetic experiences. *Brain Sciences, 5*, 69–91.

Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *Neuroimage, 12*, 14–19.

Riss, D., Hamzavi, J-S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., Baumgartner, W-D., Gstoettner, W., & Arnoldner, C. (2014). FS4, FS4-p, and FSP: A 4-month crossover study of 3 fine structure sound-coding strategies. *Ear and Hearing, 35*, e272–e281.

Rocca, C. (2012). A different musical perspective: Improving outcomes in music through habilitation, education, and training for children with cochlear implants. *Seminars in Hearing, 33*, 425–433.

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

Roden, I., Kreutz, G., & Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: a longitudinal study. *Frontiers in Psychology,* 3.

Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: Temporal and structural factors investigated using functional magnetic resonance imaging. *Journal of Neuroscience, 31,* 3843–3852.

Ronkainen, R. (2011). Enhancing listening and imitation skills in children with cochlear implants: the use of multimodal resources in speech and language therapy. *Journal of Interactional Research in Communication Disorders, 2,* 245–269.

Ross, B., Snyder, J. S., Aalto, M., McDonald, K. L., Dyson, B. J., Schneider, B., et al. (2009). Neural encoding of sound duration persists in older adults. *Neuroimage, 47,* 678–687.

Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition, 99,* 113–129.

Saarikallio, S. (2010). Music as emotional self-regulation throughout adulthood. *Psychology of Music, 39,* 307–327.

Salmi, J., Rinne, T., Koistinen, S., Salonen, O., & Alho, K. (2009). Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain Research, 1286,* 155–164.

Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology, 119,* 332–341.

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.

Särkamö, T., Tervaniemi, M., Laitinen, S., Numminen, A., Kurki, M., Johnson, J. K., et al. (2014). Cognitive, emotional, and social benefits of regular musical activities in early dementia: Randomized controlled Study. *Gerontologist, 54,* 634–650.

Schorr, E. A., Fox, N. A., van Wassenhove, V., & Knudsen, E. I. (2005). Auditory–visual fusion in speech perception in children with cochlear implants. *PNAS, 102,* 18748–18750.

Schön, D., Gordon, R., Campagne, A., Magne, C., Astesano, C., Anton, J.-L., et al. (2010). Similar cerebral networks in language, music and song perception. *Neuroimage, 51,* 450–461.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41,* 341–349.

Schönwiesner, M., Novitski, N., Pakarinen, S., Carlson, S., Tervaniemi, M., & Näätänen, R. (2007). Heschl's gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes. *Journal of Neurophysiology, 97,* 2075–2082.

Schröger, E., Giard, M. H., & Wolff, C. (2000). Auditory distraction: event-related potential and behavioral indices. *Clinical Neurophysiology, 111,* 1450–1460.

Seppänen, M., Pesonen, A.-K., & Tervaniemi, M. (2012). Music training enhances the rapid plasticity of P3a/P3b event-related brain potentials for unattended and attended target sounds. *Attention Perception & Psychophysics, 74,* 600–612.

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

Sharma, A., Campbell, J., & Cardon, G (2015). Developmental and cross-modal plasticity in deafness: Evidence from the P1 and N1 event related potentials in cochlear implanted children. *International Journal of Psychophysiology, 95,* 135–144.

Sharma, A., Dorman, M., Spahr, A., & Todd, N. W. (2002b). Early cochlear implantation in children allows normal development of central auditory pathways. *The Annals of Otology, Rhinology & Laryngology. Supplement, 189,* 38–41.

Sharma, A., Dorman, M. F., & Kral, A. (2005). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hearing Research, 203,* 134–143.

Sharma, A., Dorman, M. F., & Spahr, A. J. (2002a). Rapid development of cortical auditory evoked potentials after early cochlear implantation. *Neuroreport, 13,* 1365–1368.

Sharma, A., Gilley, P. M., Dormant, M. F., & Baldwin, R. (2007). Deprivation-induced cortical reorganization in children with cochlear implants. *International Journal of Audiology, 46,* 494–499.

Sharma, A., Kraus, N., McGee, T. J., & Nicol, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Evoked Potentials-Electroencephalography and Clinical Neurophysiology, 104*, 540–545.

Sharma, A., Nash, A. A., & Dorman, M. (2009). Cortical development, plasticity and re-organization in children with cochlear implants. *Journal of Communication Disorders, 42*, 272–279.

Singer, J. & Wilett, J. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. USA: Oxford University Press.

Stabej, K. K., Smid, L., Gros, A., Zargi, M., Kosir, A., & Vatovec, J. (2012). The music perception abilities of prelingually deaf children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology, 76*, 1392–1400.

Steinbrink, C., Groth, K., Lachmann, T., & Riecker, A. (2012). Neural correlates of temporal auditory processing in developmental dyslexia during German vowel length discrimination: An fMRI study. *Brain and Language, 121*, 1–11.

Stevens, K. N. (1998). *Acoustic phonetics*. London, UK: The MIT Press.

Stöbich, B., Zierhofer, C. M., & Hochmair, E. S. (1999). Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy. *Ear and Hearing, 20*, 104–116.

Stordahl, J. (2002). Song recognition and appraisal: A comparison of children who use cochlear implants and normally hearing children. *Journal of Music Therapy, 39*, 2–19.

Straatman, L. V., Rietveld, A. C. M., Beijen, J., Mylanus, E. A. M., & Mens, L. H. M. (2010). Advantage of bimodal fitting in prosody perception for children using a cochlear implant and a hearing aid. *Journal of the Acoustical Society of America, 128*, 1884–1895.

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.

Sucher, C. M., & McDermott, H. J. (2007). Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. *Hearing Research, 230*, 80–87.

Takahashi, H., Rissling, A. J., Pascual-Marqui, R., Kirihara, K., Pela, M., Sprock, J., et al. (2013). Neural substrates of normal and impaired preattentive sensory discrimination in large cohorts of nonpsychiatric subjects and schizophrenia patients as indexed by MMN and P3a change detection responses. *Neuroimage, 66*, 594–603.

Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences, 29*, 382–390.

Tervaniemi, M., & Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Research Reviews, 43*, 231–246.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Experimental Brain Research, 161*, 1–10.

Tervaniemi, M., Medvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., van Zuijen, T. L., et al. (2000). Lateralized automatic auditory processing of phonetic versus musical information: A PET study. *Human Brain Mapping, 10*, 74–79.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*, 53–71.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion, 4*, 46–64.

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research, 16*, 145–161.

Timm, L., Agrawal, D., Viola, F. C., Sandmann, P., Debener, S., Büchner, A., et al. (2012). Temporal feature perception in cochlear implant users. *Plos One, 7*.

Timm, L., Vuust, P., Brattico, E., Agrawal, D., Debener, S., Büchner, A., et al. (2014). Residual neural processing of musical sound features in adult cochlear implant users. *Frontiers in Human Neuroscience, 8*, 181–181.

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review, 9*, 335–340.

Trainor, L. J., Desjardins, R. N., & Rockel, C. (1999). A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Australian Journal of Psychology, 51*, 147–153.

Trehub, S. E., & Thorpe, L. A. (1989). Infant's perception of rhythm – Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology-Revue Canadienne De Psychologie, 43*,

217–229.

Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1987). Organizational processes in infants perception of auditory patterns. *Child Development, 58,* 741–749.

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.

Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *Neuroreport, 9,* 3557–3560.

Trollinger, V. L. (2003). Relationships between pitch-matching accuracy, speech fundamental frequency, speech range, age, and gender in American English-speaking preschool children. *Journal of Research in Music Education, 51,* 78–95.

Vainio, M., & Järvikivi, J. (2007). Focus in production: Tonal shape, intensity and word order. *Journal of the Acoustical Society of America, 121,* EL55–EL61.

Välimaa, T. T., Määttä, T. K., Löppönen, H. J., & Sorri, M. J. (2002a). Phoneme recognition and confusions with multichannel cochlear implants: Consonants. *Journal of Speech Language and Hearing Research, 45,* 1055–1069.

Välimaa, T. T., Määttä, T. K., Löppönen, H. J., & Sorri, M. J. (2002b). Phoneme recognition and confusions with multichannel cochlear implants: Vowels. *Journal of Speech Language and Hearing Research, 45,* 1039–1054.

van Zuijen, T. L., Simoens, V. L., Paavilainen, P., Näätänen, R., & Tervaniemi, M. (2006). Implicit, intuitive, and explicit knowledge of abstract regularities in a sound sequence: An event-related brain potential study. *Journal of Cognitive Neuroscience, 18,* 1292–1303.

Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. D., & McDermott, H. J. (2005). Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies. *Journal of the Acoustical Society of America, 117,* 3126–3138.

Virtala, P., Huotilainen, M., Putkinen, V., Makkonen, T., & Tervaniemi, M. (2012). Musical training facilitates the neural discrimination of major versus minor chords in 13-year-old children. *Psychophysiology, 49,* 1125–1132.

Vogel, I., & Raimy, E. (2002). The acquisition of compound vs. phrasal stress: The role of prosodic constituents. *Journal of Child Language, 29,* 225–250.

Volpe, U., Mucci, A., Bucci, P., Merlotti, E., Galderisi, S., & Maj, M. (2007). The cortical generators of P3a and P3b: A LORETA Study. *Brain Research Bulletin, 73,* 220–230.

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language, 38,* 133–149.

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music - Brain responses to rhythmic incongruity. *Cortex, 45,* 80–92.

Wan, C. Y., & Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *Neuroscientist, 16,* 566–577.

Wan, C. Y., Zheng, X., Marchina, S., Norton, A., & Schlaug, G. (2014). Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. *Brain and Language, 136,* 1–7.

Wang, S., Liu, B., Dong, R., Zhou, Y., Li, J., Qi, B., et al. (2012). Music and lexical tone perception in Chinese adult cochlear implant users. *Laryngoscope, 122,* 1353–1360.

Warren, J. D., Jennings, A. R., & Griffiths, T. D. (2005). Analysis of the spectral envelope of sounds by the human brain. *Neuroimage, 24,* 1052–1057.

Wechsler, D. (1997). Wechsler adult intelligence scale, 3rd Edn. New York, NY: Psychological Corporation.

Wechsler, D. (2005). Wechsler adult intelligence scale, 3rd Edn. Helsinki: Psykologien Kustannus Oy.

Wechsler D. (2010). Wechsler Intelligence Scale for Children – 4rd Edn: Manual. Helsinki: Psykologien Kustannus Oy.

Welch, G. F. (1985). A schema theory of how children learn to sing in-tune. *Psychology of Music, 13,* 3–18.

Wells, B., Peppe, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31,* 749–778.

West, B. T. (2009). Analyzing longitudinal data with the Linear Mixed Models procedure in SPSS. *Evaluation & 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.

Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research, 242,* 3–21.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature, 352,* 236–238.

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., Haden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 2468–2471.

Winkler, I., Tervaniemi, M., Schröger, E., Wolff, C., & Näätänen, R. (1998). Preattentive processing of auditory spatial information in humans. *Neuroscience Letters, 242,* 49–52.

Woods, D. L., & Alain, C. (2009). Functional imaging of human auditory cortex. *Current Opinion in Otolaryngology & Head and Neck Surgery, 17,* 407–411.

Woods, D. L., Stecker, G. C., Rinne, T., Herron, T. J., Cate, A. D., Yund, E. W., et al. (2009). Functional maps of human auditory cortex: Effects of acoustic features and attention. *Plos One, 4.*

Yabe, H., Saito, F., & Fukushima, Y. (1993). Median method for detecting endogenous event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 87, 403–407.

Yucel, E., Sennaroglu, G., & Belgin, E. (2009). The family oriented musical training for children with cochlear implants: Speech and musical perception results of two year follow-up. *International Journal of Pediatric Otorhinolaryngology, 73,* 1043–1052.

Zatorre, R. J., & Salimpoor, V. N. (2013). From perception to pleasure: Music and its neural substrates. *Proceedings of the National Academy of Sciences of the United States of America, 110,* 10430–10437.

Zeng, F.-G. (2002). Temporal pitch in electric hearing. *Hearing Research, 174,* 101–106.

Zeng, F.-G. (2004). Trends in cochlear implants. *Trends in Amplification, 8,* 1–34.

APPENDIX 1. The clusters extracted from the questionnaire, the questions included in each cluster and partial correlations (age controlled; r_p) between the mean of the answers included in cluster A and the answers given by the parents.

Clusters	Questions	r_p
Cluster A	B20A How often have the siblings played an instrument with the child between measurements (the child has been playing or singing along)?[1]	.046
	(the child has been playing or singing along)?[1]	.407
"Music	b23 How often has your child heard his/her parents play during the last year?[1]	.425
activity at	B28 How often has your child heard his/her parents play an instrument between measurements?[1]	.315
home"	b3 Does your child play an instrument at home? If yes, how often would you estimate?[1]	.641**
	b8 Does/did your child's daycare include music or singing hours? How many times a week?	-.044
	b28 How often has your child heard his/her parents play on previous years?[1]	.232
	b29 How often has your child heard his/her parents play during the first year after implantation?[1]	.309
	b15 How often has your child heard his/her siblings play an instrument during the last year?[1]	.512
	b17 How often has your child heard his/her siblings play on previous years?[1]	.524
	b18 How often has your child heard his/her siblings sing on previous years?[1]	**.710**
	B19 How often has your child heard his/her siblings play an instrument between measurements?[1]	.684**
	B1A Has your child been playing an instrument at home between measurements?[1]	.698**
	B22B How often have the siblings sung with the child before first measurements (child has been playing or singing along)?[1]	.569*
	b16 How often has your child heard his/her siblings sing during the last year?[1]	**.743***
	B22A How often have the siblings sung with the child between measurements (child has been playing or singing along)?[1]	.367
	b24 How often has your child heard his/her parents sing during the last year?[1]	.707**
	b26 How often has your child heard his/her parents sing on previous years?[1]	**.732***
	b27 How often has your child heard his/her parents sing during the first year after implantation?[1]	.641**
	B2A Has your child been singing at home during the time between measurements?[1]	.414
	B21 How often has your child heard his/her siblings sing between measurements?[1]	.633**
	B23 How often did you parents sing in front of the child between measurements?[1]	.699**
	B24 How often did you parents sing interacting with your child i.e. the child was listening to you keeping ye contact with you and/or tried to participate in the singing (e.g. sang along) between measurements?[1]	.490*
	B26 How often did you parents sing interacting with your child (see B24 above) before the previous measurements?[1]	.398
Cluster B	B10a If the child responds to the music on TV, how does he/she respond? a. gets anxious or irritated; b. smiles or laughs; c. makes sounds; d. claps spontaneously; e. dances spontaneously; f. moves according to the song spontaneously; g. sings lyrics spontaneously; h. asks questions; i. never responds in any way; j. other.[2]	
	B11 How many times a week did your child watch (and listen to) children's music videos or DVDs between measurements?_ Less frequently than weekly_	
	B14 How many times a week did your child watch (and listen to) children's music videos or DVDs before the measurements?_ Less frequently than weekly_	
Cluster C	B15 How many times a week did your child listen to music from CD:s (without visualization) before the measurements?_ Less frequently than weekly_	
Cluster D	B4 Does your child sing at home? If yes, how often?[1]	
Cluster E	E0 How many times in a week does the child have music lessons at school/daycare?	
Cluster F	b11a How many times a week has your child been listening to music (CDs, DVDs, television) on his/her free time (at home, car journeys etc.) during the last year?_ Less frequently than weekly _	
	b11b How many times a week has your child been listening to music (CDs, DVDs, television) on his/her free time (at home, car journeys etc.) before the last year?_ Less frequently than weekly _	
Cluster G	B13 How many times a week did your child watch (and listen to) children's programs, videos or DVDs that had singing and other music in the background before the previous measurements?_ Less frequently than weekly _	
Cluster H	b7 Has your child had a supervised music hobby already previously for example in a music school? What kind of hobby was it? (e.g. musical play school, rhythm group, band, playing an instrument). For how many months has your child had the hobby?	
Cluster I	E4 How many minutes in a week does the child have music lessons/singing at school/daycare?	
Cluster J	b10 Has your child attended other supervised musical activities outside the home? (e.g. ballet, other dance, rhythmic gymnastics, aerobics)? For how many months approximately has the child attended the activities?	

*Children with CIs, df = 17; Normal hearing children, df = 20 – 24; * p ≤ .050; ** p ≤ .010; *** p < .001; b = question at T1; B = question at T2; B and b were answered by parents; E = was answered by personnel at school or daycare; [1]Every week_ every other week _ occasionally_ not at all_ if weekly, how many times a week; [2]Based on van Besouv et al, 2010.*