Comparison of Two Music Training Approaches on Music and Speech Perception in Cochlear Implant Users

Christina D. Fuller[1,2,3], John J. Galvin III[1,2,3,4,5], Bert Maat[1,2,3], Deniz Başkent[1,2,3], and Rolien H. Free[1,2,3]

Abstract

In normal-hearing (NH) adults, long-term music training may benefit music and speech perception, even when listening to spectro-temporally degraded signals as experienced by cochlear implant (CI) users. In this study, we compared two different music training approaches in CI users and their effects on speech and music perception, as it remains unclear which approach to music training might be best. The approaches differed in terms of music exercises and social interaction. For the pitch/timbre group, melodic contour identification (MCI) training was performed using computer software. For the music therapy group, training involved face-to-face group exercises (rhythm perception, musical speech perception, music perception, singing, vocal emotion identification, and music improvisation). For the control group, training involved group nonmusic activities (e.g., writing, cooking, and woodworking). Training consisted of weekly 2-hr sessions over a 6-week period. Speech intelligibility in quiet and noise, vocal emotion identification, MCI, and quality of life (QoL) were measured before and after training. The different training approaches appeared to offer different benefits for music and speech perception. Training effects were observed within-domain (better MCI performance for the pitch/timbre group), with little cross-domain transfer of music training (emotion identification significantly improved for the music therapy group). While training had no significant effect on QoL, the music therapy group reported better perceptual skills across training sessions. These results suggest that more extensive and intensive training approaches that combine pitch training with the social aspects of music therapy may further benefit CI users.

Introduction

Cochlear implants (CIs) are prosthetic devices that enable severely deafened individuals to hear again. After speech, music is the second most important auditory signal for CI users. However, adult CI users have difficulty with music perception (Drennan & Rubinstein, 2008; Gfeller et al., 2000; Philips et al., 2012). Music perception is much poorer in CI users compared to normal-hearing (NH) listeners (Limb & Roy, 2014; McDermott, 2004), and CI users report low levels of music enjoyment (Fuller et al., 2013; Lassaletta et al., 2008; McDermott, 2004). Device-related factors, patient-related factors, and the nature of electric stimulation all contribute to the relatively poor music perception and enjoyment in CI users (for an overview, see

[1]Department of Otorhinolaryngology/Head and Neck Surgery, University Medical Center Groningen, University of Groningen, the Netherlands
[2]Graduate School of Medical Sciences, University of Groningen, the Netherlands
[3]Research School of Behavioral and Cognitive Neurosciences, University of Groningen, the Netherlands
[4]House Ear Institute, Los Angeles, CA, USA
[5]Department of Head and Neck Surgery, David Geffen School of Medicine, UCLA, CA, USA

Corresponding author:
Christina D. Fuller, Department of Otorhinolaryngology/Head and Neck Surgery, University Medical Center Groningen, University of Groningen, BB20 P.O. Box 30.001, 9700 RB Groningen, the Netherlands.
Email: c.d.fuller@umcg.nl

Başkent, Gaudrain, Tamati, & Wagner, 2016; Limb & Roy, 2014; Looi, Gfeller, & Driscoll, 2012). Because of the limited insertion depth and the position of the electrodes relative to healthy neurons, there is often a tonotopic mismatch between the acoustic input and the cochlear place of stimulation. Because of the limited number of electrodes, and spread of excitation, there is only limited spectral resolution. The direct electric stimulation of the nerve only gives an approximation of the fine-tuned nerve responses to normal acoustic stimuli. As such, CI users are only provided with coarse spectral envelope information along with slowly varying temporal envelope information. While speech perception in quiet is possible using primarily temporal envelope cues (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995), music requires fine-structure cues that are important for perceiving the rich and dynamic acoustic cues of music, including pitch (Mehta & Oxenham, 2017; Shannon, Fu, & Galvin, 2004; Smith, Delgutte, & Oxenham, 2002). These fine structure cues are generally not provided or well-perceived in CIs. Thus, CI users listen to a spectro-temporally degraded, tonotopically mismatched representation of sound, which greatly limits music perception and appreciation (see for reviews Limb & Roy, 2014; Looi & She, 2010; Looi et al., 2012; McDermott, 2004). Among the primary musical elements (rhythm, pitch, melody, and timbre), only rhythm is well represented by CIs, with comparable rhythm perception between NH and CI listeners (Gfeller et al., 2007; Kong, Cruz, Jones, & Zeng, 2004). CI users may also experience deafness-related changes in the auditory system that may affect music perception (Limb & Roy, 2014; Looi et al., 2012). Postlingually deafened CI users often experience a period of auditory deprivation with different effects on the peripheral and central auditory pathway. Also the etiology of the hearing loss and survival patterns of spiral ganglia play an important role (Blamey et al., 2013). These patient-related factors add to device-related factors and can further degrade music perception (Başkent et al., 2016; Limb & Roy, 2014; Looi et al., 2012; McDermott & Oxenham, 2008).

There are two general approaches to improving music perception in CI users: improvement of the device or improvement of CI users' perceptual abilities. This study is based on the latter approach, using musical training to improve perception. Recent research with NH listeners has shown that years of intensive, long-term music training, as is typically experienced by musicians, may benefit pitch perception (Besson, Schon, Moreno, Santos, & Magne, 2007; Marques, Moreno, Castro, & Besson, 2007), rhythm perception (Chen, Penhune, & Zatorre, 2008), vocal identification (Dmitrieva, Gel'man, Zaitseva, & Orlov, 2006; Thompson, Schellenberg, & Husain, 2004), and voice timbre identification (Chartrand & Belin, 2006). Başkent & Gaudrain

(2016) showed a large musician advantage for speech understanding in the presence of competing speech, a task that depends strongly on segregation according to voice cues, including voice pitch (Assmann & Summerfield, 1990; Brungart, 2001). However, previous studies have also shown mixed results for musician advantages in speech perception (Boebinger et al., 2015; Clayton et al., 2016; Deroche, Limb, Chatterjee, & Gracco, 2017; Maden, Whiteford, & Oxenham, 2017; Morse-Fortier, Parrish, Baran, & Freyman, 2017; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Ruggles, Freyman, & Oxenham, 2014; Swaminathan et al., 2015; Zendel & Alain, 2012), with some studies showing substantial musician advantages and others showing only weak effects. Thus, while musical training clearly benefits music perception, a within-domain effect, the benefits for speech perception, a cross-domain effect, are less clear.

When auditory signals are degraded, as in the case of CI, very little is known about the effects of long-term music training on auditory, music, and speech perception. Fuller et al. (2014) studied NH musicians ($\geqslant 10$ years of music training) and nonmusicians listening to acoustic simulations of CI signal processing. While performance was poorer with the CI simulations than with unprocessed signals, the musician advantage for music perception persisted in the CI simulations. However, musician advantages for speech perception with the CI simulations were limited, with no advantage for word or sentence identification in quiet and most noise conditions, but with a significant advantage for vocal emotion identification, which depends partially on the perception of voice pitch cues (Gilbers et al., 2015).

Music training in CI users has been shown to improve music perception in terms of melodic contour identification (MCI), familiar melody recognition, timbre identification, and musical pitch perception (Fu, Galvin, Wang, & Wu, 2015; Galvin, Eskridge, Oba, & Fu, 2012; Galvin, Nogaki, & Fu, 2007; Gfeller et al., 2002; Oba, Fu, & Galvin, 2011; Petersen, Mortensen, Hansen, & Vuust, 2012; Vandali, Sly, Cowan, & Van Hoesel, 2015). However, it remains unclear whether music training can also improve speech perception in CI users. Petersen et al. (2012) investigated music training in newly implanted pre- and postlingually deafened adult CI users; there was also a control group of CI users that received no music training. Music training consisted of weekly 1-hr private music training for 6 months. The training focused on pitch, rhythm, and timbre via singing, playing instruments, and listening exercises. Both the training and control groups significantly improved their speech perception after 6 months of training. The authors concluded that this effect may not have been because of music training per se, as adaptation to electric hearing during the first 6 months of implant use may have contributed to improved performance in both groups.

However, the music training group did exhibit better overall music perception, as well as accelerated identification of emotional prosody, compared with the control group. Lo, McMahon, Looi, and Thompson (2015) studied the effects of MCI training on speech perception in CI users. Results showed improved consonant recognition and speech prosody perception after training, but no benefit for sentence recognition in babble. Looi, Wong, and Loo (2016) compared a music appreciation training program (MATP) with a focused music listening (FML) training program in CI users. In the MATP training, participants listened to various pieces of music and then were tested for discrimination of these pieces. In the FML training, participants listened to music while performing other tasks; the FML group served as a control for the MATP group, in that music perception was not explicitly trained. While music perception significantly improved for the MATP group, there was no improvement in speech understanding in noise for either group. Taken together, these studies suggest possible cross-domain effects for music training in CI users.

Auditory training using speech stimuli has been shown to be effective in CI users (Fu & Galvin, 2008; Ingvalson, Lee, Fiebig, & Wong, 2013; Oba et al., 2011; Stacey & Summerfield, 2007, 2008; Stacey et al., 2010). Bottom-up auditory training (e.g., with simple stimuli or phonemes) has been shown to improve both perception of trained (within-domain) and untrained (cross-domain) stimuli (Amitay, Hawkey, & Moore, 2005; Moore, Rosenberg, & Coleman, 2005; Wright, Buonomano, Mahncke, & Merzenich, 1997). Top-down training may improve central cognitive processing which may help CI users to extract cues from degraded signals in general (Fu & Galvin, 2007; Gfeller, 2001). It remains unclear which approach to music training might be best to improve both music and speech perception (Looi et al., 2012; Gfeller, Guthe, Driscoll, & Brown, 2015).

Besides the potential benefits seen in auditory perception, music training may also be beneficial for subjective factors. Music therapy has been shown to positively influence the quality of life (QoL) in different patient populations (terminally ill patients in Hilliard, 2003; elective brain surgery patients in Walworth, Rumana, Nguyen, & Jarred, 2008). Recently, Hütter, Argstatter, Grapp, and Plinkert (2015) studied the benefits of individualized music therapy program, which involved ten 50-min sessions that were specifically addressed to individual needs of adult CI users. The program focused on the perception of musical stimuli, speech prosody, and complex acoustic situations, and training was begun shortly after initial activation of the speech processor. The preliminary results showed improvements in subjective reports of music perception and overall hearing.

In this study, two musical training approaches and one nonmusical control group were compared in postlingually deafened adult CI users: (a) *Pitch/timbre:* Individual computer-based pitch and timbre perception training (as in Galvin et al., 2007, 2012; Lo et al., 2015); (b) *Music therapy:* Group music therapy, which included both listening to and playing music; and (c) *Control:* Group therapy that did not include music or auditory training. These approaches differed in several ways: social interaction (individual computer training vs. group therapy), methodology (auditory-only vs. auditory-motor vs. nonmusical training), environment (static computer-based training vs. dynamic group therapy), and perceptual mechanism (more bottom up with the computer-based pitch and timbre training vs. more top down with the group therapy). Research questions included: (a) Can pitch/timbre training or group music therapy improve CI users' perception of music (within-domain effect) and speech (cross-domain effect)? (b) Which training method is most effective for CI users? Answers to these questions may indicate whether computer-based music training or group music therapy could be a valuable addition to the current CI rehabilitation programs.

Methods

Participants

In total, 19 postlingually deafened, adult CI users were recruited via the University Medical Center Groningen (UMCG). All participants were native Dutch speakers, had used their CI for longer than 1 year, and had no neurological disorders. Table 1 shows demographic characteristics for the three participant groups. The mean age at testing was 69.1 years (range = 56–80). The mean age at implantation was 62.8 years (range = 46–77). The mean amount of CI experience was 6.3 years (range = 3–13). One participant was a bilateral CI user and four participants were bimodal CI users. Because of the small number of participants, no across-group matching was attempted in terms of demographic variables (e.g., gender, age at testing, CI experience, etc.): Participants were randomly distributed across groups. Before the study started, written and oral information about the protocol was provided, and written informed consent was obtained from all participants. Travel costs and testing time were reimbursed in accordance with the department policy.

Test Stimuli and Procedures

The overall study design is illustrated in Figure 1. Before (Week 1) and after training (Week 8), all participants were tested for a variety of speech and music perception tasks; QoL was also assessed using a questionnaire. These are the same outcome measures as used by

Table 1. CI Participant Demographic Information.

Training group	Participant	Age at test (yrs)	Age at CI (years)	Deaf age (years)	CI exp (years)	Etiology	Device	Strategy	Word identification (% correct)
	A1	70	58	46	12	Unknown	CI24R K	Ace	83
	A2	71	68	30	3	Progressive hearing loss	CI24RE CA	Ace	89
	A3	78	75	10	3	Unknown	CI24RE CA	MP3000	78
Pitch/timbre	A4	73	63	27	10	Unknown	CI24R CA	Ace	75
	A5	73	68	35	5	Unknown	HiRes 90K Helix	HiRes-Sw/ Fidelity	94
	A6	73	68	35	5	Unknown	CI512	Ace	64
	B1	57	46	8	11	Unknown	CI24R CS	Ace	89
	B2	56	51	Unknown	5	Unknown	Hires90khelix	HiRes-S	72
	B3	67	61	Unknown	6	Unknown	Hires90khelix	HiRes-P w/Fedility 120	67
Music therapy	B4	71	67	Unknown	4	Unknown	CI24RECA	ace	92
	B5[a]	69	66	50	3	Unknown	CI512	ace	94
	B6[a]	66	56	20	10	Unknown	CI24RCA	MP3000	58
	B7	59	56	45	3	Unknown	CI24RECA	ace	94
	C1	71	65	Unknown	6	Unknown	Hires 90Khelix	HiRes-S w/Fidelity 120	44
	C2	65	52	Unknown	13	Sudden deafness	CI24RECA	Ace	68
Control group	C3[a]	74	68	69	6	Trauma	CI24RECA	Ace	89
	C4[a]	80	77	43	3	Unknown	CI24RECA	MP3000	83
	C5[b]	74	67	50	7	Unknown	Hires90khelix	HiRes-S	92
	C6	66	62	Unknown	4	Unknown	Hires90khelix	HiRes-S	75

Note. Age at CI = age at cochlear implantation; Deaf age = age at start hearing loss; CI = cochlear implant; CI exp = CI experience.
[a]Bimodal user, age at CI is shown for the first device.
[b]Bilateral CI user.
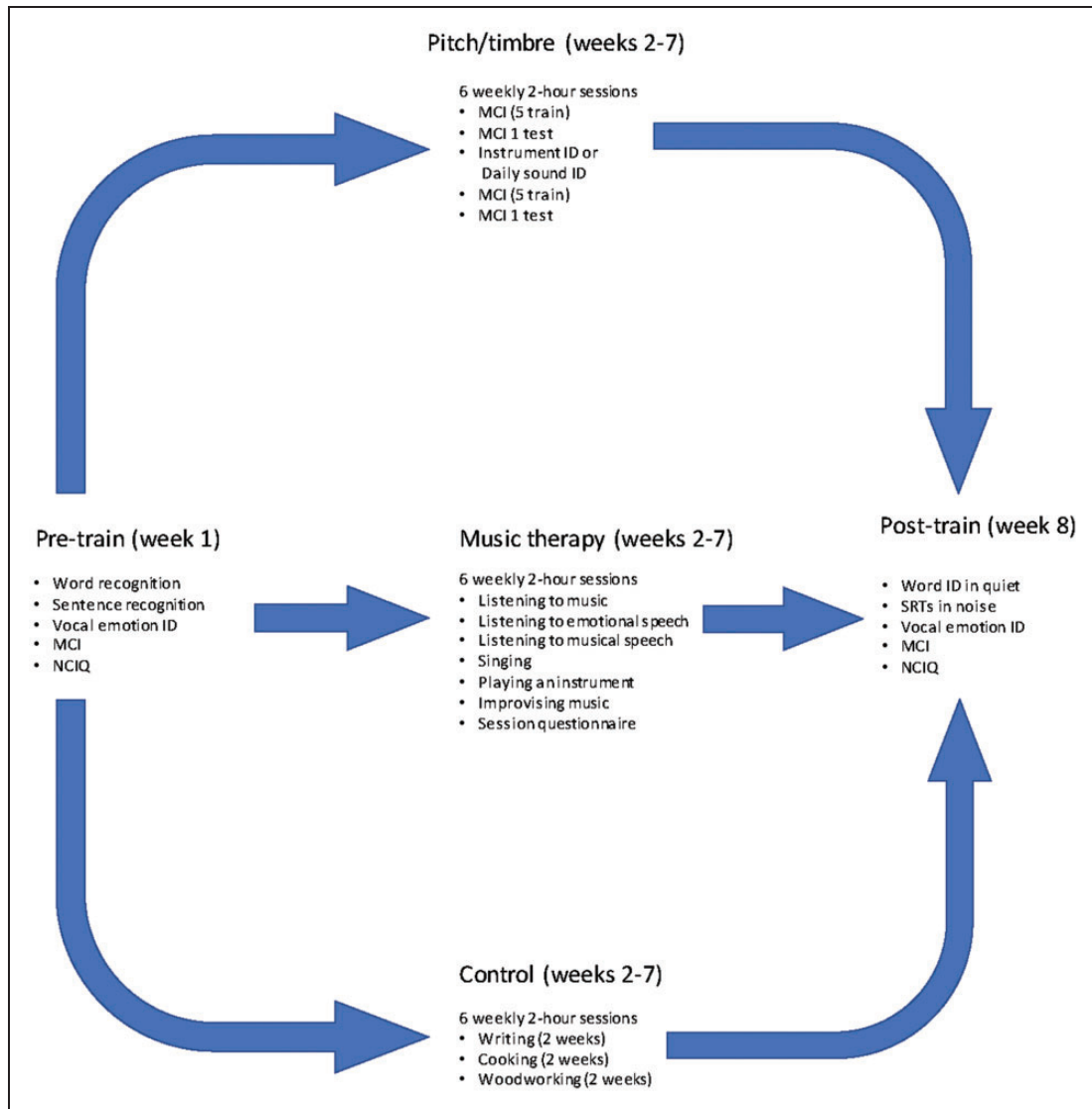
Figure 1. Flowchart of the study design.

Fuller et al. (2014) when testing NH musicians and non-musicians listening to CI simulations. All participants were tested using their clinical CI devices and daily settings; bimodal CI users removed their hearing aid during the tests. The single bilateral CI user was tested while wearing both CIs.

All speech and music tests were administered in an anechoic chamber at UMCG. Stimuli were presented at 65 dBA from a single loudspeaker (Tannoy Precision 8D; Tannoy Ltd., North Lanarkshire, UK), placed 1 m away from the participant. Sound presentation level was calibrated using a KEMAR manikin and a sound level meter (Type 2610, Brüel Kjær and Sound & Vibration Analyzer). Custom software was used to test word and sentence identification (http://tigerspeech.com/istar) and to test MCI and vocal emotion identification

(Angelsound^{TM}; Emily Shannon Fu Foundation, www.angelsound.tigerspeech.com). All stimuli were played via a Windows computer with an Asus Virtuoso Audio Device soundcard (ASUSTeK Computer Inc. Fremont, USA) connected to digital-to-analog converter (DA10; Lavry Engineering Inc., Washington, USA). Responses for the closed-set tasks were collected via touch screen monitor (A1 AOD 1908, GPEG International, Woolwich, UK). Verbal responses for open-set word and sentence identification were scored by the experimenter in an adjacent room, as well as recorded using a DR-100 digital voice recorder (Tascam, California, USA) to double-check responses as needed. Altogether, baseline (and post-training) performance measures required approximately 4 hr to complete. There was no experimenter blinding.

Word Identification. Stimuli included digital recordings of meaningful, monosyllabic Dutch words in CVC format—for example, bus (bus in English), vaak (often), nieuw (new), and so on—taken from the clinically used the nederlande vereniging voor audiologie (NVA) corpus developed by Bosman and Smoorenburg (1995). Twelve lists of 12 words each, produced by a female talker, were used for testing. Stimuli were normalized to have the same root-mean-square (RMS) amplitude (65 dBA).

Word identification was tested for four conditions: (a) quiet, (b) steady, speech-shaped noise (SSN) at 10 dB signal-to-noise ratio (SNR), (c) steady SSN at 5 dB SNR, and (d) steady SSN at 0 dB SNR. The four conditions were tested in order according to SNR. One of the 12 lists was randomly selected (without replacement) to test each condition; as such, no list was repeated within subjects. The words were presented in random order within a list. The participant was asked to repeat the word as accurately as possible, and if in doubt, to guess. The observer recorded the response and scored the phonemes correctly repeated. Stimuli were only played once; no feedback was provided.

Sentence Identification. Stimuli were meaningful and syntactically correct Dutch sentences with a semantic context, for example "De bal vloog over de schutting" (The ball flew over the fence; Plomp & Mimpen, 1979). The corpus consists of digital recordings of 10 lists of 13 sentences each (four to eight words per sentence) spoken by a female talker. Sentence identification was measured in quiet and in three types of noise: (a) steady SSN (provided with the stimulus set), (b) fluctuating SSN (provided with the set), and (c) six-talker babble (Dreschler, Verschuure, Ludvigsen, & Westermann, 2001). One list, randomly selected (without replacement), was used to test each condition; no list was repeated per participant per session.

For sentence identification in quiet, a sentence was randomly selected from the test list and presented to the participant, who was asked to repeat the sentence as accurately as possible. The observer scored the number of correctly identified words in the sentence; scores were reported in terms of percentage correct. Sentence identification in noise was measured using an adaptive one-up/one-down procedure, converging on the speech reception threshold (SRT) which was defined as the SNR that produced 50% correct whole sentence identification (Plomp & Mimpen, 1979). During testing, speech and noise were presented at the target SNR. If the participant repeated all words correctly, the SNR was reduced by 2 dB; if the participant did not repeat all words correctly, the SNR was increased by 2 dB. The initial SNR was set to +2 dB for the steady SSN condition, and to +6 dB for the fluctuating SSN and babble conditions. Note that the first sentence was repeated and

the SNR was increased until the participant repeated the entire sentence correctly. The average of the reversals in SNR between trials 4 and 13 was reported as the SRT.

Vocal Emotion Identification. Stimuli consisted of digital recordings of a nonsense word (nutohɔmsɛpikɑŋ) (Gilbers et al., 2015; Goudbeek & Broersma, 2010) produced according to four target emotions ("joy," "anger," "relief," and "sadness") by two male and two female Dutch talkers. The four target emotions were selected to represent all corners of the emotion matrix: (a) joy (high arousal, positive valence), (b) anger (high arousal, negative valence), (c) relief (low arousal, positive valence), and (d) sadness (low arousal, negative valence). Two productions of each emotion from each talker were used, for a total of 32 tokens (4 talkers × 4 emotions × 2 utterances). For further details of acoustic cues regarding the vocal emotion stimuli, see Gilbers et al. (2015).

Vocal emotion identification was measured using a four-alternative forced-choice (4AFC) closed-set task. Before formal testing, participants were first familiarized with the task using the same target emotions but produced by four other talkers that were not used during testing. During familiarization and formal testing, a stimulus was randomly selected from the set and presented to the participant, who responded by clicking on one of the four response choices shown onscreen and labeled according to target emotion. During familiarization, audiovisual feedback was provided. If the participant answered correctly, visual feedback was provided to confirm the correct response. If the participant answered incorrectly, audiovisual feedback was provided, with repeated presentation of the correct response and the participant's incorrect response. During formal testing, no feedback was provided. The software automatically calculated the percentage correct score.

Melodic Contour Identification. MCI was measured using methods and stimuli as in Galvin, Fu, and Oba (2009) and Fuller et al. (2014). Stimuli consisted of nine melodic contours with five notes each that varied in pitch pattern: "Rising," "Flat," "Falling," "Flat-Rising," "Falling-Rising," "Rising-Flat," "Falling-Flat," "Rising-Falling," and "Flat-Falling." The spacing between successive notes in the contours was 1, 2, or 3 semitones. The lowest note in a contour was A3 (220 Hz). The duration of each note in the contour was 250 ms, and silent interval between the notes was 50 ms. Contours were played by MIDI piano and organ instruments (Roland Sound Canvas GS with Microsoft Wavetable synthesis).

MCI was measured with the piano and organ alone, and in the presence of a simultaneously presented masker (flat contour played by the piano). When testing with the piano masker, the target was either the piano (same timbre) or the organ (different timbre). The base pitch

of the masker was either A3 (220 Hz; overlapping with the target pitch) or A5 (880 Hz; nonoverlapping with the target pitch). The onset and offset of the masker was the same as the target. Thus, a total of six conditions were tested: (a) piano alone, (b) organ alone, (c) piano with A3 piano masker, (d) piano with A5 piano masker, (e) organ with A3 piano masker, and (f) organ with A5 piano masker. Electrodograms for the different test stimuli can be found in Galvin, Fu, and Shannon (2009).

MCI was measured using a closed-set 9AFC task. During testing, a stimulus would be randomly selected (without replacement) and presented to the participant, who would respond by clicking on one of the nine response choices shown onscreen. During each test run of the six test conditions, each stimulus was presented twice, for a total of 54 trials (9 contours × 3 semitone spacings × 2 repeats). No feedback was provided. Scores were reported in terms of percentage correct, directly calculated by the testing software.

Health-Related QoL—Nijmegen Cochlear Implant Questionnaire. Before and after training, participants were asked to complete the Nijmegen Cochlear Implant Questionnaire (NCIQ), a validated CI-specific health-related QoL questionnaire (Hinderink, Krabbe, & Broek, 2000). The questionnaire consisted of different domains and subdomains that each included 10 statements with a 5-point response scale. The three general domains were (a) physical functioning (subdomains: sound perception basic, sound perception advanced, speech production), (b) psychological function (subdomain: self-esteem), and (c) social functioning (subdomains: activity, social interaction). The response score scale ranged between 0 and 100. The total score was calculated as the average score across the six subdomains.

Training Groups, Stimuli, and Procedures

After completing baseline measures, participants were randomly divided into three training groups: (a) Pitch/timbre (*n* = 6), (b) Music therapy (*n* = 7), and (c) Control (*n* = 6).

Pitch/Timbre Training. The pitch/timbre training group received six weekly 2-hr training sessions of computerized training for MCI (Fu et al., 2015; Galvin et al., 2007, 2012) and instrument identification. A 15-min break was provided in the middle of each training session. All training sessions were performed in a quiet room in the lab using loudspeakers (Logitech Z110) connected to a computer. All training sessions were performed using custom software (Angelsound™; Emily Shannon Fu Foundation, http://www.angelsound.tigerspeech.com/).

At the beginning of each training session, a written explanation of the exercises for that particular session

was provided. Participants were trained with each of six instruments: glockenspiel, piano, organ, clarinet, trumpet, and violin. Stimuli were MIDI instruments (Roland Sound Canvas GS with Microsoft Wavetable synthesis); examples of spectra, waveforms, and electrodograms for the different instruments can be found in Galvin, Fu, and Shannon (2009). Across training exercises, the level of difficulty was increased by reducing the spacing between notes in the contours from six semitones to one semitone. During training, a contour would be presented and the participant responded by clicking on one of the nine response choices shown onscreen. If the participant responded correctly, a new contour would be presented. If the participant answered incorrectly, audiovisual feedback was provided in which the correct answer and the participant's response were repeatedly played for comparison, after which a new contour was presented. MCI was also retested (without feedback) after completing five training exercises.

In each training session, participants were also trained for instrument identification and daily-life sound identification. These additional training exercises were included to diversify the training and to keep participants engaged during training. For the instrument identification training, stimuli consisted of melodic contours played by one of the six instruments used in the MCI training. For the daily-life sound identification training, stimuli consisted of sounds commonly encountered in everyday life (e.g., baby crying, cat meowing, car honking, water running, etc.). During the instrument identification or daily-life sound training, a stimulus would be presented and the participant would click on one of the response choices (six choices for instrument identification training, two to six choices for the daily-life sound training) shown onscreen. If the participant responded correctly, a new stimulus would be presented. If the participant answered incorrectly, audiovisual feedback was provided in which the correct answer and the participant's response were repeatedly played for comparison, after which a new stimulus was presented.

Music Therapy. Music therapy training consisted of six 2-hr group sessions, with a 15-min break in each session. The music therapy sessions were organized under the supervision of three music therapy students and their supervisor from the Hogeschool Utrecht, Department of Creative Therapy, Amersfoort. All sessions were held in the activity room of the rehabilitation center of the CI team of the Northern Netherlands and participants were accompanied by the music therapy students and one member of the CI team.

The music therapy training was social and dynamic, and consisted of auditory training (listening to speech and music) and auditory-motor training (playing an instrument, singing). Multimodal training

(auditory + motor) has been suggested to enhance neuroplasticity (Herholz & Zatorre, 2012). The music therapy targeted more central cognitive processing than the computerized pitch/timbre training described earlier. Music therapy included six types of therapy and training exercises: (a) music perception, (b) musical speech perception, (c) emotional speech perception, (d) singing, (e) playing an instrument, and (f) improvising music. Detailed description of the framework of the music therapy and the different tasks can be found in online Appendix. The music therapy was interactive, and the interactions between therapists and clients and feedback for each session guided the interactions for the following session (Migchelbrink & Brinkman, 2000).

At the end of each training session, participants also completed a (nonvalidated) questionnaire to obtain feedback on the session and to track self-reported progression. The questionnaire consisted of four domains (perception of rhythm, perception of musical speech, music perception, and playing music). Participants were asked to rate their abilities in each domain using a number ranging from 1 (*poorest ability*) to 10 (*highest ability*).

Control Group. Training for the control group consisted of six 2-hr group sessions, with a 15-min break in each session. The control group participated in interactive training activities (writing, cooking, and woodworking) that did not include music. The control group experienced a similar dynamic, interactive training environment as the music therapy group, where they had to actively listen to instructions and work with each other, but without explicit music training. Thus, comparing training outcomes between the music therapy and control groups would provide insight regarding training benefits because of music training or to the social interaction.

All training sessions for the control group were conducted at the School for the Deaf and were supervised by a member of the CI team and a social worker who explained the tasks and answered any questions. The first two training sessions involved a writing course conducted by a professional writing coach. The third and fourth training sessions involved a cooking course during which the participants collaborated to prepare different dishes. The fifth and sixth sessions involved woodworking (building a birdhouse) under the supervision of a woodworking teacher.

Results

Word Identification

Figure 2 shows boxplots of word identification scores before and after training for the three participant groups. Performance generally worsened as the SNR

was reduced, with no clear differences before or after training or among participant groups. Table 2 shows the mean, minimum, and maximum change in performance after training. A split-plot repeated measures analysis of variance (RM ANOVA) was performed on the data shown in Figure 2, with training (pre, post) and SNR (quiet, 10 dB, 5 dB, 0 dB) as within-subjects factors and training group (pitch/timbre, music therapy, control) as between-subjects factors; Greenhouse–Geisser correction was applied. Results showed a significant effect for SNR, $F(2,28) = 88.5$, $p < .001$], but not for training, $F(1,14) = 0.2$, $p = .704$, or training group, $F(2,14) = 0.8$, $p = .487$. There was a significant interaction only between training and SNR, $F(2.6,5.1) = 5.0$, $p = .005$.

Sentence Identification

Figure 3 shows boxplots of sentence identification in quiet and SRTs in noise before and after training for the three participant groups. Performance was generally poorer with the fluctuating SSN and babble than with the steady SSN, with no clear differences among participant groups and no clear training effects. Table 3 shows the mean, minimum, and maximum change in performance after training. A split-plot RM ANOVA was performed on the sentence identification in quiet data, with training as the within-subjects factor and training group as the between-subjects factor; Greenhouse–Geisser correction was applied. Results showed no significant effects for training, $F(1,16) = 1.0$, $p = .339$] or training group, $F(2,16) = 1.2$, $p = .328$; there were no significant interactions $F(2,16) = 1.3$, $p = .307$. A split-plot RM ANOVA was also performed on the sentence identification in noise data, with training and noise type (steady SSN, fluctuating SSN, babble) as the within-subjects factor and training group as the between-subjects factor; Greenhouse–Geisser correction was applied. Results showed a significant effect for noise type, $F(1.7,30) = 82.4$, $p < .005$, but not for training, $F(1,16) = 0.1$, $p = .979$, or training group, $F(2,16) = 0.2$, $p = .817$; there were no significant interactions ($p < .05$ in all cases).

Vocal Emotion Identification

Figure 4 shows boxplots for vocal emotion identification scores before and after training for the three participant groups. There was a substantial improvement in performance for the music therapy group. Table 4 shows the mean, minimum, and maximum change in performance after training. A split-plot RM ANOVA was performed on the data in Figure 4, with training as the within-subjects factor and training group as the between-subjects factor; Greenhouse–Geisser correction
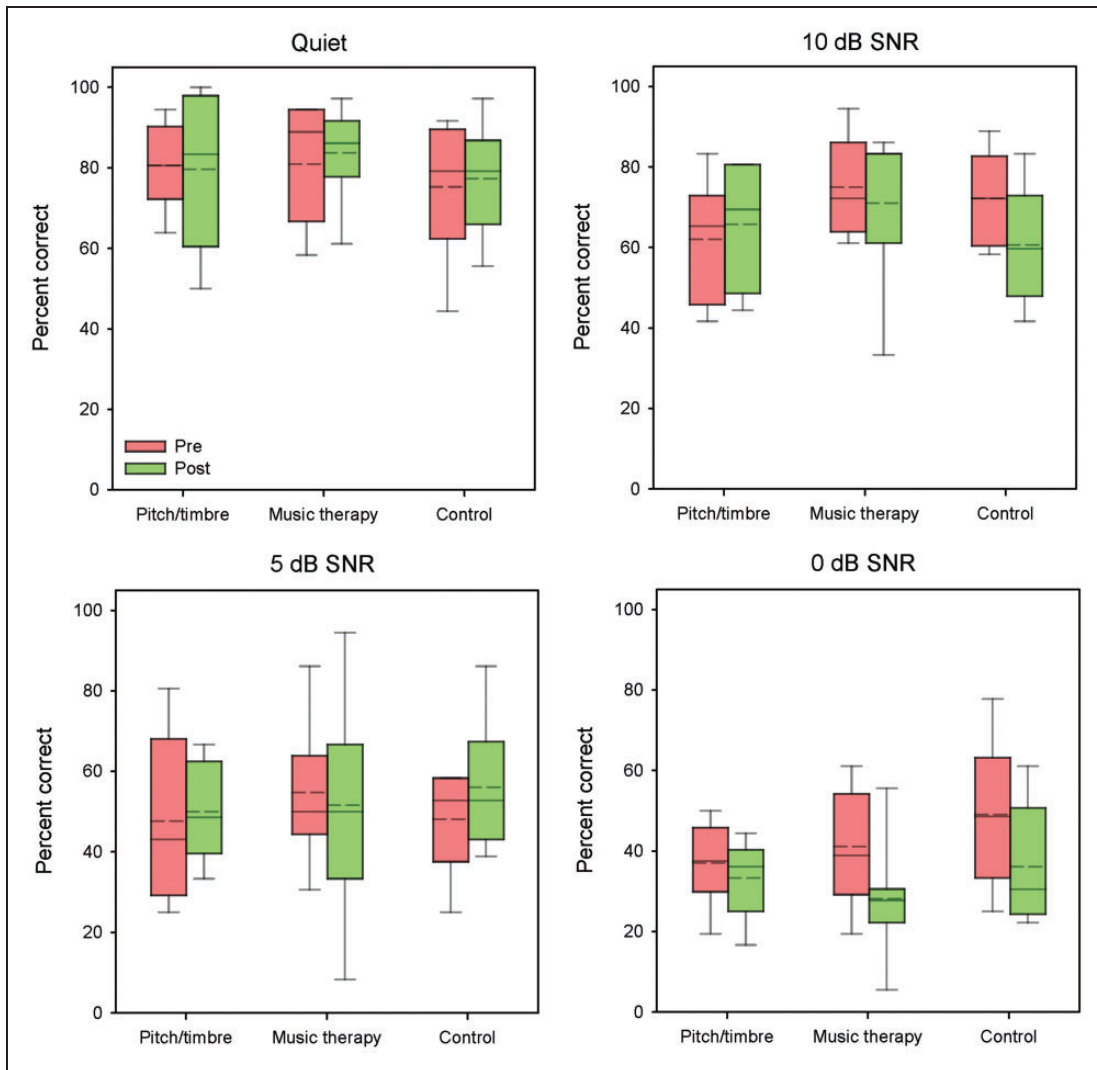
Figure 2. Boxplots of word identification scores in quiet and in noise before and after training, for the three participant groups. The boxes show the 25th and 75th percentiles, the error bars show the 5th and 95th percentiles, the solid line shows the median, and the dashed line shows the mean.
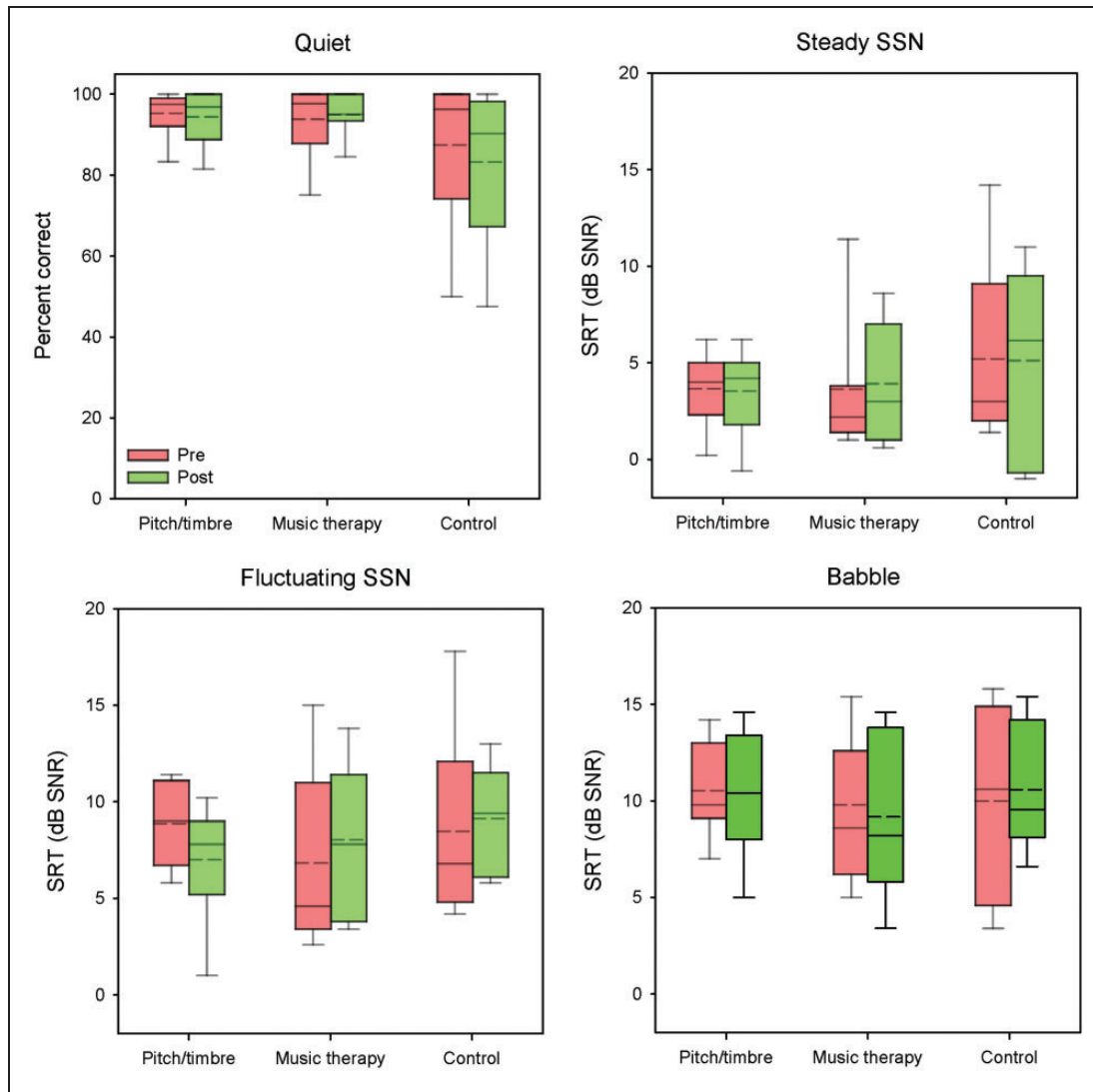
Table 2. Mean, Minimum, and Maximum Change in Word Identification Performance After Training (Posttrain–Pretrain), in Percentage Points.

	Quiet			10 dB SNR			5 dB SNR			0 dB SNR		
Group	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Pitch/timbre	−0.9	−19.4	13.9	3.7	−2.8	16.7	2.3	−19.5	27.8	−3.7	−22.2	11.1
Music therapy	2.8	−8.3	27.8	−4.0	−33.3	25.0	−3.2	−36.1	22.3	−1.2	−19.4	30.6
Control	2.0	−13.9	14.9	−11.6	−36.1	2.8	7.9	−13.9	27.8	−13.0	−36.1	0.0

Note. Positive values indicate a training benefit. SNR = signal-to-noise ratio.

was applied. Results showed no significant effects for training, $F(1,16) = 3.9$, $p = .067$, or training group, $F(2,16) = 2.1$, $p = .159$; there were no significant interactions, $F(2,16) = 1.5$, $p = .263$. A one-way RM ANOVA was also performed on the data for the music therapy group, with training as the within-subjects factor. Results showed a significant effect for training, $F(1,6) = 9.3$, $p = .022$.

Figure 3. Boxplots of sentence identification scores in quiet and SRTs in different types of noise before and after training, for the three participant groups. The boxes show the 25th and 75th percentiles, the error bars show the 5th and 95th percentiles, the solid line shows the median, and the dashed line shows the mean.

Table 3. Mean, Minimum, and Maximum Change in Sentence Identification Performance After Training (Posttrain–Pretrain).

Group	Quiet (percentage points)			Steady SSN (dB)			Fluctuating SSN (dB)			Babble (dB)		
	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Pitch/timbre	−0.9	−3.8	1.3	−0.1	1.6	−2.0	−1.9	0.0	−4.8	−0.1	0.8	−2.0
Music therapy	1.1	−5.3	19.9	0.3	3.2	−2.8	1.2	4.4	−3.6	−0.6	1.2	−2.8
Control	−4.2	−11.8	0.0	−0.1	5.5	−5.2	0.7	5.2	−4.8	0.6	5.5	−5.6

Note. For sentence identification in quiet, positive values indicate a training benefit. For sentence identification in noise, negative values indicate a training benefit. SNR = signal-to-noise ratio.

Melodic Contour Identification

Figure 5 shows boxplots of MCI scores for the piano and organ targets before and after training for the three participant groups. For the pitch/timbre training group,

mean scores were generally better after the MCI training. Table 5 shows the mean, minimum, and maximum change in performance after training. A split-plot RM ANOVA with training, target (piano, organ), and masker (no
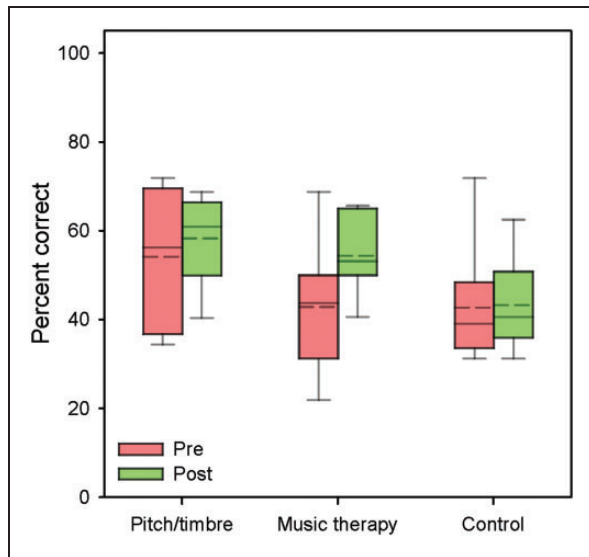
Figure 4. Boxplots of vocal emotion identification scores before and after training for the three participant groups. The boxes show the 25th and 75th percentiles, the error bars show the 5th and 95th percentiles, the solid line shows the median, and the dashed line shows the mean.

Table 4. Mean, Minimum, and Maximum Change in Vocal Emotion Identification After Training (Posttrain–Pre-train), in Percentage Points.

Group	Mean	Min.	Max.
Pitch/timbre	−0.9	−3.8	1.3
Music therapy	1.1	−5.3	19.9
Control	−4.2	−11.8	0.0

Note. Positive values indicate a training benefit.

masker, A3, A5) as the within-group factor was performed on the data in Figure 5; Greenhouse–Geisser correction was applied. Results showed a significant effect for masker, $F(2,30)=17.9$, $p<.001$, but not for training, $F(1,15)=1.9$, $p=.192$, target, $F(1,15)=2.3$, $p=.148$, or training group, $F(2,15)=2.9$, $p=.083$. Significant interactions were observed between training and training group, $F(2,15)=5.9$, $p=.013$, and among training, masker, target, and training group, $F(3.4,25.7)=3.0$, $p=.041$. Because of the substantial training effect for the pitch/timbre group, two-way RM ANOVAs were performed for the piano target and organ target data for the pitch/timbre group, with training and masker as the within-subjects factors. For the piano target, results showed a significant effect for training, $F(1,10)=7.0$, $p=.045$, but not for masker, $F(2,10)=2.3$, $p=.149$]; there was no significant interaction, $F(2,10)=0.5$, $p=.630$. For the organ target, results showed a significant effect for masker, $F(2,10)=4.1$, $p=.049$, but not for

training, $F(1,10)=5.6$, $p=.064$; there was a significant interaction, $F(2,10)=14.2$, $p=.001$. Post hoc Bonferroni pairwise comparisons showed significant effects for training for the no masker and A5 masker conditions ($p<.05$ in both cases), and that post-training performance was significantly better for the no masker than the A3 masker condition ($p<.05$).

## Quality of Life

Figure 6 shows boxplots for the total NCIQ scores (averaged across the six subdomains) before and after training for the three participant groups. Table 6 shows the mean, minimum, and maximum change in performance after training for total NCIQ scores. A split-plot ANOVA with training as the within-subjects factor and training group as the between-subjects factor was performed on the data shown in Figure 6; Greenhouse–Geisser correction was applied. Results showed no significant effect for training, $F(1,16)<0.1$, $p=.928$, or training group, $F(2,16)=0.3$, $p=.747$; there were no significant interactions, $F(2,16)=0.8$, $p=.454$.

Figure 7 shows boxplots for NCIQ scores for each subdomain before and after training for the three participant groups. Table 6 shows the mean, minimum, and maximum change in performance after training for each subdomain. A split-plot ANOVA with training and subdomain (sound perception basic, sound perception advanced, speech production, self-esteem, activity limitations, social interactions) as the within-subjects factors and training group as the between-subjects factor was performed on the data shown in Figure 7. Greenhouse–Geisser correction was applied. Results showed a significant effect for subdomain, $F(3.9,7.8)=22.5$, $p<.001$, but not for training, $F(1,16)<0.1$, $p=.927$, or training group, $F(2,16)=0.3$, $p=.747$; there were no significant interactions ($p<.05$ in all cases).

Subjective Survey Music Therapy Group

Figure 8 shows boxplot of ratings for different survey questions completed by members of the music therapy group at the end of each training session. A two-way RM ANOVA was performed on the data shown in Figure 8, with training session (1, 2, 3, 4, 5, and 6) and survey question (rhythm perception, musical speech perception, music perception, and playing music) as within-subjects factors. Results showed significant effects for training session, $F(5,75)=9.6$, $p<.001$, and survey question, $F(3,75)=7.3$, $p=.003$; there were no significant interactions, $F(15,75)=1.4$, $p=.190$. Post hoc Bonferroni pairwise comparisons showed that ratings were significantly higher for Sessions 3 to 6 relative to Session 1 ($p<.05$ in all cases) and significantly higher for Sessions 5 and 6 relative to Session 2 ($p<.05$ in both
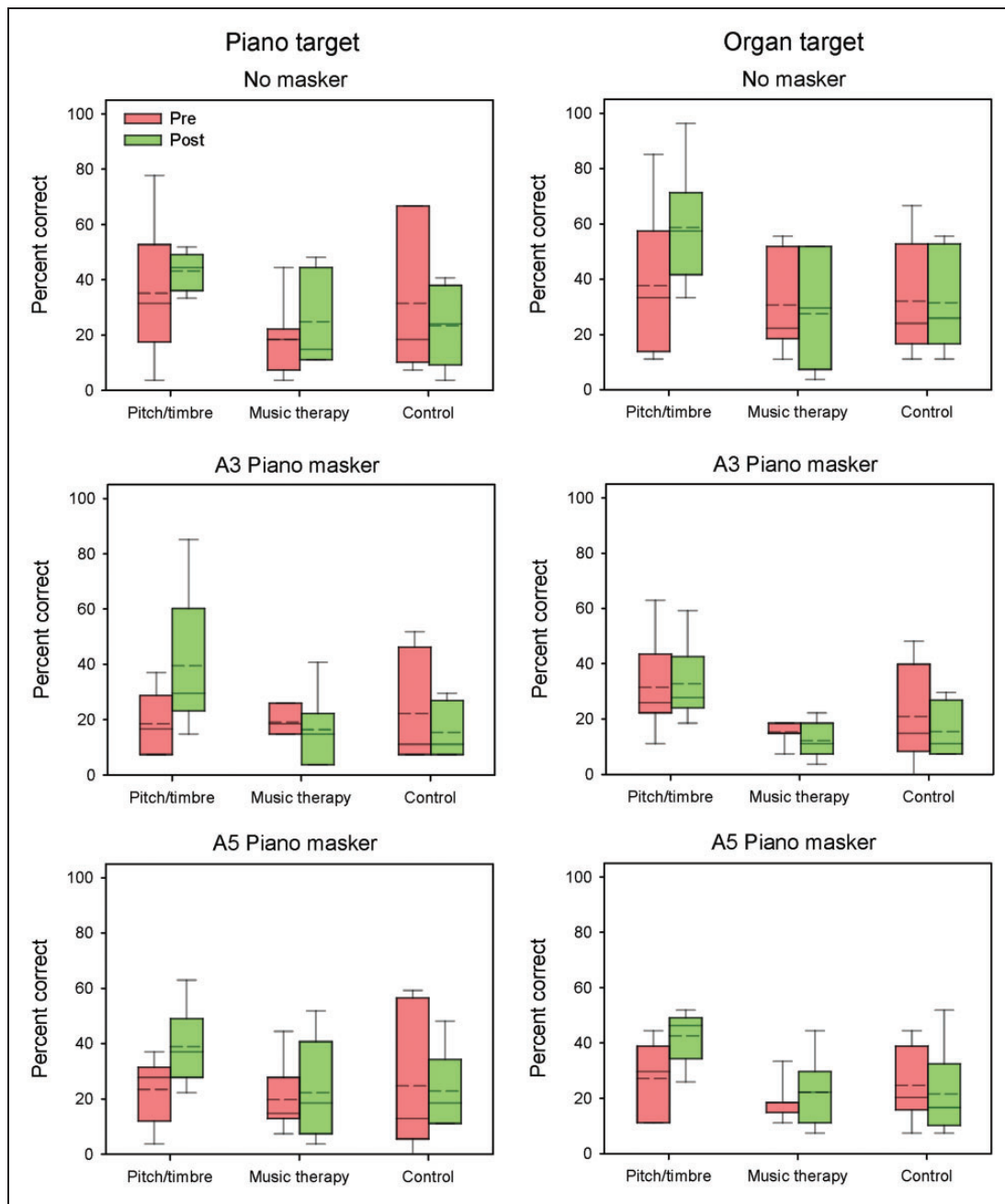
Figure 5. Boxplots of MCI scores with the piano (left column) and organ targets for the no masker (top row), overlapping A3 piano masker (middle row), and nonoverlapping A5 piano masker (bottom row) before and after training for the three participant groups. The boxes show the 25th and 75th percentiles, the error bars show the 5th and 95th percentiles, the solid line shows the median, and the dashed line shows the mean.

cases). Music perception and playing music were rated significantly better than musical speech perception ($p < .05$ in both cases).

Discussion

The main research questions of this study were: (a) Can pitch/timbre training or group music therapy improve CI users' perception of music (within-domain effect) and/or speech (cross-domain effect)? (b) Which training method is most effective for CI users? Behavioral data showed a significant within-domain effect (improved MCI performance) only for the pitch/timbre training group and a small cross-domain effect (improved vocal emotion identification) only for the music therapy group. Word or sentence identification in quiet or in noise did not

Table 5. Mean, Minimum, and Maximum Change in MCI Performance After Training (Posttrain–Pretrain), in Percentage Points.

		No masker			A3 piano masker			A5 piano masker		
	Group	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Piano target	Pitch/timbre	8.0	−29.6	44.5	21.0	−11.1	63.0	15.4	−7.4	37.0
	Music therapy	6.3	−3.7	22.2	−2.6	−14.8	14.8	5.3	−11.1	40.7
	Control	−8.0	−29.6	7.4	−6.8	−25.9	3.7	−1.9	−29.6	11.1
Organ target	Pitch/timbre	21.0	7.4	51.9	1.2	−11.1	14.8	15.4	3.7	37.1
	Music therapy	−3.2	−11.1	7.4	−3.2	−14.8	3.7	3.7	−11.1	11.1
	Control	−0.6	−11.1	7.4	−5.6	−18.5	11.1	−3.1	−11.1	7.4
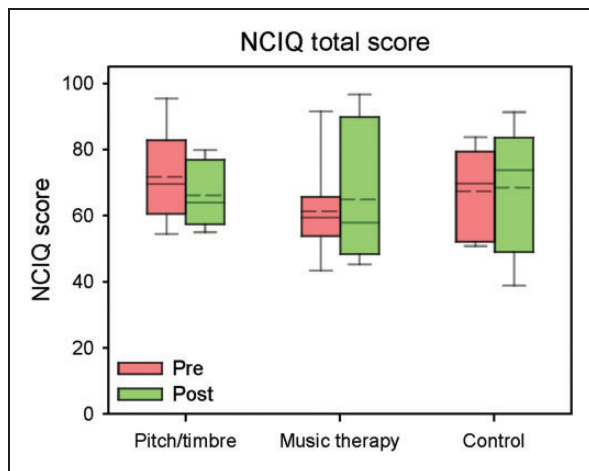
Figure 6. Boxplots of total NCIQ scores (averaged across all subdomains) before and after training for the three participant groups. The boxes show the 25th and 75th percentiles, the error bars show the 5th and 95th percentiles, the solid line shows the median, and the dashed line shows the mean.

significantly improve with training for any of the three participant groups. Other than the improved MCI performance for the pitch/training group, there were no significant differences across training methods. The subjective NCIQ showed no significant effect of training, in line with the generally weak training benefits observed for the behavioral measures. For the music therapy group, self-reported perception appeared to improve across training sessions. Subsequently, we discuss the results in greater detail.

Within-Domain Effects

MCI training in the pitch/timbre group significantly improved MCI performance, consistent with previous studies (Galvin et al., 2007, 2012). Training benefits were observed for the piano target and, to a greater extent, for the organ target when there was no masker. The greater improvement for the organ is in line with

Galvin, Fu, and Oba (2008), who reported that mean MCI performance in CI users was poorest with piano and best with organ. Perhaps, the organ is more easily trained in CI users because its spectral-temporal content is less complex than other instruments such as the piano (see Figure 4 in Galvin, Fu, & Shannon, 2009). Note that while participants trained with six instruments without a masker, performance also improved for the masker conditions. Lo et al. (2015) showed that the largest improvements in MCI performance occurred during the first 2 weeks of training (possibly indicating task-related learning), with the maximum overall improvement observed after 4 to 6 weeks of training. Unfortunately, because MCI performance was not tracked across training sessions for the pitch/timbre group, the rate of learning is unknown. In future studies, it would be worthwhile to extend the duration of training and to test performance during the training to better observe the rate of improvement and where the training effect saturates.

The music therapy and control groups showed no improvement in the MCI performance. Note that these groups interacted with the MCI test and stimuli only twice (before and after training), while the pitch/timbre group received 6 hr of MCI training. As such, there was a greater possibility for task-specific learning for the pitch/timbre group. In future studies that compare training methods, repeatedly testing baseline until achieving asymptotic performance may reduce the possibilities of procedural learning effects. It is also possible that extending/intensifying the training for the music therapy group might also improve melodic pitch perception.

Cross-Domain Effects

Word and Sentence Identification. In all three training groups, no transfer of learning to speech perception (words or sentences) was observed. This finding is not in agreement with the preliminary findings of Patel (2014) and Lo et al. (2015), but is in line with the outcomes from Petersen et al. (2012). Note that the small cross-domain effects were observed only for two

Table 6. Mean, Minimum, and Maximum Change in NCIQ Scores for Each of the Subdomains and For the Total NCIQ Score After Training (Posttrain–Pretrain).

Group	Sound perception basic			Sound perception advanced			Speech production		
	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Pitch/timbre	−5.8	−47.5	32.5	3.7	−37.5	27.5	−2.1	−17.5	15.0
Music therapy	−2.5	−22.5	32.5	11.1	2.5	40.0	−2.5	−17.3	17.5
Control	1.7	−10.0	15.0	1.0	−7.5	11.1	7.9	0.0	12.5

Group	Self−esteem			Activity limitations			Social interaction		
	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Pitch/timbre	−7.5	−40.0	11.1	−8.7	−55.0	8.1	−12.9	−55.0	6.1
Music therapy	−1.0	−21.1	17.5	8.4	−8.5	37.8	8.1	−12.8	40.0
Control	−2.0	−12.5	10.0	−3.6	−29.4	6.9	1.6	−20.0	15.0

Group	Total NCIQ score		
	Mean	Min.	Max.
Pitch/timbre	−5.5	−40.4	8.0
Music therapy	3.6	−6.6	30.9
Control	1.1	−12.0	7.6

Note. Positive values indicate a training benefit. NCIQ = Nijmegen Cochlear Implant Questionnaire.

participants in Patel (2014). Lo et al. (2015) showed a positive effect of musical training on prosody perception (question vs. statement) and consonant discrimination in 16 CI users. Petersen et al. showed no effect of musical training program on speech understanding in noise in 18 CI users.

In NH listeners, musician advantages for speech understanding in noise have been generally weak or inconsistent (Fuller et al., 2014; Parbery-Clark, Strait, Anderson, Hittner, & Kraus, 2011; Parbery-Clark et al., 2009; Ruggles et al., 2014; Zendel & Alain, 2012). Presumably, musicians have better pitch perception that allows for better segregation of speech and maskers. Alternatively, musician effects for segregation may be based on other acoustic cues besides voice pitch, and music training may improve working memory and overall pattern perception, which in turn may improve segregation and spatial hearing abilities (Başkent & Gaudrain, 2016; Clayton et al., 2016). It should be noted that the 6 hr of training used in the study is not comparable with the years of training experienced by musicians. It should also be noted that CI users experience auditory deprivation and greatly reduced spectro-temporal resolution, which is not experienced by NH musicians. CI users are also much more heterogeneous as a group than NH listeners. There is great variability in CI performance for a variety of outcome measures, because of device- and patient-related factors (e.g., electrode–neural interface, duration of deafness, age at implantation, CI experience, etc.). With these issues in

mind, cross-domain benefits for music training may be hard-won in CI users.

Another explanation for the present lack of strong cross-domain effects may be because of the speech listening tasks (i.e., word and sentence identification in quiet and in noise). Previous studies have shown greater benefits for music training for perception of pitch-mediated speech (Fuller et al., 2014; Patel, 2014). Music training has also been shown to benefit perception of speech with low linguistic content (consonant identification in Lo et al., 2015; syllable perception in Zuk et al., 2013). Word and sentence identification, as used in this study, are rich in linguistic content and as such, do not as strongly depend on perception of voice pitch. In future music training studies, it may be interesting to include speech outcome measures that differ in terms of linguistic content or importance of voice pitch cues.

Vocal Emotion Identification. Vocal emotion identification was improved only in the music therapy group. Unlike the pitch/timbre and control groups, the music therapy group received specific training for emotion identification. In one exercise, one member of the group was asked to select an emotion from a list written on a chalkboard and play this emotion on an instrument; the other group members were then asked to identify the emotion. In another exercise, a song or story with emotional content was sung or spoken by a session leader, and group members were asked to identify the emotion. These training exercises might have contributed to the positive
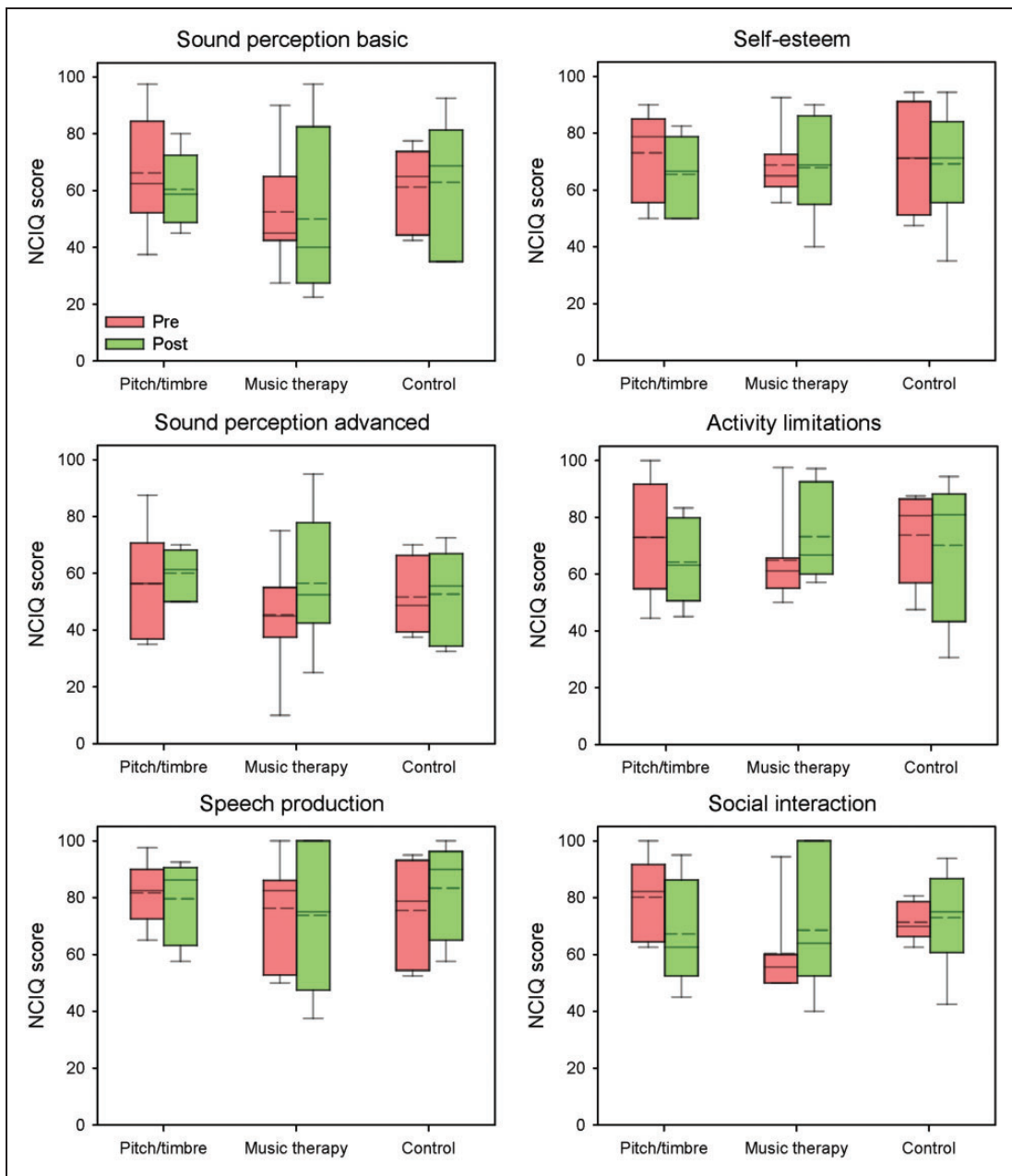
Figure 7. Boxplots of NCIQ scores for each subdomain before and after training, for the three participant groups. The boxes show the 25th and 75th percentiles, the error bars show the 5th and 95th percentiles, the solid line shows the median, and the dashed line shows the mean.

effect of training on emotion identification in the music therapy group.

Another factor that may have contributed to better emotional identification in the music therapy group is the dynamic nature of the training, which combined listening, singing, and playing an instrument in a social context, similar to the training methods used by Petersen et al. (2012). Such an approach may target more global cognitive changes, in contrast to the more bottom-up MCI training in the pitch/timbre

group (note that Peterson et al. also included MCI training as part of their music therapy). Petersen et al. found that musically trained CI users were more quickly able to detect emotional prosody in meaningful sentences and words than were the CI users who received no music training (control group). However, after 6 months, detection of emotional prosody was not significantly different between the music training and control groups. Note that Petersen et al. worked with newly implanted CI users, who generally
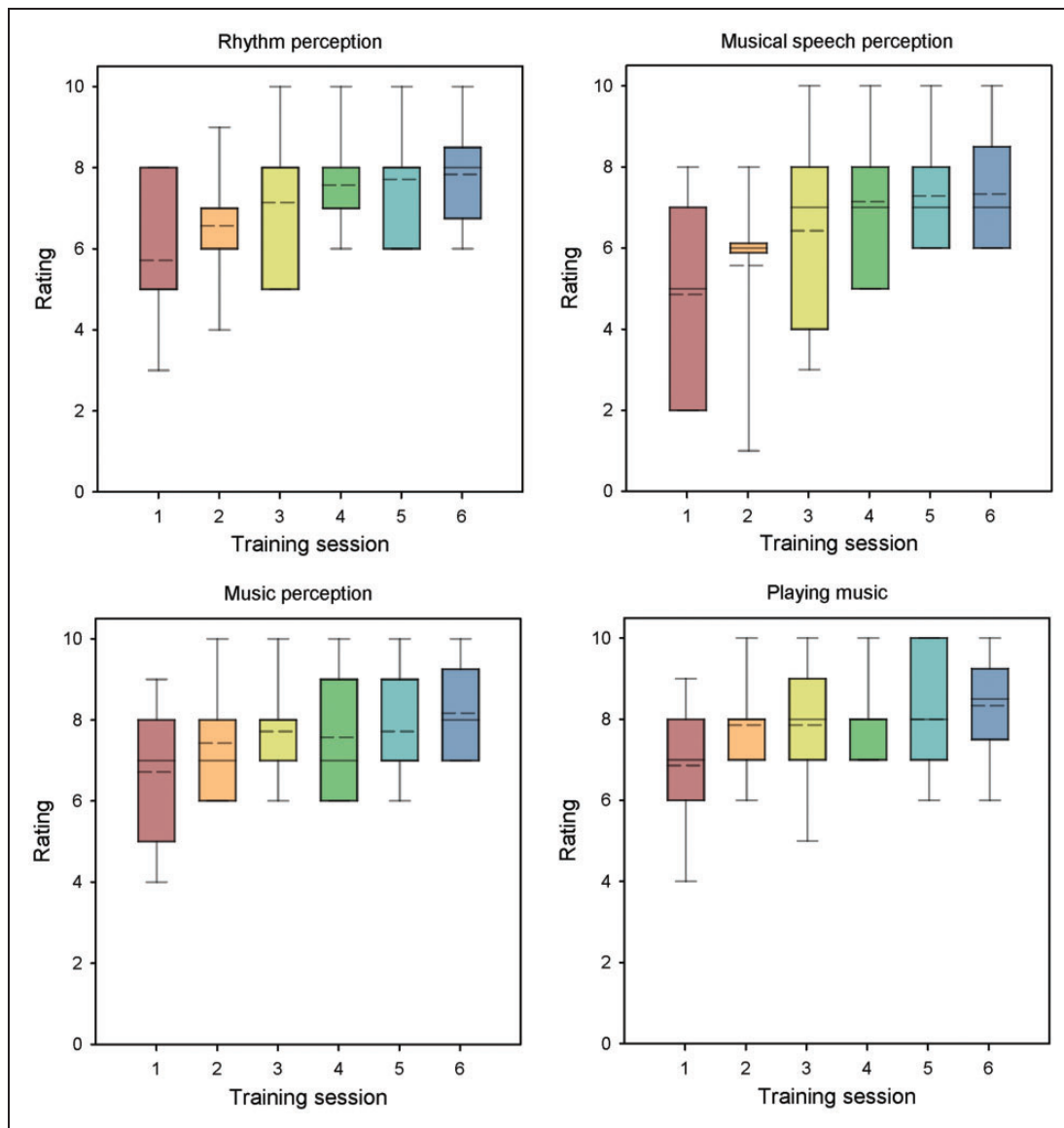
Figure 8. Boxplots of survey scores collected in the music therapy group at the end of each training session. The boxes show the 25th and 75th percentiles, the error bars show the 5th and 95th percentiles, the solid line shows the median, and the dashed line shows the mean.

experience the greatest adaptation to electric hearing during the first 6 months of implant use.

As noted earlier, musical training may especially benefit speech perception tasks that depend strongly on perception and processing of voice pitch cues (Banse & Sherer, 1996; Başkent & Gaudrain, 2016). While pitch cues strongly contribute to emotion identification, other acoustic cues that covary with F0 also contribute, such as duration (longer for sad, shorter for happy), overall amplitude (higher for happy, lower for sad), and tempo and pausing (Hubbard & Assmann, 2013; Luo, Fu, Wu, & Hsu, 2009). Vocal emotion identification has been shown to be poorer in CI users than NH listeners (House, 1994; Jiam, Caldwell, Deroche,

Chatterjee, & Limb, 2017; Luo, Fu, & Galvin, 2007; Pereira, 2000). Gilbers et al. (2015) suggested that NH listeners attend to mean pitch for emotion identification, whether listening to unprocessed stimuli or to CI simulations, while CI users seem to attend to the pitch ranges conveyed by the temporal modulations. Fuller et al. (2014) found a significant musician advantage for emotion identification for NH participants listening to acoustic CI stimulations. Thus, even with spectro-temporal degradation similar to that in real CI users, long-term musical training appeared to benefit emotion identification. As such, music training in CI users may also improve emotion identification in CI users, as occurred within the present music therapy group.

Subjective Measures

In all three training groups, no effect of training was observed on NCIQ scores, in contrast to the positive effects previously shown in other patient groups (Hilliard, 2003; Walworth et al., 2008). Note that the population of this study (CI users) was quite different from those in previous studies (terminally ill patients in Hilliard, 2003; elective brain surgery patients in Walworth et al., 2008). It may be that the short period of training in this study was not sufficient to affect QoL, as QoL is complex and multidimensional (Donnelly & Walsh, 1996). Hilliard (2003) and Walworth et al. (2008) did not report the time frame of the music therapy. It is also possible that the health-related, disease-specific questionnaire (NCIQ) used in our study did not capture the changes in QoL that may have been affected by the training. A more specific questionnaire that focuses on aspects of QoL that may be expected to improve with music training might better capture such effects.

In the music therapy group only, a survey was administered at the end of each training session to capture any subjective changes in terms of rhythm perception, musical speech perception, music perception, and playing music. The surveys were conducted in the music therapy group to guide the interactions for the subsequent sessions, as in Migchelbrink and Brinkman (2000). Results showed that subjective ratings improved across sessions for all domains (Figure 8). Anecdotal reports suggested that the music therapy participants felt better about their perceptual skills. They reported that they better understood other talkers' emotions, listened to music more often, and enjoyed music more. Participants were enthusiastic about the music therapy, similar to CI participants in the Hütter et al. (2015) and Petersen et al. (2012) studies. These self-reports of improved speech and music perception are encouraging and should be more deeply investigated, as the OPERA hypothesis (Patel, 2011, 2012, 2014) states that emotion and attention are factors in music activities that elicit higher benefits from training. Indeed, feeling positive about the training experience may motivate CI users to continue to train and better benefit from the training. Unfortunately, no subjective ratings were obtained in the pitch/timbre or control groups, mainly because of limited testing times, preventing us from a more direct comparison on such aspects between different groups.

Training Methods

In this study, we used training approaches that differed in terms of the amount of social interaction, as well as the type of training (targeting more bottom-up vs. higher cognitive processing). For most measures, there were no significant differences between the training methods. As discussed earlier, this may be because some outcome measures (e.g., word and sentence identification) may not have been sufficiently sensitive to perceptual abilities that might have been improved by particular training methods (e.g., improved voice pitch perception).

Computer-based musical training as was used in this study has been shown to be an effective within-domain training method in CI users earlier (Fu et al., 2015; Galvin et al., 2007, 2012; Patel, 2014). Our findings add to this literature, showing the effectiveness of bottom-up training for a specific task. While there were no significant cross-domain effects for the pitch/timbre group, some participants experienced substantial gains in speech performance after the MCI training (see maximum change in performance in Tables 2–4). The control group generally did not exhibit such gains in speech performance, possibly indicating an advantage of targeted computerized training over untargeted training. Computerized training indeed may present a number of advantages. It allows for repeated training sessions using large numbers of trials and feedback in a simple setting with minimal supervision. More importantly, such training can be targeted to improve specific perceptual abilities (e.g., monosyllable word training improved phoneme identification in Fu & Galvin, 2007, 2008; MCI training improved melodic pitch perception in Galvin et al., 2007, 2012). Such training can also be easily modified to accommodate different levels of performance by adjusting the level of difficulty (e.g., varying the semitone spacing for the MCI training). Finally, such training provides accessible, low-cost rehabilitation in CI users.

Because of its multimodal, dynamic, and social nature, music therapy, while still targeted, may be a more engaging approach toward auditory rehabilitation. The music therapy focused on real-life stimuli in a group of CI users. The exercises differed in difficulty and direct feedback was provided. The music therapy was considered to target more top-down processing, as participants had to produce and listen to emotional speech and real music (as opposed to the melodic contours in the pitch/training group). Using more complex stimuli has been shown to lead to greater perceptual enhancement in NH listeners using CI simulations (Loebach & Pisoni, 2008). The music therapy group was the only training group to exhibit improved speech-related task performance, though it is unclear whether the vocal emotion training directly contributed to the improved emotion identification. While there was no significant improvement on average in word or sentence identification, individual data indicated that some participants experienced substantial post-training gains (see the maximum change in performance in Tables 2 and 3). Subjectively, the music therapy group also reported improved music perception skills, as well as enthusiastic overall reactions to the sessions. Such enthusiasm can elicit positive emotions, and thereby enhance attention (Gfeller et al.,

2015; Herholz & Zatorre, 2012; Patel, 2011, 2012, 2014) and motivation to continue with training. Note, however, that the music therapy did not translate to better MCI performance. This suggests that it may be important to direct attention to key cues (i.e., some bottom-up training component) to maximize the benefit of music therapy. The pitch/timbre and music therapy training differed in terms of social interaction; the music therapy and control groups both involved social interaction, but differed in terms of training exercises. The control group did not exhibit any significant improvements for any of the outcome measures, suggesting that social interaction alone was not sufficient to show an improvement in the behavioral or subjective measures of this study.

Note that there was no experimental blinding of the study groups. Given that the vocal emotion and MCI tasks were closed-set and that the participant entered responses directly within software, blinding was likely not a major issue in this study. Future studies may include blinding participation in training or control groups to avoid experimenter bias.

The duration of this study was short (one 2-hr session per week for 6 weeks) and the total amount of training or therapy provided was small (~12 hr). Such a schedule may not be optimal for training, but was designed to resemble a rehabilitation program that might be feasible in rehabilitation clinics. While some studies have shown that training is most effective when it consists of short training sessions over a longer period of time (e.g., Gfeller et al., 2015), we chose to set up a shorter training period with long sessions. Training benefits have been observed in previous CI studies that differed in terms of the total time of training (5 days to 6 months; Driscoll, 2012; Fu & Galvin, 2007; Galvin et al., 2007; Lo, McMahon, Looi, & Thompson, 2015; Petersen et al., 2012), as well as the frequency and duration of training sessions (e.g., 1 hr/week for 6 months, 15 min/day for 4 days/week for 6 weeks, 3 hr/day over a 5-day period; Galvin et al., 2007; Lo et al., 2015; Petersen et al., 2012). More intensive, frequent, but shorter training sessions over a longer period of time may yield even greater benefits for CI users' music and speech perception.

Conclusions

In this study, outcomes for two types of music training (pitch/timbre, music therapy) were compared, along with a control group that received no music training. The training approaches differed in terms of targeting more bottom-up or top-down processes, and in terms of social interaction. There was a significant within-domain effect of music training only for the pitch/timbre group. There was a significant cross-domain effect (better vocal emotion identification) only for the music therapy group. There was no significant benefit of training for any

outcome measure for the control group. The present results suggest that computerized music training or group music therapy may be useful additions to rehabilitation programs for CI users, many of which are mainly based on speech. Note that the present music training approaches are only two of many approaches that might benefit CI users' music perception, speech performance and QoL. Further research is needed to determine the best combination of training exercises to allow CI users to remain engaged, and attending to important cues for speech and music.

Supplemental Material

Supplementary material for this article is available online.

References

Amitay, S., Hawkey, D. J., & Moore, D. R. (2005). Auditory frequency discrimination learning is affected by stimulus variability. *Perception and Psychophysics, 67*(4), 691–698.

Assmann, P. F., & Summerfield, Q. A. (1990). Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *Journal of the Acoustical Society of America*, 88, 680–697. doi: 10.1121/1.399772.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614–636.

Başkent, D., & Gaudrain, E. (2016). Musician advantage for speech-on-speech perception. *Journal of the Acoustical Society of America*, 139(3), EL51–EL56. doi: 10.1121/1.4942628.

Başkent, D., Gaudrain, E., Tamati, T. N., & Wagner, A. (2016). Perception and psychoacoustics of speech in cochlear implant users. In A. T. Cacace, E. de Kleine, A. G. Holt, & P. van Dijk (Eds.), *Scientific foundations of audiology: Perspectives from physics, biology, modeling, and medicine* (pp. 285–319). San Diego, CA: Plural Publishing. ISBN 13:978-1-59756-652-0.

Besson, M., Schon, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience*, 25(3–4), 399–410.

Blamey, P., Artieres, F., Baskent, D., Bergeron, F., Beynon, A., Burke, E.,...Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology & Neuro-Otology*, 18(1), 36–47. doi: 10.1159/000343189.

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *Journal of the Acoustical Society of America*, 137(1), 378–387. doi: 10.1121/1.4942628.

Bosman, A. J., & Smoorenburg, G. F. (1995). Intelligibility of Dutch CVC syllables and sentences for listeners with normal hearing and with three types of hearing impairment. *Audiology*, 34(5), 260–284. doi: 10.3109/00206099509071918.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109(3), 1101–1109. doi: 10.1121/1.1345696.

Chartrand, J. P., & Belin, P. (2006). Superior voice timbre processing in musicians. *Neuroscience Letters*, 405(3), 164–167. doi: 10.1016/j.neulet.2006.06.053.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex (New York, N.Y.: 1991)*, 18(12), 2844–2854. doi: 10.1093/cercor/bhn042.

Clayton, K. K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., & Kidd, G. K. Jr. (2016). Executive function, visual attention and the cocktail party problem in musicians and non-musicians. *PLOS ONE*, 11(7), e0157638. doi: 10.1371/journal.pone.0157638.

Deroche, M. L. D., Limb, C. J., Chatterjee, M., & Gracco, V. L. (2017). Similar abilities of musicians and non-musicians to segregate voices by fundamental frequency. *The Journal of the Acoustical Society of America*, 142, 1739. doi: 10.1121/1.5005496.

Dmitrieva, E. S., Gel'man, V. Y., Zaitseva, K. A., & Orlov, A. M. (2006). Ontogenetic features of the psychophysiological mechanisms of perception of the emotional component of speech in musically gifted children. *Neuroscience and Behavioral Physiology*, 36(1), 53. doi: 10.1007/s11055-005-0162-6.

Donnelly, S., & Walsh, D. (1996). Quality of life assessment in advanced cancer. *Palliative Medicine*, 10(4), 275–283. doi: 10.1177/026921639601000402.

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(5), 779–789.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., & Westermann, S. (2001). ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International collegium for rehabilitative audiology. *Audiology*, 40, 148–157. doi: 10.3109/00206090109073110.

Driscoll, V. D. (2012). The effects of training on recognition of musical instruments by adults with cochlear implants. *Seminars in Hearing*, 33(4), 410–418. doi: 10.1055/s-0032-1329230.

Fu, Q. J., & Galvin, J. J. (2007). Computer-assisted speech training for cochlear implant patients: Feasibility, outcomes, and future directions. *Seminars in Hearing*, 28(2), 1–11. doi: 10.1055/s-2007–973440.

Fu, Q. J., & Galvin, J. J. (2008). Maximizing cochlear implant patients' performance with advanced speech training procedures. *Hearing Research*, 242(1–2), 198–208 doi: 10.1016/j.heares.2007.11.010.

Fu, Q. J., Galvin, J. J., Wang, X., & Wu, J. L. (2015). Benefits of music training in mandarin-speaking pediatric cochlear implant users. *Journal of Speech, Language, and Hearing Research*, 58(1), 163–169. doi: 10.1044/2014_JSLHR-H-14-0127.

Fuller, C. D., Galvin, J. J., Maat, B., Free, R. H., & Başkent, D. (2014). The musician effect: Does it persist under degraded pitch conditions of cochlear implant simulations? *Frontiers in Neuroscience*, 8, 179. doi: 10.3389/fnins.2014.00179.

Fuller, C. D., Mallinckrodt, L., Maat, B., Başkent, D., & Free, R. H. (2013). Music and quality of life in early-deafened late-implanted adult cochlear implant users. *Otology & Neurotology*, 34(6), 1041–1047. doi: 10.1097/MAO.0b013e31828f47dd.

Galvin, J. J., Eskridge, E., Oba, S., & Fu, Q. J. (2012). Melodic contour identification training in cochlear implant users with and without a competing instrument. *Seminars in Hearing*, 33, 399. doi: 10.1055/s-0032-1329227.

Galvin, J. J., Fu, Q. J., & Nogaki, G. (2007). Melodic contour identification by cochlear implant listeners. *Ear and Hearing*, 28(3), 302–319. doi: 10.1097/01.aud.0000261689.35445.20.

Galvin, J. J., Fu, Q. J., & Oba, S. (2008). Effect of instrument timbre on melodic contour identification by cochlear implant users. *The Journal of the Acoustical Society of America*, 124(4), EL189–EL195. doi: 10.1121/1.2961171.

Galvin, J. J., Fu, Q. J., & Oba, S. I. (2009). Effect of a competing instrument on melodic contour identification by

cochlear implant users. *The Journal of the Acoustical Society of America*, *125*(3), EL98–EL103. doi: 10.1121/1.3062148.

Galvin, J. J., Fu, Q. J., & Shannon, R. V. (2009). Melodic contour identification and music perception by cochlear implant users. *Annals of the New York Academy of Sciences*, *1169*, 518–533. doi: 10.1111/j.1749-6632.2009.04551.x.

Gfeller, K. (2001). Aural rehabilitation of music listening for adult cochlear implant recipients: Addressing learner characterisics. *Music Therapy Perspectives*, *19*, 88–95. doi: 10.1093/mtp/19.2.88.

Gfeller, K., Christ, A., Knutson, J. F., Witt, S., Murray, K. T., & Tyler, R. S. (2000). Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *Journal of the American Academy of Audiology*, *11*(7), 390–406.

Gfeller, K., Guthe, E., Driscoll, V., & Brown, C. J. (2015). A preliminary report of music-based training for adult cochlear implant users: Rationales and development. *Cochlear Implants International*, *16*(Suppl. 3), S22–S31. doi: 10.1179/1467010015Z.000000000269.

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., & Olszewski, C. (2007). Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear and Hearing*, *28*(3), 412–423. doi: 10.1097/AUD.0b013e3180479318.

Gfeller, K., Witt, S., Adamek, M., Mehr, M., Rogers, J., Stordahl, J., & Ringgenberg, S. (2002). Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients. *Journal of the American Academy of Audiology*, *13*(3), 132–145.

Gilbers, S., Fuller, C., Gilbers, D., Broersma, M., Goudbeek, M., Free, R., & Başkent, D. (2015). Normal-hearing listeners' and cochlear implant users' perception of pitch cues in emotional speech. *i-Perception*, *6*(5), 0301006615599139. doi: 10.1177/0301006615599139.

Goudbeek, M., & Broersma, M. (2010). The Demo/Kemo corpus: A principled approach to the study of cross-cultural differences in the vocal expression and perception of emotion. In N. Cazolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperdis, M. Rosner, & D. Tapias (Eds.), *7th International Conference on Language Resources and Evaluation (LREC 2010)* (pp. 2211–2215). Valletta, Malta: ELRA.

Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron*, *76*(3), 486–502. doi: 10.1016/j.neuron.2012.10.011.

Hilliard, R. E. (2003). The effects of music therapy on the quality and length of life of people diagnosed with terminal cancer. *Journal of Music Therapy*, *40*(2), 113–137. doi: 10.1093/jmt/40.2.113.

Hinderink, J. B., Krabbe, P. F., & Broek, P. V. D. (2000). Development and application of a health-related quality-of-life instrument for adults with cochlear implants: The Nijmegen cochlear implant questionnaire. *Otolaryngology–Head and Neck Surgery: Official Journal of American Academy of Otolaryngology–Head and Neck Surgery*, *123*(6), 756–765. doi: 10.1590/2317-1782/20172017007.

House, D. (1994). Perception and production of mood in speech by cochlear implant users. In *3rd International Conference on Spoken Language Processing (ICSLP)*.

Hubbard, D. J., & Assmann, P. F. (2013). Perceptual adaptation to gender and expressive properties in speech: The role of fundamental frequency. *The Journal of the Acoustical Society of America*, *133*(4), 2367–2376. doi: 10.1121/1.4792145.

Hütter, E., Argstatter, H., Grapp, M., & Plinkert, P. K. (2015). Music therapy as specific and complementary training for adults after cochlear implantation: A pilot study. *Cochlear Implants International*, *16*(Suppl. 3), S13–S21. doi: 10.1179/1467010015Z.000000000261.

Ingvalson, E. M., Lee, B., Fiebig, P., & Wong, P. C. (2013). The effects of short-term computerized speech-in-noise training on postlingually deafened adult cochlear implant recipients. *Journal of Speech, Language, and Hearing Research*, *56*(1), 81–88. doi: 10.1044/1092-4388(2012/11-0291).

Jiam, N. T., Caldwell, M., Deroche, M. L., Chatterjee, M., & Limb, C. J. (2017). Voice emotion perception and production in cochlear implant users. *Hearing Research*, *352*, 30–39. doi: 10.1016/j.heares.2017.01.006.

Kong, Y. Y., Cruz, R., Jones, J. A., & Zeng, F. G. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear and Hearing*, *25*(2), 173–185. doi: 10.1097/01.AUD.0000120365.97792.2F.

Lassaletta, L., Castro, A., Bastarrica, M., Perez-Mora, R., Herran, B., Sanz, L., . . . Gavilan, J. (2008). Musical perception and enjoyment in post-lingual patients with cochlear implants. *Acta Otorrinolaringologica Espanola*, *59*(5), 228–234. doi: 10.1016/S2173-5735(08)70228-X.

Limb, C. J., & Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hearing Research*, *308*, 13–26. doi: 10.1016/j.heares.2013.04.009.

Lo, C. Y., McMahon, C. M., Looi, V., & Thompson, W. F. (2015). Melodic contour training and its effect on speech in noise, consonant discrimination, and prosody perception for cochlear implant recipients. *Behavioural Neurology*, *2015*, 352869. doi: 10.1155/2015/352869.

Loebach, J. L., & Pisoni, D. B. (2008). Perceptual learning of spectrally degraded speech and environmental sounds. *The Journal of the Acoustical Society of America*, *123*(2), 1126–1139. doi: 10.1121/1.2823453.

Looi, V., Gfeller, K., & Driscoll, V. (2012). Music appreciation and training for cochlear implant recipients: A review. *Seminars in Hearing*, *33*(4), 307–334. doi: 10.1055/s-0032-1329222.

Looi, V., & She, J. (2010). Music perception of cochlear implant users: A questionnaire, and its implications for a music training program. *International Journal of Audiology*, *49*(2), 116–128. doi: 10.3109/14992020903405987.

Looi, V., Wong, Y., & Loo, J. H. Y. (2016). The effects of training on music perception and appreciation for cochlear implant recipients. *Advances in Otolaryngology,* article ID 6286514, 12 pp. doi: 10.1155/2016/6286514.

Luo, X., Fu, Q. J., & Galvin, J. J. (2007). Vocal emotion recognition by normal-hearing listeners and cochlear implant

users. *Trends in Amplification*, *11*(4), 301–315. doi: 10.1177/ 1084713807305301.

Luo, X., Fu, Q.-J., Wu, H.-P., & Hsu, C.-J. (2009). Concurrent-vowel and tone recognition by Mandarin-speaking cochlear implant users. *Hearing Research*, *256*(1), 75–84. doi: 10.1016/j.heares.2009.07.001.

Maden, S. M. K., Whiteford, K. L., & Oxenham, A. J. (2017). Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds. *Scientific Reports*, *7*, 12624. doi: 10.1038/ s41598-017-12937-9.

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*(9), 1453–1463. doi: 10.1162/jocn.2007.19.9.1453.

McDermott, J. H. (2004). Music perception with cochlear implants: A review. *Trends in Amplification*, *8*(2), 49–82.

McDermott, J. H., & Oxenham, A. J. (2008). Music perception, pitch, and the auditory system. *Current Opinion in Neurobiology*, *18*(4), 452–463. doi: 10.1016/ j.conb.2008.09.005.

Mehta, A. H., & Oxenham, A. J. (2007). Vocoder simulations explain complex pitch perception limitations experienced by cochlear implant users. *Journal of the Association of Research in Otolarygology*, *18*(6), 789–802. doi: 10.1007/ s10162-017-0632-x.

Migchelbrink, F., & Brinkman, F. (2000). *Praktijkgericht onderzoek in zorg en welzijn* [Practice-oriented research in care and welfare.]. Amsterdam, the Netherlands: SWP.

Moore, D. R., Rosenberg, J. F., & Coleman, J. S. (2005). Discrimination training of phonemic contrasts enhances phonological processing in mainstream school children. *Brain and Language*, *94*(1), 72–85. doi: 10.1016/ j.bandl.2004.11.009.

Morse-Fortier, C., Parrish, M. M., Baran, J. A., & Freyman, R. L. (2017). The effects of musical training on speech detection in the presence of informational and energetic masking. *Trends in Hearing*, *21*. doi: 10.1177/2331216517739427.

Oba, S. I., Fu, Q. J., & Galvin, J. J. (2011). Digit training in noise can improve cochlear implant users' speech understanding in noise. *Ear and Hearing*, *32*(5), 573–581. doi: 10.1097/AUD.0b013e31820fc821.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, *30*(6), 653–661. doi: 10.1097/ AUD.0b013e3181b412e9.

Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., & Kraus, N. (2011). Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. *PLoS One*, *6*(5), e18082. doi: 10.1371/ journal.pone.0018082.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers of Psychology*, *29*(2), 142. doi: 10.3389/ fpsyg.2011.00142.

Patel, A. D. (2012). The OPERA hypothesis: Assumptions and clarifications. *Annals of the New York Academy of Sciences*, *1252*, 124–128. doi: 10.1111/j.1749-6632.2011.06426.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, *308*, 98–108. doi: 10.1016/ j.heares.2013.08.011.

Pereira, C. (2000). The perception of vocal affect by cochlear implantees. In S. B. Waltzman, & N. L. Cohen (Eds.), *Cochlear implants* (pp. 343–345). New York, NY: Thieme Medical.

Peretz, I., & Zartorre, R. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*, 89–114.

Petersen, B., Mortensen, M. V., Hansen, M., & Vuust, P. (2012). Singing in the key of life: A study on effects of musical ear training after cochlear implantation. *Psychomusicology: Music, Mind, and Brain*, *22*(2), 134–151. doi: 10.1037/a0031140.

Philips, B., Vinck, B., Vel, E. D., Maes, L., D'Haenens, W., Keppler, H., ... Dhooge, I. (2012). Characteristics and determinants of music appreciation in adult CI users. *European Archives of Oto-Rhino-Laryngology*, *269*(3), 813–821. doi: 10.1007/s00405-011-1718-4.

Plomp, R., & Mimpen, A. M. (1979). Speech-reception threshold for sentences as a function of age and noise level. *The Journal of the Acoustical Society of America*, *66*(5), 1333–1342. doi: 10.1121/1.383554.

Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PloS One*, *9*(1), e86980. doi: 10.1371/journal.pone.0086980.

Shannon, R. V., Fu, Q. J., & Galvin, J. J. (2004). The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Oto-laryngologica.Supplementum*, (552), 50–54. doi: 10.1080/ 03655230410017562.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303–304. doi: 10.1126/ science.270.5234.303.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*(6876), 87–90. doi: 10.1038/416087a.

Stacey, P. C., Raine, C. H., O'Donoghue, G. M., Tapper, L., Twomey, T., & Summerfield, A. Q. (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *International Journal of Audiology*, *49*(5), 347–356. doi: 10.3109/14992020903397838.

Stacey, P. C., & Summerfield, A. Q. (2007). Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *The Journal of the Acoustical Society of America*, *121*(5 Pt1), 2923–2935. doi: 10.1121/ 1.2713668.

Stacey, P. C., & Summerfield, A. Q. (2008). Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. *Journal of Speech, Language, and Hearing Research*, *51*(2), 526–538. doi: 10.1044/1092-4388(2008/038).

Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd, G. Jr., & Patel, A. D. (2015). Musical training, individual differences and the cocktail party problem. *Scientific Reports*, *5*, 11628. doi: 10.1038/srep11628.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, 4(1), 46–64. doi: 10.1037/1528-3542.4.1.46.

Vandali, A., Sly, D., Cowan, R., & van Hoesel, R. (2015). Training of cochlear implant users to improve pitch perception in the presence of competing place cues. *Ear and Hearing*, 36(2), e1–e13. doi: 10.1097/AUD.0000000000000109.

Walworth, D., Rumana, C. S., Nguyen, J., & Jarred, J. (2008). Effects of live music therapy sessions on quality of life indicators, medications administered and hospital length of stay for patients undergoing elective surgical procedures for brain. *Journal of Music Therapy*, 45(3), 349–359.

Wright, B. A., Buonomano, D. V., Mahncke, H. W., & Merzenich, M. M. (1997). Learning and generalization of auditory temporal-interval discrimination in humans. *Journal of Neuroscience*, 17(10), 3956–3963.

Zendel, B. R., & Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychology and Aging*, 27(2), 410. doi: 10.1037/a0024816.

Zuk, J., Ozernov-Palchik, O., Kim, H., Lakshminarayanan, K., Gabrieli, J. D., Tallal, P., & Gaab, N. (2013). Enhanced syllable discrimination thresholds in musicians. *Plos One*, 8(12), e80546. doi: 10.1371/journal.pone.0080546.