International Journal of Audiology 2014; Early Online: 1–8

informa
healthcare

Original Article

Music mixing preferences of cochlear implant recipients: A pilot study

Wim Buyens*,†,‡, Bas van Dijk*, Marc Moonen† & Jan Wouters‡

*Cochlear Technology Centre Belgium, Mechelen, Belgium †Department of Electrical Engineering (ESAT-STADIUS), KULeuven, Heverlee, Belgium, and ‡Department of Neurosciences (ExpORL), KULeuven, Leuven, Belgium

Int J Audiol Downloaded from informahealthcare.com by 213.224.16.158 on 01/30/14
For personal use only.

Abstract

Objective: Music perception and appraisal are generally poor in cochlear implant recipients. Simple musical structures, lyrics that are easy to follow, and clear rhythm/beat have been reported among the top factors to enhance music enjoyment. The present study investigated the preference for modified relative instrument levels in music with normal-hearing and cochlear implant subjects. *Design:* In experiment 1, test subjects were given a mixing console and multi-track recordings to determine their most enjoyable audio mix. In experiment 2, a preference rating experiment based on the preferred relative level settings in experiment 1 was performed. *Study sample:* Experiment 1 was performed with four postlingually deafened cochlear implant subjects, experiment 2 with ten normal-hearing and ten cochlear implant subjects. *Results:* A significant difference in preference rating was found between normal-hearing and cochlear implant subjects. The latter preferred an audio mix with larger vocals-to-instruments ratio. In addition, given an audio mix with clear vocals and attenuated instruments, cochlear implant subjects preferred the bass/drum track to be louder than the other instrument tracks. *Conclusions:* The original audio mix in real-world music might not be suitable for cochlear implant recipients. Modifying the relative instrument level settings potentially improves music enjoyment.

Key Words: Cochlear implant; music perception; multi-track; mixing

Introduction

A cochlear implant (CI) is a medical device enabling severe to profoundly deaf people to perceive sounds by electrically stimulating the auditory nerve using an electrode array implanted in the cochlea (see Loizou (1998) for an introduction to cochlear implants). The CI's sound processor uses a sound coding strategy to determine an electrical stimulation pattern from the incoming sound. The main focus of signal processing research for CIs has been on strategies to improve speech understanding in both quiet and noisy environments. Most CI recipients reach good speech understanding in quiet surroundings, however, music perception and appraisal generally remain poor (see McDermott, 2004, for a review). Mirza et al (2003) reported a significant degradation in music enjoyment after implantation by means of comparing a self-assessment scale in the period before deafness and after implantation. A similar decline in listening habits and music enjoyment after implantation was indicated by Tyler et al (2000), Leal et al (2003), Lassaletta et al (2008), and Migirov et al (2009).

Studies on music perception with CI subjects suggested a preference for simple monophonic melodies and rhythmic sounds, whereas complex polyphonic music pieces such as pop, rock, or classical orchestral music were indicated as unpleasant, noisy, or even annoying. Gfeller et al (2000) reported simple musical structures and clear rhythm/beat amongst the top factors that enhance musical enjoyment for CI subjects. The effect of complexity on the appraisal of songs was studied by Gfeller et al (2003) with CI subjects and normal-hearing (NH) subjects with pop, country, and classical music. For CI subjects a negative correlation was found between complexity and appraisal, whereas for NH subjects this correlation was positive. The study also showed that CI subjects rated classical music as more complex than pop and country music. Several plausible explanations were indicated by Gfeller et al (2003). On the one hand, from the standpoint of objective complexity, the structural characteristics of classical music tend to be more complex than pop or country music. Classical music tends to use more complex rhythmic structures and harmonic changes compared to pop and country music, which mainly consists of short simple melodies built over simple and repetitive rhythms and harmonic changes. On the other hand, the pop and country items studied by Gfeller et al (2003) had lyrics while the classical items did not. The speech-type information from the lyrics was indicated as an additional factor regarding complexity for CI subjects. The lyrics might have provided some CI subjects with enough useable acoustic information to more easily follow the sequence of events. In addition, both pop and country music

Correspondence: Wim Buyens, Cochlear Technology Centre Belgium, Schaliënhoevedreef 20 I, B-2800 Mechelen, Belgium. E-mail: WBuyens@cochlear.com

(Received 5 March 2013; accepted 6 December 2013)

ISSN 1499-2027 print/ISSN 1708-8186 online © 2014 British Society of Audiology, International Society of Audiology, and Nordic Audiological Society
DOI: 10.3109/14992027.2013.873955

RIGHTSLINK

Abbreviations

AGC	Automatic gain control
CI	Cochlear implant
dB	Decibel
F0	Fundamental frequency
ICC	Intraclass correlation coefficient
NH	Normal hearing
SD	Standard deviation
SNR	Signal-to-noise ratio

are likely to contain a strong, simple beat which is more suitable for transmission through current-day implants than the structural features of instrumental classical songs.

Along with the enjoyment of rhythm and beat, the performance for CI subjects on rhythmic pattern perception tasks is nearly as good as for NH subjects (e.g. Kong et al, 2004; McDermott, 2004; Looi et al, 2008). Low frequency modulations, such as rhythm and beat, are well encoded in the electrical stimulation pattern with current CIs, whereas the accurate transmission of spectral information remains challenging (Loizou, 1998). Due to the limited access to fundamental frequency (F0) and timbre cues, the segregation of competing voices and/or music instruments is difficult. Galvin et al (2009) showed significantly worse performance in the melodic contour identification task in the presence of a competing instrument as masker. Zhu et al (2011) reported on the inability of most CI subjects to use or combine pitch, timbre and/or timing cues in this melodic contour identification task with a competing masker, resulting in poor segregation. The masking effect of piano music on speech perception was also studied for NH and hearing-impaired subjects by Ekström & Borg (2011). A piano composition played at high speed in a low octave had greater masking effect than when played at a low speed in a high octave. Quality ratings for music with a single instrument, a solo instrument with background music, and ensemble music were studied by Looi et al (2007), which indicated that CI subjects judged music with multiple instruments as less pleasant than music played by a single instrument. Donnelly et al (2009) reported near-chance-level performance for CI subjects in recognizing two- and three-pitch stimuli, demonstrating perceptual fusion of multiple pitches as single-pitch units.

The cited studies on music perception, in which music with simple musical structures and clear rhythm/beat were generally preferred over more "complex" music, all assessed existing music with an intrinsic level of complexity. No attempt was made to reduce the complexity of the "complex" songs themselves in order to increase music enjoyment. The present study investigated the effect of complexity reduction in "complex" songs on music enjoyment with CI subjects. From the many possible structural sources of complexity within music, the current study focused primarily on one source of complexity, which is the addition of background harmony or percussion to a solo voice. Other sources of complexity such as more complicated melody lines or more complex harmonies were not considered. Music heard on CD, radio, TV, and at concerts usually contains background harmony and rhythm, as opposed to being comprised of only a solo melody. The present study investigated the effects of modifying the relative amplitude of the lead vocal melody with respect to the harmonic and percussion accompaniments in real-world music, and the resulting listener sound-quality ratings. The world of all possible forms/styles of music is enormous and cannot be fully addressed in one study. Consequently, the scope of this study was focused on one commonly heard type of music, pop music, which is constituted of a music band accompanying a vocalist singing the main melody. Such music bands typically include piano, guitar, bass guitar and drums.

The primary research question of the current study was to address the preference for modified relative instrument level settings of real-world pop music with postlingually deafened CI subjects and NH subjects, including the influence of song familiarity. In order to answer this question, two experiments were conducted. In experiment 1, CI test subjects (N = 4) were given a mixing console and multi-track recordings with which they were asked to make an audio mix that sounded most enjoyable to them. In experiment 2, a preference rating experiment was performed by NH (N = 10) and CI subjects (N = 10) based on the preferred relative level settings for the different tracks from experiment 1.

Methods

In this study, it was hypothesized that CI subjects prefer a different audio mix than NH subjects. This hypothesis was investigated in a first experiment with a mixing console and multi-track recordings, and in a second more controlled experiment with a preference rating analysis with predefined audio mixes.

Sound material

For this music experiment multi-track recordings from well-known as well as unfamiliar pop songs have been collected. Songs were chosen to consist of a vocal track with the main melody and background music with different complexity level, ranging from basic accompaniment with piano or guitar to a full music band typically consisting of piano, guitar, bass guitar, and drums. In Table 1 the selected songs with the corresponding instrument tracks are shown. Songs 1 and 2 consist of sung lyrics with basic accompaniment of one instrument (piano or guitar), songs 3 and 4 include additional accompaniment with bass guitar, whereas songs 5 and 6 have additional drums.

The different tracks were collected either from artists and record companies, or by gathering professional backing tracks and by recording the vocals at the Pop and Jazz division of the Municipal Academy of Music in Antwerp (Belgium). Only the multi-track recordings of song 2 (Table 1) were provided by the original artist. The mean dynamic range of all tracks equals 10.1 dB (SD 2.7 dB),

Table 1. Multi-track recordings for experiment 1 (song 6) and experiment 2 (song 1–6). The recordings include vocals and background music as indicated. Only the recordings of song 2 were provided by the original artist (V = vocals, P = piano, G = guitar, B = bass guitar, D = drums).

	Song	*V*	*P*	*G*	*B*	*D*
1	Hallelujah (Leonard Cohen)	x	x			
2	Before I Go (Papermouth)	x		x		
3	Michel (Anouk)	x		x	x	
4	Hey Jude (excerpt A) (The Beatles)	x		x	x	
5	Hey Jude (excerpt B) (The Beatles)	x		x	x	x
6	Dock of the Bay (Otis Redding)	x	x	x	x	x

Int J Audiol Downloaded from informahealthcare.com by 213.224.16.158 on 01/30/14
For personal use only.

RIGHTSLINK

Int J Audiol Downloaded from informahealthcare.com by 213.224.16.158 on 01/30/14
For personal use only.

which is acceptable for the type of music used in this experiment (Pleasurize Music Foundation, 2013).

The separately recorded tracks were mixed down to MONO wav-files (44.1 kHz). From these tracks, small excerpts were selected based on the characteristics of the background music. These excerpts took on average 29 seconds and were played in a loop during the experiment. All tracks were normalized with the loudness measurement from the ReplayGain specification, which is a standard that describes a way of calculating and representing the ideal replay gain for a given track or album, and that provides guidance for players to make the required level adjustment during playback (ReplayGain, 2013). The loudness measurement in this ReplayGain specification includes three steps. In the first step, the audio signal is filtered by an inverted approximation of the equal loudness curves to account for the frequency specific sensitivity of the human ear. A cascade of a 10th order IIR filter designed by MATLAB's "yulewalk" function with a 2nd order Butterworth high-pass filter provides a good approximation. The filter coefficients for both filters are listed in ReplayGain (2013). In the second step, the RMS level of the filtered signal is calculated every 50 ms. The 95th percentile of all calculated RMS levels is finally chosen to represent the overall perceived loudness. The summation of the normalized vocal track and instrument tracks was assumed to represent the studio mix for normal hearing listeners. This assumption has been confirmed by a semi-professional studio mixer. In experiment 1, the multi-track recordings for song 6 (see Table 1) including vocal, piano, guitar, bass guitar, and drums were combined in three configurations: a configuration with two mix channels containing vocals and background music respectively, a configuration with three mix channels containing vocals, piano/guitar, and bass/drum respectively and a configuration with five mix channels containing vocals and all instruments separately. The mix channels for the different configurations were derived by adding the corresponding tracks and by normalizing them with ReplayGain. For comparing the results of the three configurations at a later stage all level settings were recalculated to individual instrument-to-vocal levels. In experiment 2 all songs listed in Table 1 were included. The sound samples can be made available by contacting the authors.

Subjects

Eleven postlingually deafened CI subjects (all using the Cochlear™ Nucleus® system) participated in the present study. A summary with demographic and etiological information can be found in Table 2. Postlingually deafened CI subjects were selected, because they might still have a memory of how music should sound. None of them was a professional musician before they deafened. All subjects used their everyday map with the default ACE sound coding strategy. The subjects signed a consent form and were paid for their travel expenses. Ethical committee approval was obtained. Subjects S1-S4 determined their preferred audio mix with a mixing console in experiment 1. The pairwise comparison experiments in experiment 2 were carried out with subjects S2-S11. Ten NH subjects with no self-reported hearing deficit also participated in experiment 2 and their age varied from 29 years to 65 years with an average of 45 years.

Experiment 1

The initial experiment was based on the multi-track recordings of song 6 from Table 1, which includes the typical instruments in a pop song: vocals, piano, guitar, bass guitar, and drums. Three different configurations have been assembled from the multi-track recordings of the selected song: a configuration with two mix channels (vocals and background music), a configuration with three mix channels (vocals, piano/guitar, and bass/drums) and a configuration with five mix channels (vocals and all instruments separately). Four CI subjects (S1–S4) were asked to determine the most enjoyable audio mix by using a mixing console and the mix channels from the different configurations. During the experiment, the multi-track excerpt continuously played in a loop while the test subject could adjust the level for each mix channel separately. One trial run was given in advance with the mixing console in order to learn the system. No specific strategy was imposed on the test subjects to perform the task. However, it was suggested to listen to each mix channel separately before starting the experiment. Initially, all mix channel sliders on the mixing console were placed in the down position in order not to prime the test subject with a specific audio mix. The music was played in free-field in a sound-treated room and was limited to a maximum level of 65 dB(A) with all mix channels at maximum level. The experiment was performed on a laptop with multi-track editing software (Cubase Essential 5) connected to a MIDI mixing console (Behringer BCF2000) and a loudspeaker (Genelec 8020A). The experiment was repeated ten times in each configuration with the selected song (song 6 from Table 1). In order to prevent the test subject from using visual cues on the mixing console, a different random level offset was applied to each individual mix channel

Table 2. Overview with demographic and etiological information about the eleven postlingually deafened CI subjects (all using the Cochlear™ Nucleus® system) who participated in experiment 1 (S1–S4) and experiment 2 (S2–S11).

Subject	*Age (years)*	*Gender*	*CI experience (years)*	*Etiology*	*Sound processor*	*Implant type*
S1	73	Female	3	Progressive	Freedom®	CI24RE(CA)
S2	59	Male	7	Unknown	Freedom	CI24R(CS)
S3	54	Female	2	Unknown	Freedom	CI24RE(CA)
S4	26	Male	8	Meningitis	Freedom	CI24R(CS)
S5	64	Male	8	Otosclerosis	CP810	CI24R(CS)
S6	60	Male	16	Otosclerosis	Freedom	CI24R(CS)
S7	42	Male	13	Progressive	Freedom	CI24R(CS)
S8	74	Female	4	Otosclerosis	Freedom	CI24RE(CA)
S9	50	Male	1	Colitis ulcerosa	CP810	CI24RE(CA)
S10	62	Male	2	Progressive	CP810	CI24RE(CA)
S11	64	Male	11	Noise	CP810	CI24R(CS)

RIGHTSLINK

for each configuration and each run (0, −2 dB, −4 dB, −6 dB, −8 dB). Consequently, the determined audio mix resulted from listening to the music rather than from looking at the mixing console. The test subject had no time constraint for the task and could take a break whenever necessary. Time to complete the mixing experiment varied between 1 hour and 2 hours. The preferred audio mix was determined by the slider levels for the different mix channels measured in dB relative to the vocals and individual instrument-to-vocals ratios were computed from the mix channel slider levels for comparison across the different configurations.

Experiment 2

In the subsequent experiment, the preferred relative level settings for CI subjects were investigated further using a pairwise comparison analysis. In this experiment all songs from Table 1 were used to include different background accompaniment. The pairwise comparison analysis included four conditions with different relative level settings for the separately recorded tracks (Table 3). Condition "Standard" was composed by adding the normalized vocal track and instrument tracks, which was assumed to represent the studio mix for normal hearing listeners. Conditions "− 6dBMusic" and "− 12dBMusic" originated from the two mix channel settings of the different CI subjects in experiment 1 with, respectively, 6-dB and 12-dB attenuation for the instruments relative to the vocals. Condition "VocalBassDrum" originated from the three mix channel settings in experiment 1, in which the bass/drum channel is less attenuated than the piano/guitar channel. In a first series of experiments, the conditions "Standard", "− 6dBMusic" and "− 12dBMusic" that differ only in vocals-to-instruments ratio as shown in Table 3 were pairwise compared. All possible pairs with these three conditions were presented in a random order and repeated three times for all six songs listed in Table 1, starting with the songs with basic accompaniment (one instrument) towards the songs with all instruments including drums. The total number of presented pairs was 54. The music was played in free-field through a loudspeaker (Genelec 8020A) at a fixed level of 65 dBSPL, which is below the AGC knee point in the CI signal path, passing the signal with 0 dB gain. 10 NH and 10 CI subjects (S2–S11) were asked to select the condition that sounded most enjoyable to them and to quantify their preference with a rating score ranging from *Imperceptible*, *Slightly better*, *Better*, *Largely better*, to *Hugely better*. An additional experiment comparing the conditions "− 12dBMusic" and "VocalBassDrum" with attenuated instruments and different bass/drum level was added for CI subjects only (S2–S11). Table 3 shows the corresponding relative level settings for the different instrument tracks. The preference rating between two conditions is indicated in percentages. The percentage scores represent the preference for the first condition over the second condition for all 180 comparisons (ten subjects, six songs, and three repetitions). The percentage scores were calculated by dividing the frequency of preference for the first condition by the total number of comparisons (180), without taking into account the degree of enjoyment. The preference for the second condition is complementary (and is not shown). The additional preference rating score from *Imperceptible* to *Hugely better* indicates the strength of the preference.

Table 3. Overview of the predefined relative level settings for the different tracks (in dB) for the conditions in the pairwise comparison of experiment 2. Condition "Standard", "− 6dBMusic", and "− 12dBMusic" differ in vocals-to-instruments ratio only. Condition "VocalBassDrum" represents an audio mix with attenuated instruments in which the bass/drum is attenuated less than the piano/guitar.

	Vocals	*Piano*	*Guitar*	*Bass*	*Drums*
"Standard"	0	0	0	0	0
"− 6dBMusic"	0	− 6	− 6	− 6	− 6
"− 12dBMusic"	0	− 12	− 12	− 12	− 12
"VocalBassDrum"	0	− 12	− 12	− 3	− 3

Statistics

Given the small sample size in the first experiment, only individual results were shown. Differences in level settings between vocals and different instruments were analysed with the non-parametric Friedman test followed by the post-hoc Wilcoxon signed-ranks test with Bonferroni correction. Intraclass correlation coefficient was used to check reliability of the data.

For the preference rating in experiment 2 the chi-square test was used to determine which of the two conditions was preferred and to analyse the effect of familiarity and singer's gender on audio mix preference. Chi-square statistics is applied on the absolute preference between two conditions (without taking into account the 5-step rating score). In the null hypothesis the frequency of preferred items are the same in both conditions. A significant chi-square test results in rejecting the null hypothesis, which means a significant preference for one of the two conditions. The correlation between audio mix preference and CI experience was analyzed with Pearson's correlation coefficient. The one-sample Wilcoxon signed-ranks test was used to study the different audio mix preferences for the three defined groups in experiment 2.

Results

Experiment 1

Figure 1 shows the preferred level settings for the instrument tracks of song 6 (Table 1) for the configuration with two, three, and five mix channels respectively. The different instruments of the song are shown on the horizontal axis, whereas the vertical axis represents the preferred level (in dB) relative to the vocal level (0 dB). Individual results are shown for the 4 CI subjects (S1–S4) including 10 repetitions per configuration per subject. In the configuration with two mix channels (top row in Figure 1), the preferred vocal level is significantly higher than the instruments level for all subjects with a vocals-to-instruments ratio ranging from 3 dB (S2) to 15 dB (S3) (Wilcoxon signed-ranks test: $Z = 2.80$, $p = .005$, $r = .63$). The configuration with three mix channels (middle row in Figure 1) demonstrates the same preference for the vocals, but in addition it distinguishes between the piano/guitar and the bass/drum mix channel. For subjects S3 and S4, the vocals are set significantly louder than the mix channel with piano/guitar (Wilcoxon signed-ranks test: $Z = 2.80$, $p = .015$, $r = .63$) and bass/drum (Wilcoxon signed-ranks test: $Z = 2.80$, $p = .015$, $r = .63$ for S3 and $Z = 2.60$, $p = .027$, $r = .58$ for S4) and the mix channel with bass/drum is set louder than the mix channel with piano/guitar for subjects S2–S4, which, however, is not significant after Bonferroni correction (Wilcoxon signed-ranks test: $Z = 2.29$, $p = .066$, $r = .51$ for S4). The results for the configuration with five mix channels are shown on the bottom row in Figure 1, but reveal no significant preference for a particular instrument after

Int J Audiol Downloaded from informahealthcare.com by 213.224.16.158 on 01/30/14
For personal use only.

RIGHTSLINK

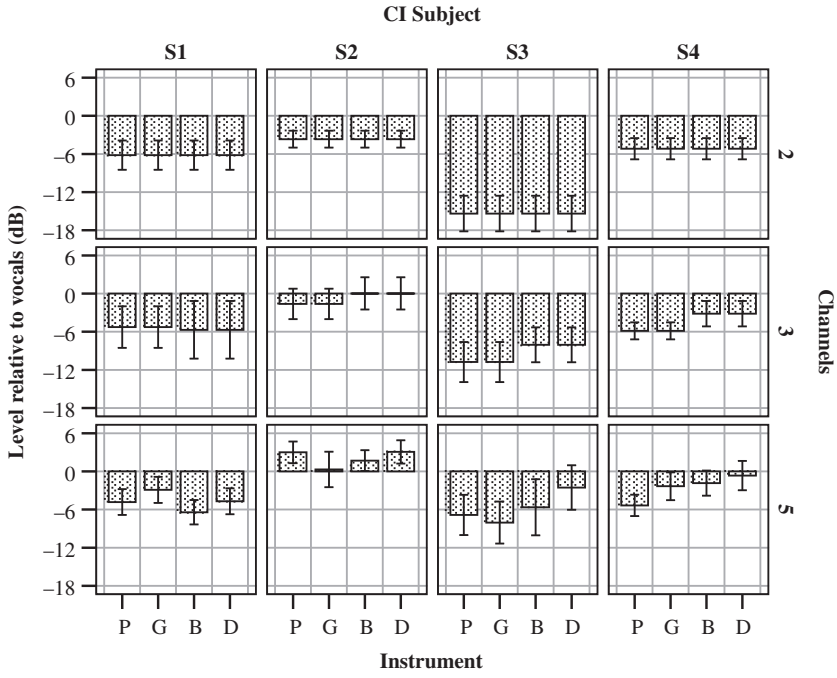

Figure 1. Results for experiment 1 indicating the preferred level settings for the different instruments (P = piano, G = guitar, B = bass, D = drums) relative to the vocal level (0 dB) for song 6 (Table 1). Individual level settings for CI subjects S1–S4 are shown in three configurations: configuration with two mix channels (all instruments together), configuration with three mix channels (piano/guitar together and bass/drum together), and configuration with five mix channels (all instruments separately available). Error bars indicate 95% confidence interval.

Int J Audiol Downloaded from informahealthcare.com by 213.224.16.158 on 01/30/14
For personal use only.

Bonferroni correction (Wilcoxon signed-ranks test: $p > .05$). The reliability of the test results has been checked with the intraclass correlation coefficient (ICC), showing high reliability for the results of subjects S2 (ICC = .89), S3 (ICC = .90) and S4 (ICC = .83) and low reliability for the results of subject S1 (ICC = .32).

Experiment 2

Figure 2 shows the results for 10 NH and 10 CI subjects (S2–S11) for the pairwise comparison between condition "Standard", "− 6dBMusic", and "− 12dBMusic" in which condition "Standard" represents the studio mix for normal hearing listeners and condition "− 6dBMusic" and "− 12dBMusic" correspond to an audio mix with the instrument tracks attenuated by 6 dB and 12 dB respectively. The results for NH subjects indicate a significant preference for the condition with an equal balance between vocals and instruments (condition "Standard") compared to condition "− 6dBMusic" (80%), $\chi^2(1) = 64.8$, $p < .001$, and condition "− 12dBMusic" (96%), $\chi^2(1) = 149$, $p < .001$. A significant preference was also found for condition "− 6dBMusic" compared to condition "− 12dBMusic" (97%), $\chi^2(1) = 161$, $p < .001$. On the contrary, the results with CI subjects indicate the opposite preference in the comparison between condition "Standard" and "− 6dBMusic". The condition with 6 dB attenuated instruments (condition "− 6dBMusic") was preferred over the condition with an equal balance between vocals and instruments (condition "Standard") with 64%, which is significant ($\chi^2(1) = 13.9$, $p < .001$). Between condition "Standard" and "− 12dBMusic", no overall preference was found over all CI subjects ($\chi^2(1) = 0.36$, $p = .55$), whereas condition "− 6dBMusic"was significantly preferred over condition "− 12dBMusic" (63%), similar to NH subjects ($\chi^2(1) = 11.8$, $p = .001$).

Following the preference rating experiment, the test subjects were asked about their familiarity with the songs provided, resulting in a classification of known (22/60) and unknown songs (38/60). No

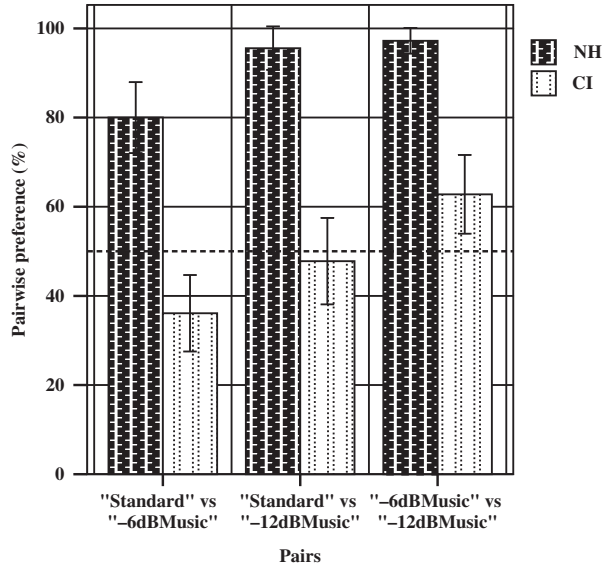

Figure 2. Results for experiment 2 indicating the pairwise comparison of condition "Standard", "− 6dBMusic", and "− 12dBMusic" with 10 NH and 10 CI subjects (S2–S11). Each bar represents the percentage scores for the preference of the first condition over the second condition for the 180 comparisons (ten subjects, six songs and three repetitions). The dashed line represents chance level. Error bars indicate 95% confidence interval.

RIGHTSLINK

effect of familiarity of the songs on the preference rating was found (Pearson $\chi^2(1) = 0.49$, $p = .49$). However, in the pairwise comparison between condition "Standard" and "− 6dBMusic", a preference of 66% was found for condition "−6dBMusic" in the group of unknown songs (95% confidence interval: 57%–75%) which is significant ($\chi^2(1) = 11.4$, $p = .001$), whereas in the group of known songs the preference (61% with 95% confidence interval: 48%–73%) was not significant, $\chi^2(1) = 2.97$, $p = .085$.

In the group of songs without drums (songs 1–4) which is equally divided between female (songs 1 and 3) and male singers (songs 2 and 4), a difference in preference rating was observed depending on the singer"s gender. For the songs with a male singer a preference of 67% was found for condition "− 12dBMusic" over condition "Standard" (95% confidence interval: 51%–83%), which is significant ($\chi^2(1) = 6.67$, $p = .010$), whereas for the songs with a female singer the preference (45% with 95% confidence interval: 27%–63%) is not significant ($\chi^2(1) = 0.60$, $p = .44$). The effect of the singer"s gender in the pairwise comparison between condition "Standard" and "− 12dBMusic" is significant (Pearson $\chi^2(1) = 5.71$, $p = .017$).

The preference for condition "− 6dBMusic" with clear vocals and attenuated instruments versus condition "Standard" is negatively correlated with the CI experience of the test subject, meaning that less-experienced CI subjects showed stronger preference for the condition with attenuated instruments than experienced subjects. This correlation is significant with a Pearson's $r(10) = -.70$, $p = .025$.

In an additional preference rating experiment with CI subjects (S2–S11), two conditions with attenuated instruments and different bass/drum level were compared (songs 5 and 6). The condition with prominent bass/drum track (condition "VocalBassDrum") obtained a significant preference of 87% (95% confidence interval: 82%–92%) over the condition with all instruments attenuated (condition "− 12dBMusic") ($\chi^2(1) = 96.8$, $p < .001$). The preference results for the pairwise comparison between condition "VocalBassDrum" and "− 12dBMusic" are compared to the preference results of the pairwise comparison between condition "Standard" and "− 12dBMusic" for songs 5 and 6 in Figure 3. Because of the individual differences in preference rating between condition "Standard" and "− 12dBMusic", the CI subjects were divided in three groups based on their preference rating, resulting in a group with significant preference for condition "Standard" (N = 4), (one-sample Wilcoxon signed-ranks test: $p < .001$), a group with no preference between "Standard" and "− 12dBMusic" (N = 3), (one-sample Wilcoxon signed-ranks test: $p = .46$), and a group with significant preference for condition "− 12dBMusic" (N = 3), (one-sample Wilcoxon signed-ranks test: $p = .003$). In the comparison between condition "VocalBassDrum" and "− 12dBMusic", all groups had significant preference for condition "VocalBassDrum" (one-sample Wilcoxon signed-ranks test: $p < .001$). It should be noted that even the group with subjects who significantly preferred condition "− 12dBMusic" with attenuated instruments over condition "Standard", also significantly preferred condition "VocalBassDrum" with prominent bass/drum over condition "− 12dBMusic".

Discussion

The primary research question of the current study was to address the preference for modified relative instrument level settings of real-world pop music with postlingually deafened CI subjects and NH subjects. In order to answer this question, two experiments were conducted. In the initial experiment 1, CI subjects determined their preferred audio mix with a mixing console and multi-track recordings. Based on the preferred relative level settings of the different tracks in experiment 1, four conditions have been defined for the pairwise comparison in experiment 2 with NH and CI subjects.

In experiment 1, CI subjects composed their preferred audio mix with all separately recorded tracks available. A significant preference for clear vocals and attenuated instruments in the audio mix is observed, especially in the configuration with two and three mix channels. However, it is noted that the overall preferred vocals-to-background music ratio is decreasing in the configurations with more mix channels available. A reason for this could be that in the configuration with two mix channels, the focus for the test subject was only on the vocals-to-background music ratio, whereas in the other configurations the focus was on the relative level differences between all tracks. The CI test subjects have also reported an increasing difficulty in making the audio mixes with increasing number of mix channels available. Because of the many repetitions of the song in the experiment, it might also be that the test subject became more familiar with the song in the more difficult five mix channel configuration towards the end, allowing more music in the audio mix.

Experiment 2 consists of a pairwise comparison analysis with predefined relative level settings for the different tracks based on the results of experiment 1 in the configuration with two mix chan-

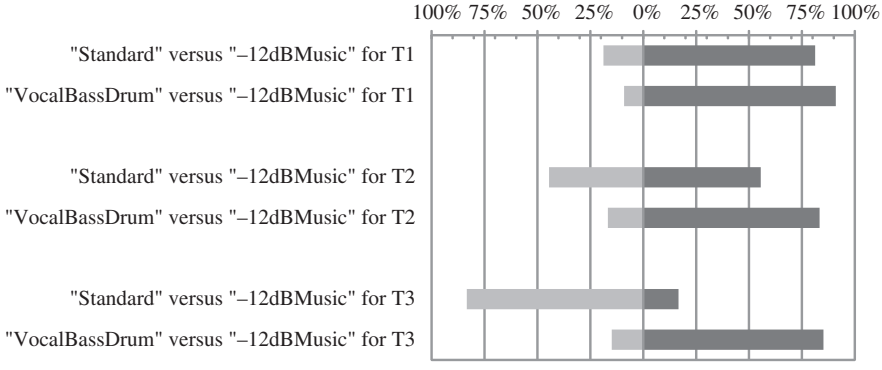

Figure 3. Results for experiment 2 indicating the preference results for the pairwise comparison of condition "Standard" versus "− 12dBMusic" and condition "VocalBassDrum" versus "− 12dBMusic" with 10 CI subjects (S2–S11) for all songs from Table 1 with a drum track available (songs 5 and 6). CI subjects are grouped based on their preference results for the comparison "Standard" versus "− 12dBMusic" (T1 (4), T2 (3), T3 (3)). The light bar indicates the preference for condition "− 12dBMusic".

Int J Audiol Downloaded from informahealthcare.com by 213.224.16.158 on 01/30/14
For personal use only.

RIGHTSLINK

Int J Audiol Downloaded from informahealthcare.com by 213.224.16.158 on 01/30/14
For personal use only.

nels and three mix channels. The results obtained in experiment 1 with two mix channels are confirmed by experiment 2. CI subjects significantly preferred the condition with clear vocals and attenuated instruments compared to the condition with equally balanced vocals and instruments. For the limited songs provided, no significant effect of familiarity with the music was found on the preference rating. The condition with clear vocals and attenuated instruments was preferred for both known and unknown songs and might help CI subjects to follow songs with lyrics more easily.

The difference in preference rating between songs with male and female singers could be explained by the difference in difficulty in the segregation between vocals and background music. Gfeller et al (2009) studied the impact of voice range on the recognition of song lyrics with CI subjects and found that the SNRs needed in song lyrics recognition for male vocalists were poorer than for female vocalists. Because segregating a male singer from background music might be more difficult than a female singer, the audio mix with attenuated instruments was preferred for songs with a male singer, whereas for songs with a female singer no significant preference was found in the pairwise comparison between condition "Standard" and "− 12dBMusic".

The additional pairwise comparison experiment between two conditions with attenuated instruments and different bass/drum level showed a significant preference for the condition with the louder bass/drum tracks. The prominent presence of the bass/drum tracks together with clear vocals provides an audio mix that was more enjoyable for CI subjects. Gfeller et al (2000) determined the ability to follow the musical score or words as a top factor for increasing musical enjoyment with CI subjects. According to the same study, features of music that enhance musical enjoyment include the presence of clear rhythm/beat and a simple musical structure. The preference results in the present study are in agreement with these findings.

The attenuation of background music is a possible way to reduce the objective complexity of a complex song. Moreover, clear vocals and clear drums might have reduced the perceived subjective complexity for CI subjects as well, because elements such as song lyrics or rhythm are more effectively transmitted and more accurately heard by CI subjects than pitch, timbre, and harmony. Low frequency modulations, such as rhythm and beat, are well encoded in the electrical stimulation pattern with current CIs, whereas the accurate transmission of spectral and fine-structure information of sounds remains challenging due to channel interactions, limited number of stimulation channels and limited dynamic range. The described reduction in complexity resulted in more enjoyable music for CI subjects. This is in agreement with the findings of Gfeller et al (2003), in which a negative correlation was determined for CI subjects between complexity and appraisal with different styles of music. Classical music was rated more complex than pop and country music, which was suggested to be due to the lack of a clear melody and lyrics in classical music. These findings correspond with the results from the present study, in which music with a clear vocal melody and attenuated instruments was determined to be more enjoyable for CI subjects.

The large variability across CI subjects indicated that the modified relative instrument level settings as studied in the present study did not necessarily increase music enjoyment for all CI subjects. A negative correlation was found between CI experience and the preference for an audio mix with clear vocals and attenuated instruments, meaning that less-experienced CI recipients benefited more from the modified relative instrument level settings of the "original" audio mix than experienced recipients did.

The "original" audio mix, which was assumed to consist of an equal balance between vocals and instruments, was significantly preferred by NH test subjects. On the contrary, CI subjects preferred an audio mix with clear vocals and attenuated instruments. However, given the limited sample size and the variability among CI subjects, these findings might not be generalizable. Moreover, the results with this style of music should not be presumed to apply to other styles of music that have important structural differences. Consequently, similar tests with other musical styles would help to determine to what extent these modified settings are beneficial for other non-pop music.

Conclusion

Normal hearing subjects significantly preferred the "original" audio mix with an equal balance between the vocal melody and the different instruments. However, this preferred mix is not necessarily the ideal audio mix for CI subjects. For the pop songs provided in the present study, CI subjects significantly preferred an audio mix with clear vocals and attenuated instruments. In addition, given an audio mix with clear vocals and attenuated instruments, CI subjects significantly preferred the bass/drum track to be louder than the other instruments. The variability in the preference results across CI subjects indicated that the modified relative instrument level settings as studied in the current study did not necessarily increase music enjoyment for all CI subjects. In particular, the data indicated that less-experienced CI subjects benefited more from the modified relative instrument level settings in the original audio mix of complex songs than experienced recipients did.

Acknowledgments

We would like to thank all CI and NH subjects for participating in this music experiment, and Astrid Van Wieringen from ExpORL - KULeuven for scheduling the CI subjects. Also thanks to all people involved in collecting and recording the multi-track recordings (in alphabetic order): Anthony Claeys, Bart Delacourt, Bart Dirckx, Ellen Peeters, Gunter Peeters, and Michel Verkempinck.

Declaration of interest: This work was supported by a Baekeland PhD grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT090274) and the Cochlear Technology Centre Belgium. Authors Wim Buyens and Bas van Dijk are employees of Cochlear Technology Centre Belgium, which is part of Cochlear Ltd.

References

Donnelly P.J., Guo B.Z. & Limb C.J. 2009. Perceptual fusion of polyphonic pitch in cochlear implant users. *J Acoust Soc Am*, 126, EL128–133.

Ekström S. & Borg E. 2011. Hearing speech in music. *Noise Health*, 13, 277–285.

Galvin J.J. III, Fu Q. & Oba S.I. 2009. Effect of a competing instrument on melodic contour identification by cochlear implant users. *J Acoust Soc Am*, 125, EL 98–103.

Gfeller K., Buzzell A., Driscoll V., Kinnaird B. & Oleson J. 2009. The impact of voice range and instrumental background accompaniment on recognition of song lyrics. *Proc APSCI*, 35–39.

Gfeller K., Christ A., Knutson J., Witt S. & Mehr M. 2003. The effects of familiarity and complexity on appraisal of complex songs by

RIGHTSLINK

cochlear implant recipients and normal-hearing adults. *J Music Ther*, XL, 78–113.

Gfeller K., Christ A., Knutson J., Witt S., Murray K.T. et al. 2000. Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *J Am Acad Audiol*, 11, 390–406.

Kong Y., Cruz R., Jones J.A. & Zang F. 2004. Music perception with temporal cues in acoustic and electric hearing. *Ear Hear*, 25, 173–185.

Lassaletta L., Castro A., Bastarrica M., Pérez-Mora R., Herrán B. et al. 2008. Changes in listening habits and quality of musical sound after cochlear implantation. *Otolaryngol Head Neck Surg*, 138, 363–367.

Leal M.C., Shin Y.J., Laborde M., Calmels M., Verges S. et al. 2003. Music perception in adult cochlear implant recipients, *Acta Otolaryngol*, 123, 826–835.

Loizou P. 1998. Introduction to cochlear implants. *IEEE Signal Proc Mag*, 15, 101–130.

Looi V., McDermott H., McKay C. & Hickson L. 2007. Comparisons of quality ratings for music by cochlear implant and hearing-aid users. *Ear Hear*, 28, 59S–61S.

Looi V., McDermott H., McKay C. & Hickson L. 2008. Music perception of cochlear implant users compared to that of hearing-aid users. *Ear Hear*, 29, 421–434.

McDermott H. 2004. Music perception with cochlear implants: A review. *Trends Amplif*, 8, 49–82.

Migirov L., Kronenberg J. & Henkin Y. 2009. Self-reported listening habits and enjoyment of music among adult cochlear implant recipients. *Ann Otol Rhinol Laryngol*, 118, 350–355.

Mirza S., Douglas S.A., Lindsey P., Hildreth T. & Hawthorne M. 2003. Appreciation of music in adult patients with cochlear implants: A patient questionnaire. *Cochlear Implants Int*, 4, 85–95.

PleasurizeMusicFoundation. 2013. http://www.pleasurizemusic.com (date last viewed 30/01/13).

ReplayGain. 2013. http://www.replaygain.org (date last viewed 30/01/13).

Tyler R.S., Gfeller K. & Mehr M.A. 2000. A preliminary investigation comparing one and eight channels at fast and slow rates on music appraisal in adults with cochlear implants. *Cochlear Implants Int*, 1, 82–87.

Zhu M., Chen B., Galvin J.J. III & Fu Q. 2011. Influence of pitch, timbre, and timing cues on melodic contour identification with a competing masker (L). *J Acoust Soc Am*, 130, 3562–3565.

Int J Audiol Downloaded from informahealthcare.com by 213.224.16.158 on 01/30/14
For personal use only.

RIGHTSLINK