KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Buyens W., Van Dijk B., Moonen M., Wouters J., ``Evaluation of a stereo music pre-proocessing scheme for cochlear implant users'', *Journal of the American Academy of Audiology*, 2017, pp. 1-9
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	http://dx.doi.org/10.3766/jaaa.16103
Journal homepage	http://www.audiology.org/publications/about-journal-american-academy-audiology
IR	https://lirias.kuleuven.be/handle/123456789/565714

(article begins on next page)



Evaluation of a Stereo Music Preprocessing Scheme for Cochlear Implant Users

DOI: 10.3766/jaaa.16103

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

Abstract

Background: Although for most cochlear implant (CI) users good speech understanding is reached (at least in quiet environments), the perception and the appraisal of music are generally unsatisfactory.

Purpose: The improvement in music appraisal was evaluated in CI participants by using a stereo music preprocessing scheme implemented on a take-home device, in a comfortable listening environment. The preprocessing allowed adjusting the balance among vocals/bass/drums and other instruments, and was evaluated for different genres of music. The correlation between the preferred settings and the participants' speech and pitch detection performance was investigated.

Research Design: During the initial visit preceding the take-home test, the participants' speech-in-noise perception and pitch detection performance were measured, and a questionnaire about their music involvement was completed. The take-home device was provided, including the stereo music preprocessing scheme and seven playlists with six songs each. The participants were asked to adjust the balance by means of a turning wheel to make the music sound most enjoyable, and to repeat this three times for all songs.

Study Sample: Twelve postlingually deafened CI users participated in the study.

Data Collection and Analysis: The data were collected by means of a take-home device, which preserved all the preferred settings for the different songs. Statistical analysis was done with a Friedman test (with post hoc Wilcoxon signed-rank test) to check the effect of ''Genre.'' The correlations were investigated with Pearson's and Spearman's correlation coefficients.

Results: All participants preferred a balance significantly different from the original balance. Differences across participants were observed which could not be explained by perceptual abilities. An effect of ''Genre'' was found, showing significantly smaller preferred deviation from the original balance for Golden Oldies compared to the other genres.

Conclusions: The stereo music preprocessing scheme showed an improvement in music appraisal with complex music and hence might be a good tool for music listening, training, or rehabilitation for CI users.

Key Words: cochlear implants, music appraisal, music processing

Abbreviations: CI = cochlear implant; DR = dynamic range; SD = standard deviation; SRT = Speech Reception Threshold

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

Corresponding author: Wim Buyens, Cochlear Technology Centre Belgium, B-2800 Mechelen, Belgium; E-mail: wbuyens@cochlear.com

This work was supported by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT090274 and IWT150280) 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.

INTRODUCTION

A cochlear implant (CI) is a medical device that provides auditory sensations to recipients with severe-to-profound hearing loss by electrically stimulating the auditory nerve using an electrode array implanted in the cochlea (Loizou, 1998). Most CI users reach good speech understanding in quiet environments, but speech performance in noise varies across recipients, and the perception of music is generally unsatisfactory for CI users (McDermott, 2004). Earlier reports demonstrated poor pitch perception, poor melody recognition, inadequate timbre perception, and a degradation of music enjoyment after implantation compared to the period before becoming deaf (Mirza et al, 2003; McDermott, 2004). A sound-coding strategy to improve music perception for CI users was proposed and evaluated by Laneau et al (2006) and Milczynski et al (2009). Interesting improvements were shown for pitch perception, melodic contour identification, and familiar melody identification, but no improvement in music enjoyment with complex music was revealed.

CI devices have been developed and optimized primarily for transmitting speech sounds. The transmission of the features in musical sounds is limited in CI devices due to the poor pitch representation (Limb, 2006; Limb and Roy, 2014). At the level of electrode-auditory nerve interface, further degradation of the musical sounds—and sound representation in general—is caused by the spread of electrical stimulation, the limited stimulation of apical cochlear regions, the potential mismatch between the frequency band mapped to an electrode (the so-called frequency allocation table) and the frequency corresponding to the actual electrode position in the cochlea due to the tonotopic organization, and the limited phase locking to auditory input with electric hearing. Furthermore, CI recipients have a considerably reduced dynamic range (DR) with a reduced number of distinguishable steps in loudness, which also leads to an additional decrease in perceived music sound quality. On the other hand, poor pitch and timbre perception do not necessarily result in poor music appreciation. Drennan et al (2015) reported on a clinical evaluation of music perception, appraisal, and experience in CI participants and showed only a weak relationship between perceptual accuracy and music enjoyment, suggesting that perception and appraisal are relatively independent for CI users. Also by Gfeller et al (2008) and Lassaletta et al (2008), no or weak association was found between music enjoyment and music perception skills. Wright and Uchanski (2012) studied music appraisal and perception skills in CI participants as well as normal-hearing participants listening to CI-simulated music. The normal-hearing participant results provided a reasonable model for many music perception skills of CI participants, but not for their rating of music appraisal. Here again, only weak or nonexistent correlations were found between appraisal ratings and music perception scores.

The effect of complexity on music appraisal was studied by Gfeller et al (2003), which showed a strong negative correlation between complexity and appraisal for CI participants. Music excerpts that were (subjectively) rated as less complex (such as country music) were appreciated more than music excerpts that were rated as more complex (such as classical music). Moreover, CI participants judged music that involved multiple instruments, on average, less pleasant than music played by single instruments (Looi et al, 2007). The reduction of complexity by modifying the relative instrument levels in the audio mix of complex music was studied with CI participants by Buyens et al (2014). The participants were asked to make their preferred audio mix by using multi-track recordings and a mixing console. Subsequently, a preference rating experiment was performed with predefined audio mixes, showing a significant preference for an audio mix with clear vocals and attenuated instruments, while preserving the bass/drum track. Clear rhythm/beat and lyrics that are easy to follow were also found among the top factors to enhance music enjoyment by Gfeller et al (2000). Since multi-track recordings are not widely available for most commercial music, signal-processing techniques are required to modify the relative instrument levels in complex music. A stereo music preprocessing scheme that is able to perform these modifications in stereo music was studied in Buyens et al (2015) and was evaluated in CI participants with pop/rock music excerpts. Results showed that this music preprocessing scheme potentially improves music appraisal by adjusting the balance between vocals/bass/drums and other instruments. Individual differences across participants were observed.

In the present study, the potential improvement in music appraisal with this stereo music preprocessing scheme was further investigated for different genres of music and in a home-listening environment, which is more comfortable and arguably more realistic than a sound booth. In the following section, the sound material and the take-home device are described, together with the design of the take-home test.

METHODS

The stereo music preprocessing scheme, which was described in Buyens et al (2015), was evaluated with postlingually deafened CI participants in a take-home test. In the "Sound Material" section, the sound material used for this evaluation is described. The "Participants" section contains the demographic and etiological information of the CI participants. In the "Take-Home Device" section, the take-home device is presented, and, finally, in the "Test Procedure" section, the test procedure is explained in detail.

Sound Material

It is obviously impossible to include all existing musical genres and subgenres in a single study. The former "Head of Music Belgium" of Mood Media (and still a professional DJ) was asked to provide a selection of different music genres with specific sound characteristics that included vocals and that were generally widespread. The six selected genres were Disco/Funk/Soul, Golden Oldies, Latin, Pop (Dance), Rock, and songs with Dutch lyrics. The first genre was Disco/Funk/Soul, which is characterized by lots of rhythm and percussion, and a prominent brass section. The next genre was Golden Oldies, which contained songs from the 1960s that are still popular nowadays. The genre Latin included Latin-American music such as salsa, merengue, and bachata. These songs have Spanish lyrics and create a tropical atmosphere. Pop (Dance) is a genre of popular music that typically contains a strong electronic beat and synthesized sounds, supporting the vocal lyrics. Rock music is a genre of popular music in which the vocals are typically complemented with electric guitars, bass guitar, piano, and drums. Finally, the songs with Dutch lyrics contained Dutch cabaret music. Typically, these songs have a simple structure and a limited number of mainly acoustic instruments. The emphasis in these songs is on the lyrics. For each genre, seven songs were chosen that were representative of the specific characteristics of that genre. The sound material consisted of the commercially available stereo recordings with sampling frequency of 44.1 kHz, with a total duration of >150 min. Table 1 lists the different genres with total duration together with their average DR (averaged over the seven songs), which gives an indication of the compression used in the final mix of the music. For the take-home test, seven playlists were compiled, each containing one song from each genre, randomized per participant. To assess the familiarity with the songs, the list of songs was provided and participants were asked to indicate whether they were familiar with the song or not.

Table 1. Sound Material for Take-Home Test, Divided into Six Music Genres with Their Total Duration and the Mean and SD of the DR Over All Songs

Music Genre	Total Duration (min:sec)	Mean DR Value (dB)	SD (dB)
Disco/Funk/Soul	24:36	10.0	1.9
Golden Oldies	22:57	11.0	1.7
Latin	29:48	9.3	2.9
Pop (Dance)	28:14	6.3	0.8
Rock	26:33	6.6	2.2
Dutch songs	25:32	8.6	2.1

Participants

Twelve postlingually deafened CI users (all Cochlear™ Nucleus®) participated in the study. They were recruited with an advertisement on social media and in newsletters and mailings from user groups. A summary with demographic and etiological information can be found in Table 2. The participants signed a consent form and were paid for their travel expenses. Ethical committee approval was obtained for this study. All participants were using the default advanced combination encoder strategy in their sound processor. They were asked to do the experiment in a comfortable and quiet environment with their program fixed to the one they are using every day and with their own preferred mixing ratio between microphone and input. No effect of sound processor was observed.

Take-Home Device

The participants received a take-home device to be able to listen to the sound material in a comfortable listening environment. This device was a commercially available iPhone5 with a custom-made application for the experiment containing all the sound material. A "Personal Audio Cable" was used to link the take-home device with the participants' CI sound processor. The three important features of the application were (a) the music library access and navigation buttons, (b) the turning wheel for adjusting the balance between vocals/bass/drums and the other instruments, and (c) the "Vote" button to store the preferred setting for each song on the device. Figure 1 shows a screenshot of the application on the take-home device.

With the music library access button, the participants could select one of the seven playlists. With the navigation buttons, the participants could navigate through the songs of the selected playlist. The song name and song progress were displayed on the screen while the song was played. The turning wheel effectively controlled the attenuation parameter in the stereo music preprocessing scheme, as described in Buyens et al (2015). The stereo music preprocessing scheme separated vocals, bass, and drums from the other instruments by exploiting the representation of harmonic and percussive components in the input power spectrogram together with the spatial information contained in typical stereo recordings. The output signal with the extracted vocals, bass, and drums was mixed together with the attenuated residual signal (other instruments). The attenuation parameter ranged from −6 to +24 dB, where positive values represented an attenuation of the residual signal, whereas negative values represented an amplification of the residual signal. An attenuation parameter of zero represented the original balance between vocals/bass/drums and other instruments. In order not to have abrupt changes when turning the wheel, a mirrored version of the scale of the attenuation parameter was added

Table 2. Demographic and Etiological Information of the 12 Postlingually Deafened CI Participants in the Study

Participant	Age (Years)	Gender	CI Experience (Years)	Etiology	Sound Processor
S1	64	Male	3	Progressive	CP810
S2	29	Male	11	Meningitis	CP810
S3	56	Female	6	Progressive	CP910
S4	62	Female	7	Unknown	CP810
S5	62	Male	9	Otosclerosis	CP810
S6	29	Female	5	Unknown	CP810
S7	67	Male	9	Streptomizine	CP910
S8	70	Male	9	Congenital	CP810
S9	64	Male	17	Otosclerosis	CP910
S10	52	Female	3	Unknown	CP810 + CP910
S11	74	Male	10	Struma	CP810
S12	45	Female	14	Meningitis	CP910

to the original scale. A random rotation of this doubled scale was then applied each time the navigation buttons were used to select a new song, in order not to prime the participants with a visual cue. Finally, the "Vote" button stored the preferred attenuation parameter setting for each song. The counter on the "Vote" button represented the number of times the preferred setting for a certain song was stored. By pushing the "Vote" button, the logging file on the take-home device was updated with one entry containing song name, song progress, preferred attenuation parameter setting, and random rotation, together with a time stamp.

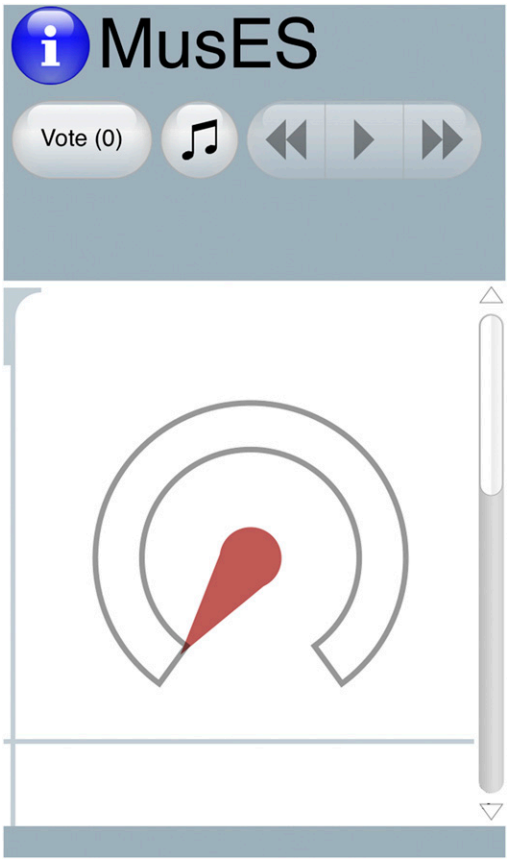

Figure 1. Screenshot of the music application on the take-home device, including the music library access and navigation buttons, the turning wheel to adjust the attenuation parameter, and the "Vote" button to store the preferred setting. (This figure appears in color in the online version of this article.)

Test Procedure

Before the start of the test with the take-home device, participants were asked to perform two initial experiments to determine their speech understanding in noise and pitch discrimination performance, and to fill in a questionnaire about their music involvement and music experience before and after implantation. The speech understanding performance was measured with an adaptive Speech Reception Threshold (SRT) Test with Leuven Intelligibility Sentence Test sentences and Speech Shaped Noise (van Wieringen and Wouters, 2008). The SRT is expressed in dB and represents the signal-to-noise ratio at 50% correct level. The average of two SRT test runs (ten sentences) was used to indicate the speech-in-noise performance. If the difference between the two SRT test results was >2 dB, a third SRT test was performed, and the average of the last two SRT test results was used. The participants' pitch discrimination abilities were measured with the pitch discrimination subtest from the University of Washington Clinical Assessment of Music Perception, which is a two-alternative forced-choice adaptive procedure to determine a threshold interval for discrimination of complex pitch direction change (Nimmons et al, 2008; Kang et al, 2009). Synthesized complex tones with three different base frequencies were used, corresponding to middle C, E above middle C, and G above middle C (262, 330, and 391 Hz). The tone with the base frequency and a higher-pitched tone determined by the random adaptive interval size were presented in random order, and the participants had to indicate which of the two tones was higher in pitch. The amplitude of each tone was randomly roved ±4 dB

to eliminate any loudness cues. The minimum tested interval was 1 semitone, and the maximum was 12 semitones or 1 octave. The test included three randomly presented trials for each base frequency and ended when eight reversals were reached. The threshold for each base frequency was calculated by averaging the last six reversals for each trial. Finally, the mean threshold over the three base frequencies was used to indicate the pitch discrimination ability. After the initial experiments, the participants were asked to complete a questionnaire that was based on the music questionnaire from Mirza et al (2003) translated into Dutch. It queried the participants about their CI experience, music involvement, and appraisal, and it included questions about playing an instrument and singing before deafness and after implantation.

After the initial experiments and the questionnaire, a take-home device was provided to the participants together with a detailed manual. During a training session, the handling of the take-home device was explained. The participants were asked to select one playlist at a time and to listen to the songs one by one while adjusting the balance between vocals/bass/drums and the other instruments by means of the turning wheel. The participants were asked to listen to all the songs from the playlist three times to have three preferred attenuation parameter settings for each song. The same procedure was repeated for all playlists. Additional instructions were provided regarding the voting moment, especially for songs with a long intro. If in the beginning of a song only small or no changes were noticed, the voting moment had to be delayed until the chorus, in which vocals and all instruments are present. After the first visit with initial experiments and training, the participants took the device home for the listening test. The participants were also asked to indicate their familiarity with the songs. The time foreseen for the take-home test was two weeks, with an optional intermediate visit if problems arose. During the final visit, the logging file of the take-home device was read and feedback from the participant was discussed.

RESULTS

The preferred attenuation parameter settings were stored in the take-home device three times for the seven songs from each of the six music genres. The average time spent before voting per song over all participants was 92 sec (with a standard deviation [SD] of 31 sec). A first data check was performed before the analysis of the preferred settings. Three possible issues were considered. The first possible issue was an unintended voting, which appeared as a preferred setting with song progress between 0 and 5 sec. The second possible issue related to an early voting during the intro of a song, whereas the instruction indicated that the preferred setting had to be adjusted after the intro with vocals and instruments present. All early voting data were discarded before further data analysis. The third possible issue was missing or incomplete data, that is, absent data in the returned take-home device after the music experiment. In total, 3.7% of the data were discarded or missing after the first data check.

Three runs with preferred attenuation parameter were registered for every song. A Friedman test with factor "Run" was performed on the data (N = 12) to check the effect of time on the preferred attenuation setting. No significant effect was found [$\chi^2(2) = 0.41$, $p = 0.82$]. Therefore, the median of the three runs was taken for further analysis. However, when looking into individual results for participant S7, a median preferred attenuation of 13 dB was observed in the first run, whereas the median preferred attenuation for the second and the third run was only 5.5 and 8.5 dB, respectively. For participant S9, the opposite trend was observed, with, in the first run, a median preferred attenuation of 5.5 dB, and in the second and third run a median preferred attenuation of 12.5 and 11 dB, respectively.

A Friedman test with factor "Genre" was performed on the median of the three measurements (N = 12) to check the effect of "Genre" on the preferred attenuation setting. A significant effect of "Genre" was found [$\chi^2(5) = 29.07$, $p < 0.001$]. The post hoc Wilcoxon signed-rank test (with Bonferroni correction) was performed to look into this significant effect. The songs from genre Golden Oldies had significantly lower preferred attenuation settings compared to songs with Dutch lyrics ($Z = -4.41$, $p < 0.001$), compared to Rock ($Z = -4.28$, $p < 0.001$), compared to Disco ($Z = -3.37$, $p = 0.015$), and compared to Latin ($Z = -3.12$, $p = 0.03$). The difference in preferred attenuation for Golden Oldies compared to Pop was not significant ($Z = -1.85$, $p = 0.96$). The mean preferred attenuation parameter setting for each genre is shown in Figure 2. There was no significant correlation between the mean DR value (Table 1) and the preferred attenuation parameter setting for each genre [Pearson's $r(6) = -0.36$, $p = 0.49$, Spearman's $\rho(6) = -0.37$, $p = 0.47$].

The effect of "Familiarity" of the songs on the preferred setting was investigated with a Mann–Whitney test. The test participants had to indicate, for each song, whether it was known or unknown. The overall group of known songs had 36.5% of the songs, the group of unknown songs 63.5%. No significant difference was found for the preferred attenuation setting between the two groups ($U = 26{,}750$, $p = 0.30$).

The mean preferred attenuation parameter setting over all songs for all participants is shown in Figure 3. Differences in mean preferred attenuation parameter settings were observed across participants ranging from 4 dB (for participant S2) to 16 dB (for participant S10). A one-sample Wilcoxon signed-rank test with test

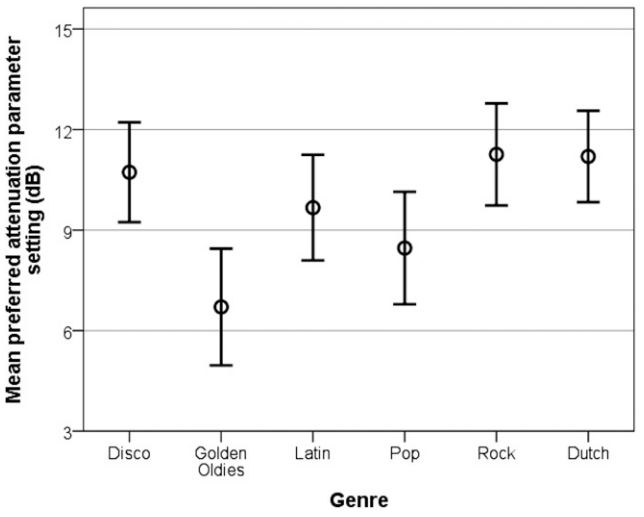

Figure 2. Mean preferred attenuation parameter setting (dB) for the different genres. Positive values represent an attenuation of the "other instruments," negative values (not shown) represent an amplification of the "other instruments," and value zero (not shown) represents the original balance. Error bars indicate 95% confidence interval.

value = 0 was used to check the preferred attenuation setting against the original setting. This test was significant ($p < 0.001$) and thus revealed a preferred attenuation setting significantly different from zero over all participants. No correlation was found between the mean preferred setting and the SRT for speech-in-noise [Pearson's $r(12) = -0.35, p = 0.26$, Spearman's $\rho(12) = -0.32$, $p = 0.31$], nor with the pitch detection performance [Pearson's $r(12) = -0.17$, $p = 0.59$, Spearman's $\rho(12) = -0.21$, $p = 0.51$], nor with the CI experience

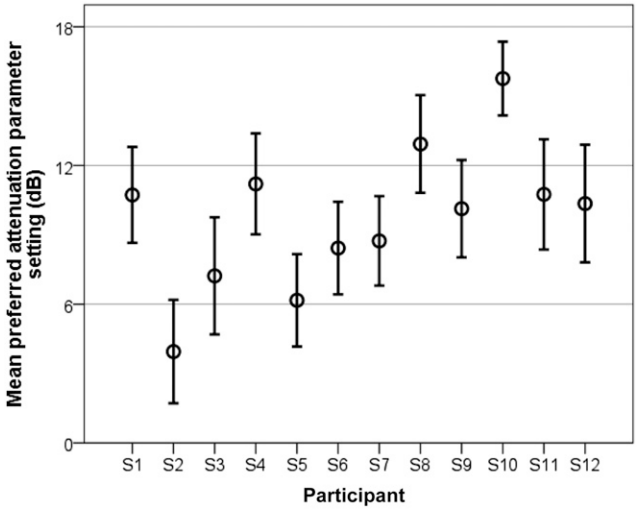

Figure 3. Mean preferred attenuation parameter setting (dB) over all songs for all participants in the take-home test. Positive values represent an attenuation of "other instruments," negative values (not shown) represent an amplification of "other instruments," and value zero represents the original balance. Error bars indicate 95% confidence interval.

[Pearson's $r(12) = -0.29$, $p = 0.36$, Spearman's $\rho(12) = -0.34, p = 0.29$], nor with any item from the music involvement questionnaire. An overview of the SRT scores for speech-in-noise, pitch discrimination scores, and preferred attenuation settings is listed in Table 3.

The range of the three preferred settings collected for one song, which was defined as the difference between the maximum preferred attenuation and the minimum preferred attenuation, could take a value from 0 to 30 dB (= 24 to −6 dB) and was a measure for the strength of the preference for the attenuation setting for the song. A small range indicated a strong preference for the attenuation setting and/or large audible effect from the music preprocessing, whereas a large range could be considered as a weak preference for the attenuation setting or small audible effect of the music preprocessing. The median range over all songs and all participants was 11 dB. A significant effect of "Genre" was found on the computed range of the three preferred settings collected for one song [Friedman with factor "Genre," $\chi^2(5) = 13.12, p = 0.022$]. Post hoc analysis with Wilcoxon signed-rank test (with Bonferroni correction) revealed a significantly lower range for Pop compared to Disco ($Z = -3.14$, $p = 0.03$), and a significantly lower range for Rock compared to Disco ($Z = -3.54$, $p < 0.001$); or, phrased in a different way, the range of the three preferred settings collected for Disco songs was significantly higher compared to the range in Pop and Rock songs.

In the questionnaire, participants were asked for their singing activities before and after implantation (see Table 3). The six participants who reported singing activities before implantation had on average a significantly smaller range over the three preferred attenuation parameter settings (Mann–Whitney test: $U = 4.5, p = 0.026$). The five participants who indicated singing activities after implantation (see Table 3) were significantly more experienced CI users (i.e., in terms of years of CI usage) (Mann–Whitney test: $U = 2, p = 0.01$) and showed a preference for lower attenuation parameter settings, although not significant (Mann–Whitney test: $U = 6$, $p = 0.073$).

DISCUSSION

The main objective of this study was to evaluate the improvement in music appraisal in CI participants with a stereo music preprocessing scheme, for different genres of music and in a comfortable listening environment. The relation between the preferred attenuation parameter settings and the speech and pitch performance was investigated as well as the participants' music involvement. The stereo music preprocessing scheme was described in Buyens et al (2015) and showed encouraging results in laboratory experiments with pop/rock song excerpts. Consequently, a take-home test was set up for the

Table 3. SRT Score for Speech-In-Noise (dB), Pitch Discrimination (Semitones), Preferred Attenuation (dB), Median Range of Three Runs, Familiarity with the Songs (%), Singing Activity before Implantation, Singing Activity after Implantation, and Music Enjoyment (0–10) of the 12 CI Participants

Participant	SRT (dB)	Pitch (Semitones)	Preferred Attenuation (dB)	Median Range (dB)	Familiarity (%)	Singing Before	Singing After	Music Enjoyment
S1	1	1	11	14	45	No	No	0
S2	8	3	4	8	31	Yes	Yes	5
S3	−4	1	7	11	45	No	No	9
S4	−2	2	11	10	48	No	No	7
S5	2	1	6	13	45	Yes	Yes	6
S6	4	3	8	14	64	No	No	10
S7	0	3	9	8	5	Yes	Yes	10
S8	0	2	13	10	33	No	No	0
S9	−1	3	10	7	38	Yes	Yes	0
S10	−2	2	16	7	41	Yes	No	9
S11	7	1	11	11	12	No	No	8
S12	9	2	10	10	31	Yes	Yes	9

assessment of the stereo music preprocessing scheme with other music genres. After validating the registered input data, the consistency of the data over the three runs was checked. No effect of "Run" was found, and thus the median of the three runs was used for further analysis. The median was preferred over the mean to exclude the influence from possible outliers. During the feedback discussion when returning the take-home device, participant S7 indicated himself the difference in voting between the first run and the second/third run. For the first run, he focused on understanding the lyrics and consequently adjusted the attenuation parameter to emphasize the vocals. For the second and the third run, however, he focused more on the song as a whole, and adjusted the attenuation parameter to a lower level than in the first run. In participant S9, a trend in the opposite direction was observed, showing a lower preferred setting for the first run and a higher preferred setting for the second and third runs.

A significant effect of "Genre" was found for the preferred setting. The preferred setting was lowest for Golden Oldies. A possible explanation for this is that these songs were less complex than the songs from the other genres, and consequently required less attenuation of the instruments. Gfeller et al (2003) showed a negative correlation between complexity and appraisal for CI users and also Buyens et al (2015) found a preference for lower attenuation of instruments for songs with a lower complexity compared to songs with a higher complexity. Another possible explanation is related to the different audio mixing trends in older recordings (such as Golden Oldies). Older recordings typically have a larger DR where vocals stand out of the background music, whereas in contemporary music, the vocals are more engraved in the music and the audio mix is heavily compressed (Vickers, 2010). Therefore, the vocals enhancement of the stereo music preprocessing scheme was less needed in Golden Oldies compared to the other genres. On the contrary, during the feedback session participants reported a strong advantage of emphasizing the vocals, especially in songs with Dutch lyrics. With the adjustment of the attenuation parameter, the participants (all Dutch speaking) were able to understand the lyrics better. Some of them indicated that this was less required for songs with English or Spanish lyrics, because that was not their mother tongue. The understanding of the lyrics as a confounding effect in the preference settings was not addressed in the current study. Nevertheless, also for the songs without Dutch lyrics, the preferred mix with clear vocals/bass/drums was significantly different from the original mix. Since these songs included lyrics in languages that were less or not familiar to them, this would suggest that the CI participants preferred an enhancement of the vocal melody even without understanding the lyrics.

All participants preferred a setting that was different from the original balance between vocals/bass/drums and the other instruments, but individual differences were observed across participants. In Buyens et al (2014), a significant negative correlation was found between the pairwise preference for an audio mix with attenuated instruments and CI experience, but this was not found in the current study. To understand the individual differences, the relationship with speech and pitch detection performance was investigated, but no correlation was found between the preferred settings and the participants' speech perception performance or pitch detection abilities. The task in the experiment was to adjust the attenuation parameter to make the music most enjoyable, which seems to be unrelated to perceptual abilities. Drennan et al (2015) also showed only weak relationships between perceptual accuracy and music enjoyment, suggesting that perception and appraisal are relatively independent for CI users. Similar conclusions were found by Gfeller et al (2008), Lassaletta et al (2008), and Wright and Uchanski (2012).

The range over the three preferred settings collected for a song indicated the strength of the preference for the determined setting or the audibility of the music preprocessing effect. This was investigated for the different genres. The median range over all songs including all genres was 11 dB. For the Disco songs, significantly weaker preference for the determined settings (i.e., higher range) was found compared to Pop and Rock, which showed a stronger preference for the determined setting. No significant difference in range was found for Golden Oldies, Latin, and songs with Dutch lyrics. The explanation for this may relate to the different instruments used in the different genres, with, on the one hand, Pop and Rock music with mostly vocals, guitar, piano, bass guitar, and drums; and on the other hand, Disco music with additional hand clapping, extensive percussion, singers in harmony, and prominent brass sections, which may make the effect of the music preprocessing less noticeable. Participants who indicated singing activities before implantation had on average a smaller range of the three preferred settings collected for one song, which indicates a stronger preference for the determined preferred settings or better ability to hear out the effect of the stereo music preprocessing scheme.

Informal feedback after the experiment was mostly positive, confirming the appreciation for clear vocals and reduction of the disturbing "chaos" in the background. A small number of participants reported a certain difficulty in finding the preferred setting, especially for unknown songs or genres, but most of the participants liked the experiment and enjoyed playing around with the balance between vocals/bass/drums and the other instruments. They reported that for most songs, a difference between good and bad was noticeable by adjusting the attenuation parameter.

CONCLUSION

A stereo music preprocessing scheme implemented on a take-home device was assessed with 12 postlingually deafened CI participants in a take-home test. The preferred setting for the adjustable attenuation parameter, which balances vocals/bass/drums against other instruments, was investigated for different genres of music. All participants preferred an attenuation parameter setting to construct a balance with attenuated instruments significantly different from the original mix. Individual differences across participants were observed. This could not be explained by perceptual abilities as speech perception or pitch detection performance. An effect of "Genre" was found, showing lower preferred attenuation settings for Golden Oldies compared to the other genres. Since the complexity of music was reduced with the stereo music preprocessing scheme, it may provide a good tool for music training or rehabilitation programs. CI recipients could start listening to "simplified" music, and gradually increase complexity by mixing the other instruments back in.

Acknowledgments. The authors thank all CI participants for participating in this study, Hans Buyens for collecting the music tracks, and Robert Cerny (dlab GmbH) and Tony Van den Eynde for their support in programming the take-home devices.

REFERENCES

Buyens W, van Dijk B, Moonen M, Wouters J. (2014) Music mixing preferences of cochlear implant recipients: a pilot study. *Int J Audiol* 53(5):294–301.

Buyens W, van Dijk B, Wouters J, Moonen M. (2015) A stereo music pre-processing scheme for cochlear implant users. *IEEE Trans Biomed Eng* 62(10):2434–2442.

Drennan WR, Oleson JJ, Gfeller K, Crosson J, Driscoll VD, Won JH, Anderson ES, Rubinstein JT. (2015) Clinical evaluation of music perception, appraisal and experience in cochlear implant users. *Int J Audiol* 54(2):114–123.

Gfeller K, Christ A, Knutson J, Witt S, Mehr M. (2003) The effects of familiarity and complexity on appraisal of complex songs by cochlear implant recipients and normal hearing adults. *J Music Ther* 40(2):78–112.

Gfeller K, Christ A, Knutson JF, Witt S, Murray KT, Tyler RS. (2000) Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *J Am Acad Audiol* 11(7): 390–406.

Gfeller K, Oleson J, Knutson JF, Breheny P, Driscoll V, Olszewski C. (2008) Multivariate predictors of music perception and appraisal by adult cochlear implant users. *J Am Acad Audiol* 19(2):120–134.

Kang R, Nimmons GL, Drennan W, Longnion J, Ruffin C, Nie K, Won JH, Worman T, Yueh B, Rubinstein J. (2009) Development and validation of the University of Washington Clinical Assessment of Music Perception test. *Ear Hear* 30(4):411–418.

Laneau J, Wouters J, Moonen M. (2006) Improved music perception with explicit pitch coding in cochlear implants. *Audiol Neurootol* 11(1):38–52.

Lassaletta L, Castro A, Bastarrica M, Pérez-Mora R, Herrán B, Sanz L, de Sarriá MJ, Gavilán J. (2008) Changes in listening habits and quality of musical sound after cochlear implantation. *Otolaryngol Head Neck Surg* 138(3):363–367.

Limb CJ. (2006) Cochlear implant-mediated perception of music. *Curr Opin Otolaryngol Head Neck Surg* 14(5):337–340.

Limb CJ, Roy AT. (2014) Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear Res* 308:13–26.

Loizou PC. (1998) Introduction to cochlear implants. *IEEE Signal Process 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(2, Suppl):59S–61S.

McDermott HJ. (2004) Music perception with cochlear implants: a review. *Trends Amplif* 8(2):49–82.

Milczynski M, Wouters J, van Wieringen A. (2009) Improved fundamental frequency coding in cochlear implant signal processing. *J Acoust Soc Am* 125(4):2260–2271.

Mirza S, Douglas SA, Lindsey P, Hildreth T, Hawthorne M. (2003) Appreciation of music in adult patients with cochlear implants: a patient questionnaire. *Cochlear Implants Int* 4(2): 85–95.

Nimmons GL, Kang RS, Drennan WR, Longnion J, Ruffin C, Worman T, Yueh B, Rubenstien JT. (2008) Clinical assessment of music perception in cochlear implant listeners. *Otol Neurotol* 29(2):149–155.

van Wieringen A, Wouters J. (2008) LIST and LINT: sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands. *Int J Audiol* 47(6):348–355.

Vickers E. (2010) The loudness war: background, speculation, and recommendations. In: *Audio Engineering Society Convention 129*. AES 129th Convention, San Francisco, CA, 2010 November 4–7, 2010.

Wright R, Uchanski RM. (2012) Music perception and appraisal: cochlear implant users and simulated cochlear implant listening. *J Am Acad Audiol* 23(5):350–365.