Effect of Compression on Musical Sound Quality in Cochlear Implant Users

Melanie Gilbert, Patpong Jiradejvong, and Charles Limb

Objectives: Cochlear implant (CI) users frequently report poor sound quality while listening to music, although the specific parameters responsible for this loss of sound quality remain poorly understood. Audio compression, which reduces the dynamic range (DR) for a given sound, is a ubiquitous component of signal processing used by both CI and hearing aid technology. However, the relative impact of compression for acoustic and electric hearing on music perception has not been well studied, an important consideration especially given that most compression algorithms in CIs were developed to optimize speech perception. The authors hypothesized that normal-hearing (NH) listeners would detect increased levels of compression more easily than CI users, but that both groups would perceive a loss of sound quality with increasing compression levels.

Design: The present study utilizes the Cochlear Implant-MUltiple Stimulus with Hidden Reference and Anchor to evaluate the listener sensitivity to increasing levels of compression applied to music stimuli. The Cochlear Implant-MUltiple Stimulus with Hidden Reference and Anchor is a tool used to assess relative changes in the perceived sound quality of music across increasingly degraded listening conditions, in both CI and NH subjects. In this study, the authors applied multiple iterations of an aggressive compression algorithm to the music clips using Adobe Audition. The test conditions included 1, 3, 5, and 20 iterations sound tokens, with the 20-iteration samples serving as the Anchor stimuli. The compressed excerpts were 5 sec in length, with five clips for each of the five common musical genres (i.e., Classical, Jazz, Country, Rock, and Hip-Hop). Subjects were also presented with a Reference excerpt, which was the original music clip without any additional compression applied. CI recipients (n = 7, 11 ears) and NH listeners (n = 10) were asked to rate the sound quality of additionally compressed music as compared to the Reference.

Results: Although both NH and CI groups could detect sound quality differences as a function of compression level, the discriminatory ability of the CI group was blunted compared to the NH group. The CI group had less variability in their responses and overall demonstrated reduced sensitivity to deterioration caused by excessive levels of compression. On average, the CI group rated the Anchor condition as only "Slightly worse" than the Reference. The music clips that were most affected by the compression were from Jazz and Hip-Hop genres and less so for Rock and Country clips. Corollary to this was a small but statistically significant impact of DR of the music clips on sound quality ratings, with narrower DR showing an association with poorer ratings.

Conclusions: These results indicate that CI users exhibit less sensitivity to sound quality changes in music attributable to high levels of compression. These findings may account for another contributing factor to the generally poor music perception observed in CI users, particularly when listening to commercially recorded music.

Key words: Cochlear implants, Compression, Music perception, Sound quality, MUSHRA.

INTRODUCTION

Cochlear implants (CIs) provide access to auditory information for individuals with severe to profound sensorineural hearing loss. While most CI users obtain some degree of open-set speech perception, even high-performing CI users vary significantly in their report of sound quality and appreciation of music (Gfeller et al. 2008; Holden et al. 2013). Music is a complex acoustic stimulus, consisting of rhythmic, melodic, harmonic, and timbral cues. CIs are limited in their ability to transmit these cues, due to technological, biological, and acoustic features of CI-mediated listening (Limb & Roy 2014). Previous studies have shown that CI users have limitations in their ability to detect sound quality deteriorations attributable to restricted low- and high-bandpass filtering, increased reverberation, and lack of fine structure processing (Roy et al. 2012a, 2012b, 2015a, 2015b).

Compression is a ubiquitous and obligatory signal processing step in both CIs and hearing aids (HAs). Compression is used to reduce the dynamic range (DR) of incoming auditory signal to ensure that speech information is being transmitted within a patient's usable DR. While compression is intended to improve the audibility and sound quality of speech stimuli, relatively little is known about how such compression impacts musical sound quality for individuals with CIs. The perception of compression by acoustic and electric hearing is different in both range and resolution. While normal-hearing (NH) ears can perceive broad intensity changes of up to 120 dB (acoustic), they also have as many as 200 discriminable steps from the softest sound heard to the loudest sound tolerable. CI users, on the other hand, have reduced tolerable intensity range of about 10 to 20 dB (electric), with a notable decrease in the number of discriminable steps to around 20 steps (Nelson et al. 1996; Zeng et al. 1998; Fu & Shannon 1999). An additional variable for CI users is that CIs use mapping functions that attempt to replicate the loudness growth of the incoming auditory signal across the DR. It is unclear how such restricted DR and resolution influence the perception of music sound quality in CI users.

Halliwell et al. (2015) examined the acute effects of input DR compression on music enjoyment in 10 Cochlear Nucleus users. CI users compared music clips with a 1.5:1 compression ratio to the original clips for three musical genres (Jazz, Country, and Classical) and found no preference among the CI users between the altered and unaltered clips. Buyens et al. (2015, 2018) created a stereo music preprocessing scheme to optimize sound quality of music for CI listeners. Their software

Department of Otolaryngology-Head & Neck Surgery, University of California, San Francisco, California, USA.

TABLE 1. CI user demographics for tested ear

Subject	Age (yrs)	Laterality	Etiology	Duration of HL (yrs)	Duration of CI Use (yrs)	Internal Device	Electrode Array	Processing Strategy
1	53.8	R	Sudden SNHL	0.5	7.2	MED-EL Sonata	Standard	FS4-p
		L	Sudden SNHL	3.6	4.0	MED-EL Concert	FLEX28	FS4-p
2	48.9	R	Idiopathic	47.5	1.4	MED-EL Synchrony	FLEX28	FS4
		L	Idiopathic	47.2	1.6	MED-EL Synchrony	FLEX28	FS4
3	53.8	R	Idiopathic	3.8	4.0	MED-EL Concert	FLEX28	FS4-p
		L	Idiopathic	0.2	4.3	MED-EL Concert	FLEX28	FS4-p
4	71.4	R	Genetic	0.4	2.9	MED-EL Concert	FLEX28	HDCIS
		L	Genetic	14.1	6.2	MED-EL Sonata	Medium	FS4-p
5	79.4	R	Noise exposure	1.0	4.0	MED-EL Concert	FLEX28	FS4-p
6	66.8	R	Idiopathic	2.7	2.1	MED-EL Concert	FLEX28	FS4-p
7	35.8	R	Genetic	15.1	0.7	MED-EL Synchrony	FLEX28	FS4

Subjects with both left and right laterality were bilaterally implanted. "Duration of HL" refers to duration of severe to profound HL before implantation.
CI, cochlear implant; FS4 and FS4-p, fine structure processing (-p indicates parallel signal processing); HDCIS, high-definition continuous interleaved sampling; HL, hearing loss; L, left; R, right; SNHL, sensorineural hearing loss. Electrode arrays: FLEX28, 28 mm; medium, 24 mm; and standard, 31.5 mm.

utilized the timing differences for various instruments and vocals inherent in stereo recordings to modify the audio mix (similar to how vocals are identified and attenuated for karaoke music). In testing their strategy on CI users, they found that the more complex the music was, the more the CI users preferred to enhance the drums, bass, and vocals of the music (and attenuate the other components). In 2018, they gave 12 CI users a take-home device to experiment with the strategy while listening for longer lengths of time to a wider variety of music, and they found that all participants preferred to attenuate some of the aspects of (i.e., make less complex) the music.

In addition to its role in CI sound processing, compression is also a common feature of modern studio recording techniques. DR compression may serve several purposes within a musical track. During the recording process, each instrument or vocal track is often recorded with compression to control for the large variation in DRs inherent in different instruments and voices. An audio engineer may then apply various levels of additional compression to each individual track or the overall mix to create the desired balance of overall sound. It is generally accepted that popular music released today has more aggressive levels of compression incorporated into the final product than the popular music of a few decades ago; indeed, there is a correlation between the overall loudness of a musical track and increased record sales (Vickers 2011). While there has been a fair amount of research published on the effect of compression settings on music perception in HA users, little is known about how compression affects perception of musical sound quality in CI users (Kirchberger & Russo 2016).

In this experiment, we used the Cochlear Implant-Multiple Stimulus with Hidden Reference and Anchor (CI-MUSHRA) approach to analyze the ability of CI users to detect changes in musical sound quality attributable to increasing levels of DR compression. The CI-MUSHRA approach allows the user to rate the sound quality of modified sound clips in comparison to an unmodified Reference version of the same sound clip, thereby minimizing the impact of preference, genre, or recording quality differences from one stimulus to the next. The primary utility of this approach is that parametric manipulations of a sound can be made to assess the subjective change in sound quality, relative to the unaltered sound (Roy et al. 2012a, 2012b, 2015a, 2015b). This approach minimizes some of the potential drawbacks of subjective rating scales and questionnaire-based approaches and

allows objective measurement of ability to detect deteriorations/manipulations in sound quality. We hypothesized that NH users would detect increased levels of compression more easily than CI users, but that both groups would perceive a loss of sound quality with increasing compression levels.

MATERIALS AND METHODS

Participants

Seven MED-EL CI recipients (58.6 ± 14.8 years old; 4 males and 3 females) and 10 NH listeners (32.4 ± 14.4 years old; 5 males and 5 females) participated in the study. Of the 7 CI recipients, 4 were bilaterally implanted and 3 were unilaterally implanted, yielding 11 CI ears tested. Six CI recipients were postlingually deafened and 1 was prelingually deafened; all used oral/aural communication as their primary method of communication. The mean duration of implant use was 3.5 ± 2.0 years. The CI users utilized a variety of devices and processing strategies, as shown in Table 1.

All participants completed a questionnaire on their musical training background, which included their total years of formal music instruction, the setting of the training (e.g., one-on-one, self-taught, classroom, etc.), specific instruments played, and age at which they began training, and how many hours per week they currently listen to music. Four members of the CI group and seven of the NH group reported formal musical training. The duration of formal musical training for the CI group was 5.2 ± 5.9 years and for the NH group was 5.9 ± 6.5 years. A single-factor analysis of variance between the CI and NH groups showed no significant difference in number of years of formal musical training ($F(1, 17) = 0.049$, $p = 0.83$). There was also no significant difference on a single factor analysis of variance between the CI and NH groups for number of hours per week they currently listen to music ($F(1, 17) = 4.234$, $p = 0.06$); however, it is likely not significant due to the small sample size. Three of the 7 CI subjects and none of the 10 NH subjects reported not listening to music in general (0 hr per week).

This study was approved by the Institutional Review Board at University of California, San Francisco (UCSF). Informed consent was obtained for all participants. Subjects were recruited via flyers posted in the CI Center and on bulletin boards at the Parnassus campus of UCSF. Subjects were also drawn from a

database, maintained by the UCSF Sound and Music Perception Lab, of persons who had expressed interest in being contacted to participate in research studies.

Musical Stimuli

Previous efforts by this lab have generated a corpus of 25 real-world music clips. The clips are 5 sec in length with five clips from each of the five common musical genres (Classical, Jazz, Country, Rock, and Hip-Hop). Within each genre, three clips were considered well known and two clips were lesser known. See Roy et al. (2012a) for additional details regarding music clip selection. In the present study, Adobe Audition 3 (San Jose, CA, USA) was used to apply multiple iterations of an aggressive compression algorithm to the music clips. Compression was applied with an input to output ratio of 10:1, attack time of 40 msec, release time of 150 msec, and look-ahead of 3 msec. To alter the amount of compression, without changing the compression parameters, we applied an iterative process to the music clips, whereby the same compression algorithm was applied multiple times. The test conditions included 1, 3, 5, and 20 iterations; the 20-iteration samples served as the Anchor. The original music clip without additional compression applied served as the Reference. All clips were balanced for loudness using RMS normalization. A 250 msec fade-in and fade-out were applied to all clips to minimize click artifacts. For samples of rock and jazz music clips with progressively more compression applied (Reference, 1, 3, 5, and 20 iterations), see Rock-MusicSamples in Supplemental Digital Content 1, http://links.lww.com/EANDH/A524, and JazzMusicSamples in Supplemental Digital Content 2, http://links.lww.com/EANDH/A525.

The DR of the original music clips (Reference stimuli) are shown in Figure 1 in the form of intensity-level distribution functions. Figure 1 illustrates the variation of intensity level by frequency and further shows that, on average, the Hip-Hop (39.9 dB) and Jazz (34.9 dB) samples we used for this study have a larger DR than do Rock (27.1 dB) and Country (29.0 dB), with Classical (32.4 dB) falling in between the other genres. It should be noted that the corpus of music clips in this experiment was limited to 25, which were only 5 sec long, and may not be generalizable to each genre as a whole.

Test Paradigm

The CI-MUSHRA is a tool that can be used to assess relative changes in the perceived sound quality of music across increasingly degraded listening conditions and to compare the response pattern of CI users to NH controls. Participants were asked to rate the relative sound quality of five randomized versions of a particular music clip on a sliding scale with the following categorical labels: "Much better," "Slightly better," "Same as reference," "Slightly worse," and "Much worse" (see Fig. 2). The sliding scale corresponded to numbers from 0 to 200; scores from 0 to 99 represented sound quality worse than the Reference, a score of 100 represented sound quality equal to the Reference stimulus, and scores from 101 to 200 represented sound quality better than the Reference stimulus. Although an Anchor (extremely poor quality) stimulus was included to encourage the use of the entire rating range, the subjects were not explicitly required to use the whole range of the sliding scale, nor were they informed whether the stimuli were expected to sound better or worse than the Reference.

Subjects were given verbal instructions on how to complete the MUSHRA task, along with a live demonstration of the software interface functionality on the computer and then required to complete a practice round, supervised by the researcher. The instructions were as follows: (1) Click "Play Reference" to listen to the Reference music clip. You may replay the Reference as often as you wish. (2) Listen to Sound A and move the slider up or down to indicate how Sound A compares to the Reference. (3) Repeat steps (1) and (2) for Sounds B–E. (4) At least one clip must be ranked "Same as Reference" because one of the clips is the Reference. (5) Click "Save and proceed" when you are satisfied with your rankings of Sounds A–E. There are 25 trials in this experiment.

During the practice round, the subjects were required to complete the MUSHRA task for at least three sets of stimuli but were allowed to do additional sets until they reported that they were comfortable using the computer and interacting with the software interface. The experiment took on average 30 to 60 min per subject. Despite the age difference between the CI and NH groups, there did not seem to be a gross difference between the CI and NH groups for difficulty interacting with the software or time needed to complete the assessment.

Test Conditions

Testing took place in a calibrated sound booth. The stimuli were presented from a speaker located at 0° azimuth at an average RMS level of 65 dBA. All CI recipients used a sound processor designated for research use only, programmed with a modified version of their preferred everyday fitting. To increase consistency of audibility of soft, average, and loud sounds across test subjects, their preferred everyday fittings were modified as follows: upper stimulation levels (MCLs) were loudness-balanced at 80%, MCLs were swept at 100% to ensure comfort, lower stimulation levels (THRs) were measured on each electrode and set at highest inaudible level, maplaw was set to 1000, and automatic gain control was set to 3:1. There were no predetermined restrictions imposed upon the patients' electrical DR (the difference between MCL and THR). Volume control was fixed at 100% and sensitivity at 75%. Microphone directionality and wind noise reduction features were disabled, when applicable. The frequency allocation and stimulation strategy were not altered to minimize acclimatization effects. CI users were given about 30 min to adapt to the relatively minor changes that were made.

NH participants listened with both ears. CI users listened with the test ear only. If the nontest ear had any residual hearing, a foam earplug was used. In one case, a CI recipient had a Lyric HA in the nontest ear that was situated deep in the ear canal and could not be removed for testing. The Lyric was therefore turned off and functioned as an earplug. All CI users were asked if they could hear the test stimuli with the nontest ear, while the earplug was in place and the CI on their test ear was turned off; all declined being able to hear the stimuli with the nontest ear. Participants completed the CI-MUSHRA using a Dell Latitude E7470. Subjects were given the option of using a touch screen, mouse, or touchpad. The computer ran the CI-MUSHRA program with MATLAB software version R2012b (Mathworks, Natick, MA, USA).

Statistical Analysis

We first examined the potential associations of hearing status (NH or CI), level of compression, years of musical training,
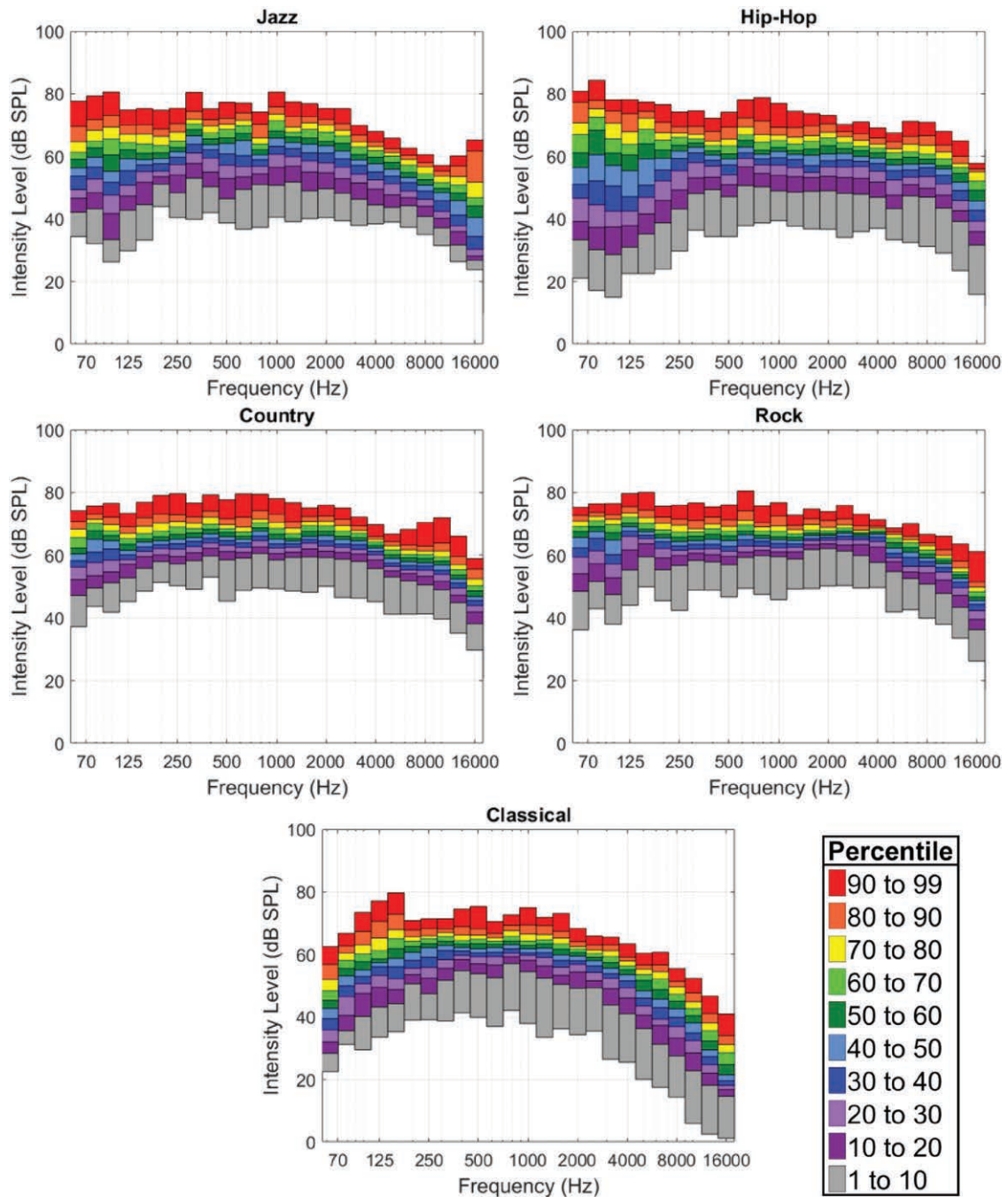
Fig. 1. Intensity level distribution functions for the Reference music stimuli. Each graph depicts the distribution of the dynamic range (*y* axis), in 1/3 octave frequency bands (*x* axis), for the five music clips belonging to each of the five musical genres tested. Each vertical column is partitioned into 10 percentiles (see legend). The dynamic range is greatest for Jazz and Hip-Hop, least for Rock and Country, with Classical falling in between the two groups.

DR, and musical genre individually with discriminatory ability in univariate regression models. Generalized estimating equations techniques were used, because, in the case of the CI group, data were analyzed at the level of the individual ear, and hearing would likely be correlated between two ears in the same person. Generalized estimating equation is a repeated measures analysis which accounts for the within-person correlation created when multiple measurements are made on the same person. It adjusts SEs measured on between-person variability for this additional within-person variability, then averages over all subjects and calculates appropriately adjusted estimates. We then

ran multivariable models in which we tested for interactions between musical genre and hearing group, and the level of compression and hearing group. Results from the statistical analysis are shown in Table 2. In the models, "hearing group" refers to whether participants were in the NH or CI group.

For DR analysis, music clips were grouped by genre and analyzed with Matlab to generate spectral power level distribution functions, expressed as percentiles (Fig. 1). This analysis was plotted using vertical subdivisions that represent power within a given 1/3 octave band, using a 125 msec window (50% overlap), similar to the approach used by Holube et al. (2010) in
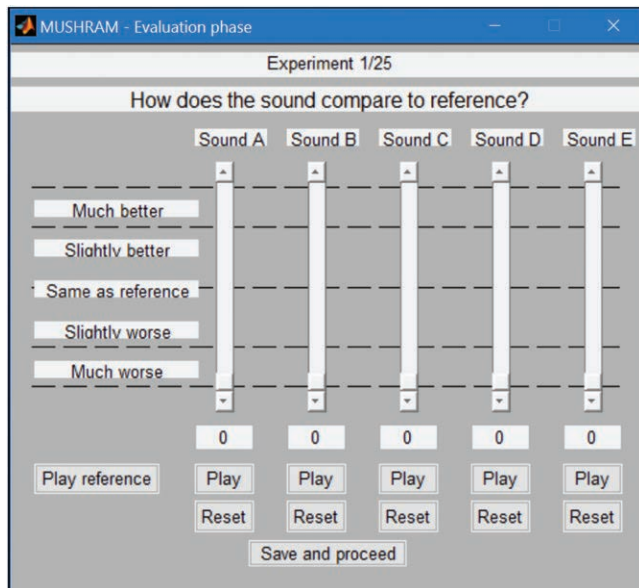
Fig. 2. Screenshot of the CI-MUSHRA software user interface for subject testing. The "Play reference" button allowed the subject to listen to the Reference (unaltered stimulus) and was freely available for comparison. Sounds A–E (columns) represent each of the five randomized music clips provided in a trial, and the "Play" button allows the user to listen to each stimulus as many times and in any order they prefer. The "Reset" button will snap the slider to the middle of the scale, labeled as "Same as reference."

their development and analysis of the International Speech Test Signal. Speech is frequently analyzed using a similar approach known as Long-term Average Speech Spectrum (LTASS). In this approach, DR is examined according to frequency bands. In most common usages, however, the Long-term Average Speech Spectrum graphs usually do not include the spectral power below the 30th percentile, because these measurements are often attributed to the noise floor of the microphone setting and recording setting (Byrne et al. 1994; Holube et al. 2010). For these stimuli, we chose to display the full-power spectrum because speech in quiet is a drastically different stimulus than music, speech and music may be recorded in different settings, and it is unclear that what portion of the DR is attributable to noise floor and recording equipment in commercially produced recordings.

RESULTS

We identified several important findings from our data analysis. Hearing group, level of compression, DR, and musical genre were each significantly associated with discriminatory ability of sound quality in univariate analysis, although overall years of musical training were not significant. The CI group was found to be significantly less sensitive to the effects of compression on musical sound quality than the NH group (Table 2, Model 1, $p < 0.001$); discriminatory ability overall for the two groups combined exhibited a dose–response effect, with more compressed music clips being increasingly associated with poorer sound quality (Table 2, Model 2, $p < 0.001$ for each level when compared to the Reference with no compression).

Overall, participants' responses differed based on musical genre (Table 2, Model 4). The sound quality rankings were averaged for each genre separately, revealing that the music clips

belonging to the Classical genre were affected least by the additional iterations of compression. When compared to Classical, the Rock genre was not rated significantly different ($p = 0.08$), but Country, Jazz, and Hip-Hop were rated slightly lower in quality ($p < 0.001$). Combining these predictors in multivariable models fitted with interaction terms, we found a significant interaction between hearing group and level of compression, but no significant difference between the CI and NH groups was found with regard to their responses to various genres (Table 2, Model 5, $p > 0.05$).

Our final model therefore included hearing group, level of compression, and an interaction term between these two. Because a significant interaction indicates that the relationship between compression level and ability to discriminate differences in sound quality is different between the two groups, estimates from the multivariable model are reported with group held constant, and compression level varied (Table 2, Model 3). We found that, because the levels of compression increased, each hearing group was able to discern the degradation of quality in the sound; however, the CI group had more modest levels of discriminatory ability at each compression level than those with normal hearing.

Figure 3 shows sound quality ratings for both NH and CI users according to level of compression, relative to the uncompressed Reference clip. The shaded regions around each line indicate their respective 95% confidence intervals provided by the statistical analysis. While all subjects on average rated the more highly compressed music clips as poorer in sound quality, the NH group demonstrated greater discriminatory ability than the CI group at each level of compression.

We examined whether the results were driven primarily by the DR of music clips and found that while DR was significantly associated with sound quality (Table 2, Model 6, $p = 0.005$), if the multivariable model was adjusted for DR, the results reported above did not change (Table 2, Model 7). The effect of DR is very small compared with the other effects (namely, level of compression and hearing group).

Finally, we further explored the potential correlation between sound quality ratings and years of formal musical training using stratified regression models, because this variable was not significant at the 0.05 level in univariate analysis. A history of formal musical training did appear to benefit the NH group in their ability to discriminate between the levels of compression (Table 2, Model 8a, $p = 0.003$). Musical training did not, however, have a significant effect for the CI group (Model 8b, $p = 0.45$).

DISCUSSION

In this study, we examined the ability of CI users and NH controls to detect changes in musical sound quality attributable to increasing levels of DR compression. We confirmed our hypothesis that CI users would exhibit less sensitivity to musical quality degradation introduced by increased levels of compression. These findings are consistent with prior work on other aspects of musical sound quality degradation, such as band-pass filtering, reverberation, etc. (Roy et al. 2012a, 2012b, 2015a, 2015b) and also similar to observations in HA users, who also demonstrate subjective impairments in sound quality (Chasin 2012). In addition, the CI users were more likely to rate the most degraded music clips as only "Slightly worse than reference" or

TABLE 2. Regression models utilizing generalized estimating equations techniques

Characteristic	Coefficient	(95% Confidence Interval)	p
Model 1: hearing group only			
NH vs. CI	20.6	±8.6	<0.001*
Model 2: level of compression only			
Reference	Ref		
1×	−9.56	±3.78	<0.001*
3×	−24.4	±6.8	<0.001*
5×	−39.8	±8.4	<0.001*
20×	−66.0	±12.4	<0.001*
Model 3: interaction between level of compression and hearing group			
NH			
Reference	Ref		
Difference between Ref and level 1×	−9.72	±5.90	0.001*
Difference between Ref and level 3×	−29.8	±9.1	<0.001*
Difference between Ref and level 5×	−49.4	±8.2	<0.001*
Difference between Ref and level 20×	−87.0	±6.6	<0.001*
CI			
Reference	Ref		
Difference between Ref and level 1×	−9.40	±4.85	<0.001*
Difference between Ref and level 3×	−19.6	±9.1	<0.001*
Difference between Ref and level 5×	−31.0	±12.5	<0.001*
Difference between Ref and level 20×	−47.0	±15.3	<0.001*
Model 4: musical genre only			
Classical	Ref		
Rock	−3.80	±4.29	0.08
Country	−7.06	±2.80	<0.001*
Jazz	−12.0	±4.7	<0.001*
Hip-Hop	−14.0	±3.8	<0.001*
Model 5: interaction between musical genre and hearing group			
This interaction is not significant			
Model 6: dynamic range only			
Dynamic range	0.35	±0.24	0.005*
Dynamic range (in multivariable model below)	0.29	±0.21	0.006*
Model 7: hearing group (interaction between hearing group and level of compression, adjusted for dynamic range)			
NH			
Reference	Ref		
Difference between Ref and level 1×	−10.8	±6.05	<0.001*
Difference between Ref and level 3×	−30.8	±9.29	<0.001*
Difference between Ref and level 5×	−50.4	±8.32	<0.001*
Difference between Ref and level 20×	−88.0	±6.77	<0.001*
CI			
Reference	Ref		
Difference between Ref and level 1×	−10.5	±4.65	<0.001*
Difference between Ref and level 3×	−20.7	±9.10	<0.001*
Difference between Ref and level 5×	−32.1	±12.3	<0.001*
Difference between Ref and level 20×	−47.9	±15.1	<0.001*
Model 8a: formal musical training			
NH	−0.88	±0.58	0.003*
Model 8b: formal musical training			
CI	−0.40	±1.04	0.45

Statistical analysis provided by biostatistician.
*Significant difference (p < 0.05).
CI, cochlear implant; NH, normal hearing.

"Same as reference." These findings are consistent with prior work showing high variability in CI users' responses to music.

Taken together, the results from this study raise important questions about the use of compression in CI-mediated listening, particularly with respect to music, and suggest that the implementation of acoustic compression to electrical stimulation of the cochlea remains poorly understood. In this study, we found that CI users are relatively insensitive to changes in auditory stimuli with varying degrees of compression, requiring considerably higher compression levels to detect a change in sound quality than NH subjects. Our CI subjects used their standard clinical processing strategies during their testing, which includes obligatory compression that is applied early on during sound processing. One potential implication of this study

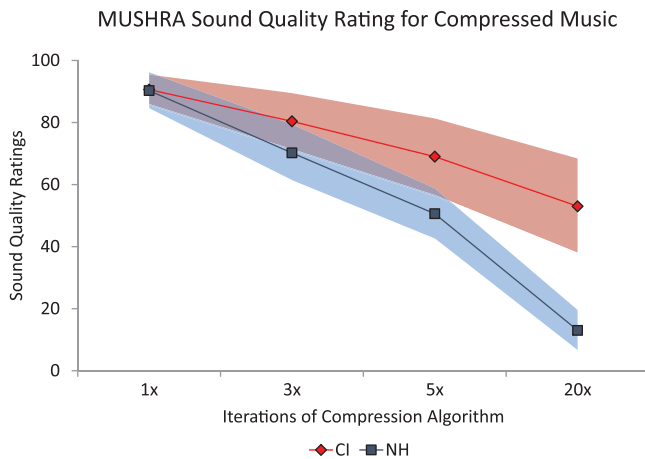## MUSHRA Sound Quality Rating for Compressed Music

Fig. 3. Sound quality rankings by cochlear implant (CI) and normal-hearing (NH) listeners to music clips that have undergone progressively more aggressive compression. The x axis indicates how many iterations of compression were applied to the Reference clip. The y axis indicates the CI or NH listeners' average rankings of the music clip as compared to their ranking of the Reference (100). Shading above and below lines represents 95% confidence intervals.

is that there may be too much compression applied to musical signals for CI users, as evinced by a relative loss of ability to discriminate changes directly attributable to compression. If so, this study raises fundamental concerns regarding the most appropriate compression settings for electrical listening, or, more likely, whether the parameters and settings for the compression algorithms commonly used for speech should be altered for music. It is plausible that variable implementation of front-end compression settings for music and for speech might constitute a viable option. Future directions of inquiry may include investigating the compression settings in the CI, accessible via the programming software, to allow for improved sound quality and music appreciation.

One of the guiding principles in eventual optimization of CIs for music is to identify limitations in the present CI processing that contribute to impairments in music perception. For example, in the previous studies of low-frequency perception, we found that CI users had a relatively poor ability to identify changes in musical sound quality attributable to low frequency removal, implying that they do not hear bass frequencies well. Hence, frequencies that improve perception of musical bass frequencies may help normalize sound quality assessments relative to NH counterparts (Roy et al. 2012a). In a similar fashion, characterizing the degree to which compression impacts (or fails to impact) sound quality ratings for CI users will allow us to have an assessment of the degree to which compression helps or hurts musical sound quality perception in relation to normal hearing. Through studies such as the present one, we hope to gain information that may help us determine the ideal amount of audio compression delivered to CI users in order for assessments of sound quality to mirror those of NH counterparts. These kinds of data will eventually contribute to a better understanding of how to optimize compression device settings for music, while subsequent research could enable prescriptive changes in CI device programming.

To examine these results with more granularity, we conducted analyses of DR and musical genre. Our results suggest two important findings: first, the DR of our music clips differs

as a function of musical genre, and second, CI-MUSHRA performance ratings for compression (in both NH and CI subjects) are related to the DR of the auditory stimulus. We found that Hip-Hop and Jazz genres had the greatest DR, whereas Rock and Country music had the least DR in the excerpts we used here. Classical excerpts fell in between these two ends, with an intermediate DR. These differences are a product of a number of factors, including the nature of the instruments and sounds utilized in each genre, the recording techniques used to capture the original source material, and also the use of additional studio compression during both original tracking but also mixing and mastering. These differences are primarily technical in nature, as one could easily alter the DR of any excerpt of music simply by using different recording/studio processing techniques.

For our CI users, however, the fact that Rock and Country were the most heavily compressed genres (i.e., had the narrowest DR) before the application of any additional compression, may explain, the finding that they showed less sensitivity to compression of these genres. Similarly, CI users displayed the greatest ability to detect changes in compression applied to source material that began with a greater DR (Jazz and Hip-Hop). In addition, our results suggest that DR (independent of musical genre) is statistically correlated to sound quality ratings on the CI-MUSHRA for both subject groups. Although this effect was small in magnitude, particularly in relation to hearing status and level of compression, it does suggest that sound quality ratings in CI users for different compression settings are affected by multiple factors. While it is not at all surprising that the ability to detect compression levels is easiest when DR is greatest, and more difficult when DR is reduced, this finding does shed light on why our subjects performed the way they did for the genres used here. Furthermore, these findings may also carry implications for when compression may be more or less helpful in musical contexts.

This study has several limitations that should be noted. First, we used music clips that were 5 sec in duration, to allow for a sufficient quantity of clips to be auditioned in a reasonable amount of time. Clips of longer duration, particularly for genres where there is considerable fluctuation of DR depending on the point in the piece (e.g., classical passages where a few solo instruments are highlighted that transition to a full orchestra), might allow for a more real-world understanding of how compression effects are perceived not only by genre or by amount, but how they evolve over time. Because the music clips included in the CI-MUSHRA task are 5 sec in length, they do not encapsulate the full breadth of each genre, and, as such, the results of the task should be interpreted with caution. In addition, because we used real-world musical excerpts from commercial recordings, these all had different acoustic properties as a function of studio recording/processing that we could not control or necessarily even identify. Therefore, the true implications of compression applied to music that is played on the radio or produced as a final commercial track require further study. Another limitation of this study was the relatively small sample size of CI users. While the results were statistically significant, they should be interpreted with caution because they may not be representative of all CI users' experience.

CONCLUSIONS

This study shows that CI users display limited ability to detect the effects of additional compression on musical stimuli

in comparison to NH listeners. Because compression remains an obligatory processing step for both CIs and HAs, further research on how this important step impacts music perception is needed. In addition, further studies are needed to determine whether alterations to CI user programs that accentuate or reduce the degree of compression applied to musical sounds might lead to improved sound quality for musical stimuli. Finally, the impact of compression on bilateral listening or bimodal listening strategies was not examined here and would also constitute important areas for future research.

REFERENCES

Buyens, W., van Dijk, B., Moonen, M., et al. (2018). Evaluation of a stereo music preprocessing scheme for cochlear implant users. *J Am Acad Audiol, 29*, 35–43.

Buyens, W., van Dijk, B., Wouters, J., et al. (2015). A stereo music preprocessing scheme for cochlear implant users. *IEEE Trans Biomed Eng, 62*, 2434–2442.

Byrne D., Dillon H., Tran K., et al. (1994). An international comparison of long-term average speech spectra. *J Acoust Soc Am, 96*, 2108–2120.

Caldwell, M. T., Jiam, N. T., Limb, C. J. (2017). Assessment and improvement of sound quality in cochlear implant users. *Laryngoscope Investig Otolaryngol, 2*, 119–124.

Chasin, M. (2012). Music and hearing aids–an introduction. *Trends Amplif, 16*, 136–139.

Fu, Q. J., & Shannon, R. V. (1999). Effect of acoustic dynamic range on phoneme recognition in quiet and noise by cochlear implant users. *J Acoust Soc Am, 106*, L65–L70.

Gfeller, K., Oleson, J., Knutson, J. F., et al. (2008). Multivariate predictors of music perception and appraisal by adult cochlear implant users. *J Am Acad Audiol, 19*, 120–134.

Halliwell, E. R., Jones, L. L., Fraser, M., et al. (2015). Effect of input compression and input frequency response on music perception in cochlear implant users. *Int J Audiol, 54*, 401–407.

Holube, I., Fredelake, S., Vlaming, M., et al. (2010). Development and analysis of an International Speech Test Signal (ISTS). *Int J Audiol, 49*, 891–903.

Holden, L. K., Finley, C. C., Firszt, J. B., et al. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear, 34*, 342–360.

Kirchberger, M., & Russo, F. A. (2016). Dynamic range across music genres and the perception of dynamic compression in hearing-impaired listeners. *Trends Hear 20*, 2331216516630549.

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

Nelson, D. A., Schmitz, J. L., Donaldson, G. S., et al. (1996). Intensity discrimination as a function of stimulus level with electric stimulation. *J Acoust Soc Am, 100*(4 Pt 1), 2393–2414.

Roy, A. T., Carver, C., Jiradejvong, P., et al. (2015a). Musical sound quality in cochlear implant users: A comparison in bass frequency perception between fine structure processing and high-definition continuous interleaved sampling strategies. *Ear Hear, 36*, 582–590.

Roy, A. T., Jiradejvong, P., Carver, C., et al. (2012a). Assessment of sound quality perception in cochlear implant users during music listening. *Otol Neurotol, 33*, 319–327.

Roy, A. T., Jiradejvong, P., Carver, C., et al. (2012b). Musical sound quality impairments in cochlear implant (CI) users as a function of limited high-frequency perception. *Trends Amplif, 16*, 191–200.

Roy, A. T., Vigeant, M., Munjal, T., et al. (2015b). Reverberation negatively impacts musical sound quality for cochlear implant users. *Cochlear Implants Int, 16* (Suppl 3), S105–S113.

Vickers, E. (2011). The loudness war: Do louder, hypercompressed recordings sell better? *J Audio Eng Soc, 59*, 346–351.

Zeng, F. G., Galvin, J. J. 3rd, Zhang, C. (1998). Encoding loudness by electric stimulation of the auditory nerve. *Neuroreport, 9*, 1845–1848.