Published in final edited form as:

Cochlear Implants Int. 2015 September ; 16(0 3): S22–S31. doi:10.1179/1467010015Z.000000000269.

A preliminary report of music-based training for adult cochlear implant users: rationales and development

Kate Gfeller,

Iowa Cochlear Implant Research Center, School of Music, Department of Communication Sciences and Disorders, The University of Iowa

Emily Guthe,

School of Music, The University of Iowa

Virginia Driscoll, and

Iowa Cochlear Implant Research Center, Department of Otolaryngology, The University of Iowa

Carolyn J. Brown

Department of Communication Sciences and Disorders, The University of Iowa

Abstract

Objective—This paper provides a preliminary report of a music-based training program for adult cochlear implant (CI) recipients. Included in this report are descriptions of the rationale for music-based training, factors influencing program development, and the resulting program components.

Methods—Prior studies describing experience-based plasticity in response to music training, auditory training for persons with hearing impairment, and music training for cochlear implant recipients were reviewed. These sources revealed rationales for using music to enhance speech, factors associated with successful auditory training, relevant aspects of electric hearing and music perception, and extant evidence regarding limitations and advantages associated with parameters for music training with CI users. This information formed the development of a computer-based music training program designed specifically for adult CI users.

Results—Principles and parameters for perceptual training of music, such as stimulus choice, rehabilitation approach, and motivational concerns were developed in relation to the unique auditory characteristics of adults with electric hearing. An outline of the resulting program components and the outcome measures for evaluating program effectiveness are presented.

Conclusions—Music training can enhance the perceptual accuracy of music, but is also hypothesized to enhance several features of speech with similar processing requirements as music (e.g., pitch and timbre). However, additional evaluation of specific training parameters and the impact of music-based training on speech perception of CI users are required.

Keywords

cochlear implants; music training; experience-based plasticity

Music Training and Experience-Based Neural Plasticity: Why Use Music?

Over the past few decades, there has been a burgeoning interest in the relationship between music training and experience-based neural plasticity, particularly for the auditory system. As people engage with music, a widespread bilateral network of brain regions is activated (frontal, temporal, parietal, subcortical) (for reviews, see Herholz & Zatorre, 2012; Patel, 2011). Music listening involves on-going encoding, organization and recall of multidimensional patterns of pitch, timbre, rhythm, and amplitude, and the segregation of one pattern from another (e.g., melody vs. harmony, contrasting patterns within counterpoint). The heightened fine-grained frequency discrimination required in response to ongoing changes in acoustic parameters has been attributed with improved auditory working memory, attention, and more rapid spectro-temporal processes at various levels of the auditory system (Patel, 2011).

Performing music (as opposed to only listening) engages and integrates multiple sensory and motor systems specific to the type of action (e.g., playing a keyboard vs. a violin), and makes demands on a variety of higher-order cognitive processes (see review by Herholz & Zatorre, 2012). Furthermore, music involvement can be associated with arousal, reward, positive mood, and social factors; these can promote attentive, sustained engagement (Herholz & Zatorre, 2012; Patel, 2011).

Mounting evidence, gathered primarily from normal hearing (NH) listeners, suggests overlap in brain networks that process acoustic features heard in both music and speech. This premise contributes to the supposition that musical training may generalize to neural encoding of speech as well as music (e.g., Besson et al., 2007; Herholz & Zatorre, 2012; Kraus et al., 2009; Patel, 2011; Shahin, 2011). The higher perceptual demands required for music listening, or what Patel (2011) refers to as greater precision, may ‘fine-tune’ the auditory system (Herholz & Zatorre, 2012; Invalson & Wong, 2013; Patel, 2011). Research suggests that this fine tuning may generalize to skills such as phonological processing, verbal memory, learning mechanisms for language, and lower perceptual thresholds for complex auditory input (Herholz & Zatorre, 2012; Kraus et al., 2009; Musacchia et al., 2007; Patel, 2011; Shahin, 2011; Wong et al., 2007). It has been hypothesized that neural changes will be more robust and that music training is more likely to generalize to speech when perception of the signal input requires precision of auditory processing, is behaviorally relevant and rewarding, and the task is actively trained (Fu & Galvin, 2007; Herholz & Zatorre, 2012; Patel, 2011).

In our own laboratory, we have recently examined the impact of long-term musical training on perceptual accuracy of several types of music and speech stimuli (Brown et al., 2014). We compared the perceptual capabilities of normal hearing (NH) adults who were either professional musicians (n=10, mean age=42.2 yrs.; completed at least 4 years as a college music major) or adults (n=10; mean age=38.4 yrs.) with little or no musical training (no music classes above elementary school). Participants were tested on pitch discrimination, timbre recognition, the *AZ-Bio* in noise condition (Spahr & Dorman, 2004), and spectral ripple discrimination (Henry & Turner, 2003), which provides a measure of spectral

resolution considered predictive of speech perception for spectrally-complex aspects of speech.

Consistent with prior studies regarding music training, the professional musicians demonstrated superior perceptual accuracy for the following measures: complex pitch discrimination ($p=.002$), pure-tone difference limens at 200 Hz ($p=.02$) and 1600 Hz ($p=.001$), and superior timbre recognition ($p=.01$). Relevant to the question of generalization to speech perception, the musicians also exhibited superior accuracy on the *AZ Bio* at 0 SNR ($p=.046$) and spectral ripple discrimination thresholds ($p=.002$). Interestingly, the musician group, while having near-normal hearing, did have greater incidences of mild hearing loss at 4K or higher; nevertheless, their perceptual accuracy as a group was greater for several measures of spectrally complex sounds than the NH group with better hearing acuity (Brown et al., 2014).

Studies linking music training with linguistic enhancements have sparked interest in music as a potential tool in (re)habilitation for persons with communication disorders, including persons with cochlear implants (CI) (e.g., Ingvalson & Wong, 2013; Kraus & Skoe, 2009). CI users, as a group, perform less accurately than NH listeners on perception of spectrally-complex features of speech, such as lexical tones, linguistic or affective prosody, and speech in background noise (Gfeller et al., 2007; See et al., 2013). Because of the hypothesized overlap in neural networks that process speech and music, and because music tends to be associated with positive mood and reward (which may help with motivation and persistence) (Patel, 2011), there has been speculation that music training may be a valuable companion to more conventional forms of speech rehabilitation (Chermak, 2010; Ingvalson & Wong, 2013; Kraus & Skoe, 2009). However, to date, little research has addressed specifically whether music-based training can transfer to speech perception tasks in CI users (Shahin, 2011). The following section addresses some factors that may influence whether music-based training can benefit adult CI recipients.

Adoption of Music Training for Cochlear Implant Recipients

As we consider the various rationales for music-based training for CI recipients, it is important to recall that the majority of studies regarding music training have focused on NH people with healthy auditory systems and long-term music training, sometimes commencing at an early age and persisting well into adulthood. Training may have included many years of lessons, participation in musical ensembles, and formal ear training (music theory). This is relevant to this discussion, because different forms of musical experience (e.g., listening vs. instrumental music training; years of exposure to one's own musical instrument, etc.) are associated with different types of changes in the auditory system. For example, some studies have shown modifications as specific as enlarged auditory cortical evoked potentials in musicians in response to the sound of their own musical instrument. In addition, music training that includes multimodal interactions (e.g., combined auditory and motor processes) has been associated with outcomes different from music listening (Herholz & Zatorre, 2012).

The neural changes that occur as a result of the short-term discrimination training and comprising music listening (which is arguably more typical in rehabilitation) are likely to differ from changes associated with music training over many years. The specific adaptations within the auditory system resulting from short-term training will differ depending upon the stimuli parameters, task difficulty, attention, or other non-specific factors that accompany the learning, and whether adaptations have behavioral relevance, and thus transfer to other tasks (Herholz & Zatorre, 2012). Research studies to date have employed diverse combinations of musical sounds combined with different forms of musical engagement (e.g., listening, playing different instruments, singing, etc.), varying format parameters, onset and length of training, and have assessed a variety of cognitive-linguistic sub-skills (Herholz & Zatorre, 2012; Patel, 2011; Shahin, 2011). Thus, it is difficult to interpret trends, and predict what particular forms of music training are more likely to generalize to enhanced speech, even for persons with typical auditory systems.

Now, let us consider some additional factors relevant to music training of adult CI recipients. From the standpoint of *duration* of training, only a small proportion of CI users are likely to have experienced early (commencing at a young age) and extensive music training equivalent to that of professional musicians (Gfeller, 2001). Enrollment in extensive music training after implantation is not realistic for many CI users, nor can one 're-instate' the neural plasticity associated with childhood (i.e., adult CI users may respond to training differently from pediatric CI users). With regard to the *type* of training, adult CI users are more likely to engage in readily accessible and affordable computer-assisted listening exercises, as opposed to the years of lessons and musical ensembles typical for professional musicians.

From the standpoint of the *acoustic signal* that will drive neural changes, normal hearing listeners can discriminate remarkably small variations in pitch, timbre, and dynamic range. In contrast, CI recipients have damaged auditory systems that undermine pitch and spectral resolution, and the acoustic representation of the musical signal conveyed by the cochlear implant is a crude representation of that heard through a healthy hearing mechanism (Limb & Roy, 2014; Looi, Gfeller & Driscoll, 2012). The impact of electric hearing on key features of music is the focus of the following section.

Music Perception and Electric Hearing

As we consider the use of music in rehabilitation for CI users, it is important to consider limitations with which the rich and complex elements of music are encoded by the CI (for reviews, see Looi, Gfeller & Driscoll, 2012; Limb & Roy, 2014). Present-day CI processing strategies usually remove the temporal fine-structure information in the stimulus waveforms and preserve the temporal envelopes extracted from 6 to 22 frequency bands; these are conveyed via the electrodes in the internal array. The electrode array has a small number of wide bandpass filters with fixed center frequencies, resulting in coarse spectral cues, and thus poor frequency resolution. This degraded input is sufficient for conveying speech perception in quiet as well as the rhythmic components of music. However, CIs are poorly suited for transmitting greater fine structure required for perception of pitch and timbre (Limb & Roy, 2014; Looi, Gfeller & Driscoll, 2012).

Pitch, has been described as the most basic organizing structure in most musical cultures (Patel, 2007). The direction and exact magnitude of sequential or concurrent pitch relations comprise melodies and harmonies. Thus, the poor spectral resolution conveyed through the CI, which undermines pitch resolution, contributes to poor perception of melody and harmony by CI users (Looi, Gfeller & Driscoll, 2012; Gfeller, 2001).

Timbre, which comprises the unique onset transients, steady state, and decay of acoustic energy of specific musical sources (instruments or voices), is important in identification of musical instruments (e.g., a flute vs. a piano) or singers (e.g., Bob Dylan vs. Taylor Swift). Because most people listen to music for enjoyment, the sound quality which results from spectral elements is as important as timbre recognition. Research on timbre perception indicates that CI recipients have significantly poorer outcomes for both instrument identification as well as ratings of sound quality, though some musical instruments are more readily recognized and also have more acceptable sound quality (Gfeller et al., 2008, 2010, Limb & Roy, 2014; Looi, Gfeller, & Driscoll, 2012).

Interestingly, despite the technical limitations of the CI in transmission of pitch and timbre, there is considerable variability among CI users in perceptual accuracy of musical structures such as pitch, melody, harmony, and timbre (Gfeller et al., 2008, 2010; Looi, Gfeller & Driscoll, 2012). Furthermore, intra-individual differences occur as a result of the frequency range and spectral characteristics of the stimuli (Vandali et al., 2015) as well as the particular response tasks (Looi, Gfeller, & Driscoll, 2012). Some CI users show remarkably accurate perception for particular stimuli, are more adept at utilizing contextual cues to accommodate the degraded auditory signal, and/or are highly motivated to seek out musical experiences, thus promoting their own experience-based plasticity. But as a group, CI users demonstrate significant limitations in perception of spectrally-complex sounds, both in music and in speech (Gfeller et al., 2008, 2010; Limb & Roy, 2014; Looi, Gfeller & Driscoll, 2012; See et al., 2013)

Keeping in mind that the high-precision listening tasks required in music training have been hypothesized as important in the transfer to linguistic processes (Patel, 2011), what is the impact of the lack of fine structure in the CI signal? One could argue that CI users are less likely than NH listeners to benefit from music training, both with regard to salient features of music as well as speech. Some preliminary electrophysiological or radiographic data suggest that CI recipients may be less able to discern small pitch changes or spectral components (e.g., Limb & Roy, 2014; Petersen et al., 2015). In addition, some music training studies suggest possible limitations to the use of more complex stimuli with multiple cues (e.g., Vandali et al., 2015). Nevertheless, some CI users are more successful than others, and some even exceed expectations in extracting sufficient fine structure from the degraded signal to support reasonable perception of some spectrally complex sounds (Gfeller et al., 2008, 2010). Those results suggest that CI users may have greater potential to extract meaningful information from the degraded signal under particular circumstances, including focused training. For example, multiple regression analyses of a group of 209 CI users indicate that more extensive musical training prior to implantation is associated with greater accuracy on several measures of music perception (Gfeller et al., 2008, 2010). Can training after implantation optimize perception as well?

Music training of pitch-based or timbral features of music

Several studies indicate that short-term computer-assisted music training can enhance perceptual accuracy by CI recipients on some aspects of music, despite the technical limitations of the CI. Through training, adult CI users have improved pitch discrimination (Vandali et al., 2015), melodic contour recognition (Galvin, Fu, & Nogaki, 2007; Galvin et al., 2012), complex (real-world) melody recognition (Gfeller, Witt, et al., 2000), timbre recognition (Driscoll, 2012; Gfeller, Witt, Adamek, et al., 2002; Gfeller, Witt, Woodworth, et al., 2002), improved appraisal ratings (sound quality) (Gfeller, Witt, Adamek, et al., 2002; Gfeller, Witt, Woodworth, et al., 2002), or general enjoyment and participation in music experiences (Gfeller, 2001; Gfeller, Mehr, & Witt, 2001; see review in Looi et al., 2012). This evidence that CI users can improve some aspects of music perception through focused training is of interest not only in relation to enhancing music enjoyment and quality of life, but also because pitch and timbre are both acoustic features associated with perception of several spectrally complex features of speech (See et al., 2013).

As we consider the clinical benefits of music training, it is important to note that individual CI users have varied considerably on rate and extent of benefit in response to various music training protocols (Driscoll, et al., 2009; Fu & Galvin, 2007; Galvin, Fu, & Nogaki, 2007; Gfeller, Witt, et al., 2000; Gfeller, Mehr, & Witt 2001; Gfeller, Witt, Adamek, et al., 2002; Gfeller, Witt, Woodworth, et al., 2002; Looi, Gfeller & Driscoll, 2012). Therefore, music training may be differentially beneficial for adult CI users, possibly due to individual differences as they interact with specific training parameters used in each study.

Correlations between music and speech perception of CI users

Also relevant to this discussion are studies indicating that those CI users who have more accurate pitch and timbre perception are also more likely to have better perceptual outcomes on spectrally-complex elements of speech, such as linguistic or affective prosody, talker identification, and speech in background noise (Gfeller et al., 2012; See, 2013). Several studies documenting statistically significant correlations between measures of pitch, timbre, and speech in noise suggest shared perceptual abilities (Drennan & Rubinstein, 2008; Gfeller et al., 2007, 2008, 2010). Thus, if short-term training can enhance perception of spectrally complex features of music, and significant correlations have been documented between perception of pitch or timbre and spectrally complex features of speech, it seems plausible that music training may also generalize to speech perception of CI users.

As we consider further the possibility of music-based training for adult CI users, let us consider those features of training that have been identified in prior research as efficacious for persons with hearing loss. The following section discusses different factors (e.g., program parameters, stimulus choices) that may influence training efficacy of music as well as speech.

Parameters for Music-based Training

Common approaches to auditory training

Two general approaches from the literature on auditory training have been applied to music training: analytic, which emphasizes bottom-up perceptual processes, and synthetic, which emphasizes top-down cognitive processes (American Academy of Audiology, 2010; Fu & Galvin, 2007; Gfeller, et al., 2001; Looi, Gfeller & Driscoll, 2012; Moore & Amitay, 2007). Analytic approaches expose the listener to increasingly difficult contrasts in acoustic features (such as isolated pitches or timbres in music, or phonemes in speech). These features are often presented using an adaptive algorithm, or in fixed but gradually increasing levels of difficulty. The intent is to increase perceptual efficiency in hearing small changes and to facilitate more efficient processing throughout the auditory system; this may generalize to tasks reliant upon similar processing skills (Fu & Galvin, 2007; Kraus et al., 2009; Moore & Amitay, 2007; Strait et al., 2009).

Synthetic approaches are designed to promote more efficient central (cognitive) processing (e.g., enhanced attention, use of contextual cues, priming), which can assist the listener in extracting sufficient useable information from the signal (Boothroyd, 2010; Fu & Galvin, 2007; Gfeller et al., 2001; Moore & Amitay, 2007). Synthetic approaches are also more likely to use stimuli that are 'connected' and more naturalistic (e.g., real voices or instruments and musical compositions, as opposed to computer-generated pure tones or unique harmonic complexes). Connected stimuli in speech training would include complete sentences or paragraphs, as opposed to brief acoustical stimuli such as isolated phonemes. In music training, connected stimuli would include listening to or playing complete musical phrases, songs, or excerpts from longer musical compositions. These sorts of tasks may require integrated perception over time of rapidly changing, multiple cues, and focused attention to particular aspects of the stimuli (e.g., focusing on a vocalist against a background accompaniment).

Synthetic training would more nearly resemble the sorts of auditory experiences associated with music lessons, ensembles, music theory, and music listening for appreciation. Consequently, it more nearly resembles the sorts of music training of participants in those studies examining experience-based plasticity of musicians. Connected stimuli can be made up of either complex computer-generated stimuli (e.g., melodies and harmonies created using MIDI technology), recordings of live singers and musical instruments, or actual music-making experiences. A synthetic listening task may include multimodal, attentional or contextual cues (e.g., non-auditory cues associated with prior listening experiences), which can assist the listener in processing the auditory signal.

Analytic and synthetic approaches can be differentially beneficial, depending upon the listening circumstances, the auditory stimulus, listener capabilities (Fu & Galvin, 2007; Song et al., 2008), and hearing history and age, which influences relevance of contextual cues or neural plasticity (Chermak, 2010). In optimal listening environments with good stimuli (e.g., segmental features of speech in quiet listening conditions, or the rhythmic component of music), analytic training may be most beneficial. On the other hand, synthetic training may help listeners to compensate in suboptimal listening conditions, or when the

acoustic signal is of poor quality (e.g., coarse representation of pitch, or sung lyrics partially masked by an accompaniment). For postlingually deaf CI users, a training protocol can include contextual references to listening experiences prior to hearing loss. Specific to music training, some CI users have improved pitch and timbre perception as a result of analytic (bottom up) or synthetic (top down) training (Driscoll, 2012; Gfeller, Witt, et al., 2000; Gfeller, Witt, Adamek, et al., 2002; Gfeller, Witt, Woodworth, et al., 2002; Fu & Galvin, 2007; Moore & Amitay, 2007).

Training stimuli characteristics

Within analytic or synthetic approaches, various parameters in regard to stimuli choice and response options (e.g., discrimination, recognition) have been utilized. Vandali et al., (2015) have noted several key features associated with successful perceptual training, including the provision of feedback, the use of highly variable auditory stimuli within a program (e.g., using several different frequencies or different spectral characteristics), perceptual fading (i.e., beginning with most extreme exemplars, and progressing to more natural sounding stimuli as perception improves) and training with primary cues alone (e.g., training on pitch alone, or timbre alone); graded variations in secondary cues are gradually introduced (Vandali et al., 2015).

In relation to choice of musical stimuli for auditory training, let us consider once again the rationale posited regarding generalization of music training to linguistic skills. In studies linking music training with enhanced linguistic functions in normal hearing individuals, the stimulus complexity of music has been identified as a key component. According to Patel (2011), the benefits of music training to enhance linguistic capabilities rests in part upon the perceptual precision required when listening to music, which can in turn generalize to speech perception tasks that share neural processing requirements.

Spectrally rich stimuli (as opposed to pure tones) in training for CI users has been supported through several studies of adults with CIs (Galvin, et al., 2012) or NH persons training with CI simulations (Loebach & Pisoni, 2008; Loebach et al., 2009). The use of more complex training stimuli (e.g., melodic contours paired with masking, complex instrumental sounds) has resulted in greater perceptual enhancement for some listeners, both for music and transfer to various linguistic tasks.

Data from published music-based training with CI users indicate highly varied results (e.g., Fu & Galvin, 2007; Fu et al. 2015), suggesting that some CI users may perceive enough discriminable features to achieve benefits from training. Other psychological factors such as attention, motivation, and use of contextual cues are also likely to influence training benefit. It is not yet clear whether CI users who access more fine structure in the first place, are therefore, more motivated to listen more frequently to music (which in turn promotes neural adaptation), or if greater exposure to musical experience can also help those with less robust auditory systems to improve (Looi, Gfeller & Driscoll, 2012).

Given the enormous universe of musical sounds possible (computer-generated or natural instrumental or vocal sounds), the numerous auditory sub-skills that could be targeted, and the variability among CI users on a host of factors (e.g., hearing history, personality traits,

life experiences, etc.), there seem to be nearly infinite factors and combinations thereof that could influence training outcomes. Thus, it is difficult to offer clear clinical recommendations for training stimuli based upon available studies. The optimal choice of stimuli may vary depending upon the targeted auditory task (Fu et al., 2015), and some approaches may be more or less successful in relation to the unique characteristics and expectations of a given CI user. These questions will require ongoing research efforts and for the present, careful clinical observation of individual change over time in response to clinical applications.

In summary, it is not yet clear, whether the rich and varied spectral characteristics in music that benefit NH listeners would also enhance speech perception in CI recipients who hear a degraded signal through a damaged auditory system (Shahin, 2011). Further research is needed to better understand potential benefits of complex musical stimuli in training protocols for CI users. In addition, other factors of music training that have been associated with benefit for speech should be addressed, more specifically, aspects of music associated with enhanced attention and motivation (emotional reward), which support persistence. These factors are the foci of the following section.

Motivation and Persistence

When we consider the role of music within our cultures, and within the lives of individuals, music is not only a rich acoustic signal, but also a socio-cultural phenomenon often associated with positive emotion (Gfeller, 2008). In most cultures, music is a potent form of entertainment, and is associated with social and culturally significant events (e.g., graduation, weddings, sporting events, parties, etc.) (Gfeller, 2008). Participating in singing or playing instruments, particularly when the task is well-matched to one's capabilities and preferences, tend to be associated with strong positive emotion and reward, such as social attention, praise, and the pleasure of hearing beautiful music (Herholz & Zatorre, 2012; Patel, 2011).

Several researchers who address music and experience-based plasticity from the standpoint of cognitive neuroscience emphasize the positive emotions or reward value associated with music-based training (Herholz & Zatorre, 2012; Patel, 2011). Positive emotions or reward can enhance attention; this can recruit more neurons to attend to subtle changes, and can increase synchrony of neural firing, and thus basic encoding of sound features. This premise reflects what good clinicians have implemented in rehabilitation from a behavioral standpoint for many years—therapy works best when the individual is highly engaged, motivated and attending to the appropriate stimuli.

However, one cannot presume that music training will have reward value for all individuals. As anyone who has ever taken music lessons can attest, practicing musical skills is not always enjoyable. Tedium or frustration can occur in the course of the numerous trials required to achieve satisfactory results, particularly if the listener lacks natural aptitude, or (even more problematic) must face difficult perceptual conditions, as is the case with CI users. A number of studies report that music listening or participation tends to be of lesser pleasure for adult CI users than their recollection of music with normal hearing (Gfeller,

Mehr, & Witt, 2001; Gfeller et al., 2008, 2010). Thus, the extent of reward associated with musical training may be considerably less for CI users, who are reliant on electric hearing, than for normal hearing persons. In short, given the technical limitations of the CI, one cannot presume that CI users would experience positive emotions within a musical context.

Those adult CI users who are strongly motivated to re-establish sufficient musical enjoyment to participate in social and aesthetic experiences including music, and who are open to 'new' forms of music appreciation, however, may find music a satisfying tool for enhancing auditory processing. The desired conditions of positive emotion and focused attention, from a clinical as well as neuroscience standpoint, emphasize the importance of selecting developmentally appropriate and engaging musical stimuli that are behaviorally relevant (Herholz & Zatorre, 2012).

With adult CI users, who typically face multiple responsibilities of daily life, and who are more likely than children to seek practical benefits to instructional experiences, training materials should be interesting, have apparent practical benefits, and require reasonable time commitment in order to promote persistent and attention (Gfeller, 2001). A potentially inherent advantage of connected musical stimuli is that musical forms usually include repeated melodic, harmonic, and rhythmic patterns. This built-in repetition helps the listener organize and predict acoustical patterns, which can enhance listening efficiency and enjoyment. This brings us to the issue of duration of training and repetition, given that repeated trials are essential for neural plasticity and learning.

Duration of training

From a logistical standpoint, generalizing studies of experience-dependent plasticity based on years of high-level music training is problematic in relation to CI recipients (Shahin, 2011). While general wisdom suggests that more training and spaced rehearsal (as opposed to massed practice) is better, it is not yet clear how much training is essential to achieve clinical benefit; sufficient training may differ depending upon stimuli, response tasks, and listener differences (Boothroyd, 2010; Fu & Galvin, 2007; Kraus et al., 2009; Moore & Amitay, 2007).

Training programs described in published studies have varied on overall program length, frequency of sessions, and length of sessions. Music training studies with CI simulations have documented significant changes from as little as 45 minutes of acute training (Loebach & Pisoni, 2008; Loebach et al., 2009). In one CI study, individuals required < 1 week, to > 4 weeks to achieve significant gains (Fu & Galvin, 2007; Galvin, Fu, & Nogaki, 2007). Research from our own lab (simulations, CI users) has shown significant improvements within 3 weeks (further consolidation over 5 wks.), though the individual benefit is variable (Driscoll et al., 2009; Driscoll, 2012).

Clear evidence and methodological guidelines for CI users have yet to emerge regarding music perception and possible generalization to speech. However, rehabilitation studies typically are limited by accessibility, compliance, and motivation of trainees; thus, most studies of adult CI recipients have involved relatively short-term training, as opposed to the long-term music training associated with many studies of music experience-based plasticity

(Looi, Gfeller & Driscoll, 2012; Vandali et al., 2015). Because experience-based plasticity requires sufficient repetition/exposures to relevant auditory stimuli, a successful program should strike a suitable balance between adequate repetition and a time requirement that is not so burdensome as to undermine initial participation and persistence.

In summary, a host of factors have been identified as relevant to success in perceptual training, and more specifically, suitability for adult CI recipients. The large number of relevant factors and options for combining these factors in a protocol makes it difficult to offer clear-cut recommendations for music-based training with this population. Nevertheless, systematic evaluation of music-based training for CI recipients is warranted, particularly to confirm whether hypothesized benefits will yield outcomes equivalent or superior to well-documented speech rehabilitation programs. Acknowledging the numerous procedural challenges associated with devising a suitable training program, we have developed a music-based training program for CI recipients that integrates program parameters or approaches identified in prior literature (as described above) that reflect various aspects of perceptual learning. The description of those principles or factors is discussed in the following sections.

Applying prior knowledge to the development of a training protocol

This section describes principles and factors that we have integrated into a computer-based music training program for adult postlingually deaf CI users. The program is made up of twelve 30-minute instructional modules self-administered at home via a computer. The following considerations have emerged from examining prior studies relevant to this topic, and we describe how each consideration informs the development of our music training program that is currently undergoing field testing:

- 1.0** *Analytic and synthetic methods provide differential benefits to auditory learning* (Fu & Galvin, 2007; Looi, Gfeller & Driscoll, 2012).
- 1.1** *Application to training:* Our program includes approximately half analytic and half synthetic exercises in each lesson, thus we are promoting various types and levels of auditory processing.
 - 1.1.1** Analytic exercises focusing on changes in pitch and in timbre present several key features associated with bottom-up perception of spectrally-complex features of speech as well as music. These exercises permit considerable control over stimulus parameters in response to the listener's capabilities.
 - 1.1.2** Top-down processing (synthetic) can be enhanced through the use of training tasks that encourage the listener to attend to specific auditory features (e.g., listening to sung lyrics against background accompaniment), or to utilize contextual or non-auditory cues. In contrast with analytic items, the synthetic exercises more nearly resemble the diverse types and blends of musical stimuli and listening tasks associated with training of musicians.
- 2.0** *The auditory system is fine-tuned by experience in processing stimuli that requires more fine-grained discrimination* (Patel, 2011).

2.1 *Application to training:* Our training program emphasizes spectrally complex stimuli. We focus on timbre and pitch for the analytic training, because both require spectral selectivity, and are strongly correlated (Donnelly & Limb, 2009; Drennan & Rubinstein, 2008). Additionally, accuracy of pitch perception can be influenced by the timbre of testing stimuli (Drennan & Rubinstein, 2008; Galvin, Fu, & Shannon, 2009; Vandali et al, 2015). While speech, like music, has pitch and timbre/tone quality elements, musical sounds are more diverse with regard to frequency range and spectral diversity (Chasin & Russo, 2004); thus, music offers a diverse and challenging pool of training stimuli, which fits Patel's recommendation (2011) for listening tasks that require greater precision than speech.

2.1.1 Analytic items include MIDI-created complex tones of acoustic piano, flute, clarinet, oboe, also saxophone, violin, cello, and trombone played over a wide frequency range.

2.1.2 Synthetic items include MIDI-created melodies and harmonies on acoustic piano, guitar; recorded live singers; and excerpts from real-world music, representing pop, country, and classical genres. These real-world excerpts include complex and rapidly changing blends of melodies, harmony, timbre, rhythm, and amplitude.

Two of the synthetic modules include sung lyrics, which introduce linguistic content into the music training protocol; it is plausible that this may help with transfer to speech processing outcomes. The modules with sung lyrics with background accompaniment might be considered a musical analogue to understanding speech in background noise, in that the listener must extract the word (lyrics) from a complex and rapidly changing blend of background sounds.

3.0 *Training is enhanced by the use of highly variable auditory stimuli* (Vandali, 2015; Boothroyd, 2010).

3.1 *Application to training:* We have included a diverse pool of instrumental and vocal music sounds that include different solo timbres, timbral blends, and complex combinations of pitch, timbre, rhythm, and amplitude. The components range from simple isolated stimuli with one relevant cue to complex and rapidly changing items with combinations of auditory cues.

3.1.1 Analytic items include 9 different instrumental timbres and a wide frequency range (110–740 Hz). Because training with a diverse sequence of stimuli, or multiple auditory cues (e.g., concurrent changes in pitch and timbre) can be more difficult for some listeners (Vandali et al., 2015), the analytic tasks within our training program initially present stimuli with one predominant cue (either timbre or pitch). Because perceptual fading can enhance training, the program commences with hearing contrasts of vastly different pitches or timbres. As the listener progresses, the comparisons become increasingly difficult.

- 3.2 Synthetic items include complex combinations of pitch, harmony, rhythm, timbre and amplitude from music of different genre, with and without sung lyrics.
- 4.0 Within the context of rehabilitation, repetition through practice and repeated trials is essential for neural plasticity and learning.
- 4.1 *Application to training:* Our computer-assisted program has been designed to provide repetition balanced with practical use and features designed to foster motivation and persistence.
 - 4.1.1 Repetition is provided through a computer program, which can be completed at home, and includes spaced practice of twelve 30-minute lessons distributed over a period of four weeks, a time considered acceptable by CI recipients trained in prior studies (Looi et al., 2012). Each sub-skill is presented multiple times. For the analytic tasks, the level of the initial trials is established from baseline testing, and the level of difficulty is increased in response to 80% accuracy.
 - 4.1.2 Ease of use is supported through content presented in larger font and good visual resolution; this facilitates ease of reading and comprehension for many adult CI users who are 50 years or older and thus have perceptual and cognitive characteristics associated with adult learners (Gfeller, 2001). The training materials utilize vocabulary suitable for non-musicians and require no reading of music notation; they are interactive in nature, including feedback on perceptual accuracy. CI users listen to various musical items, and then indicate what was heard through on-screen prompts. The responses are automatically saved to the computer, which permits on-going assessment of perceptual accuracy as well as program compliance.
 - 4.1.3 Because the adult CI users in the training group were postlingually deafened, and grew up with exposure to music as part of everyday life, persistence is encouraged through the choice of ecologically relevant music (e.g., relevant to real life listening), and computer-generated feedback on accuracy (Gfeller, 2001). These materials, which reflect these concepts, (summarized in Table 1) are an extension of music training that we have used successfully for over two decades with adults with hearing losses (Driscoll et al., 2009; Driscoll, 2012; Gfeller, Christ, et al., 2000; Gfeller, Witt, et al., 2000; Gfeller, Mehr, Witt, 2001; Gfeller, Witt, Adamek, et al., 2002; Gfeller, Witt, Woodworth, et al., 2002; Gfeller et al., 2007).

Training Program Evaluation

A novel aspect to the design of this training program is that progress is measured from baseline to post-training using both behavioral tests and evoked potentials. We will examine if auditory evoked responses can help us document changes in neural processing that result from training, and how these changes relate to behavioral and functional outcomes. Hopefully, evoked potentials could help to identify individuals most likely to benefit from training.

The following behavioral measures are being gathered: (1) Pure tone frequency discrimination (difference limens) at 200, 800, and 1600 Hz (Gfeller et al., 2007, 2008); (2) musical (complex) pitch discrimination (Gfeller et al., 2007, 2008); (3) familiar song recognition (Gfeller et al., 2012); (4) timbre discrimination and recognition (Gfeller, Witt, Adamek, et al., 2002; Gfeller, Witt, Woodworth, et al., 2002; Gfeller et al., 2008); (5) complex melody recognition test (Gfeller et al., 2012); (6) *AZ Bio* sentences at three signal-to-noise ratios (0, -5, and -10 dB) (Spahr & Dorman, 2004); (7) spectral ripple discrimination (Henry & Turner, 2003).

The specific auditory evoked potential that we use to assess outcome is the Acoustic Change Complex (ACC). The ACC is a recording of neural activity at the level of the auditory cortex that is associated with discrimination between two ongoing sounds (Martin & Boothroyd, 2000). It can be evoked using a passive listening paradigm. It has been successfully recorded using spectrally complex acoustic stimuli presented in the sound field and processed by a hearing aid or a cochlear implant (Brown et al., 2008; Friesen & Tremblay, 2006; Kirby & Brown, 2015). Previous work from our lab and others has shown that changes in these cortical auditory evoked potentials may parallel changes in perception associated with training (Brown et al., 2014; Tremblay & Kraus, 2002). We have designed the training program so that many of the stimuli used for training (e.g. pitch and/or timbre contrasts) can also be used to evoke the ACC response.

We hope this electrophysiologic measure of how the auditory signal is coded at the cortex, combined with behavioral measures of how accurately these contrasts are perceived, and how malleable they are to the effects of training, will help us better understand the variance in perceptual accuracy commonly observed among CI recipients. The findings from this study with postlingually deafened adult CI users may also be useful in refining habilitative practices for pediatric patients.

Summary

In summary, this paper has provided an overview of rationales for music-based training for adult CI recipients, focusing on training parameters associated with successful learning. This forms the basis for a computer-based training program that is currently being field tested with postlingually deaf adult CI users. An overview of the training components as well as outcome measures has been provided.

Acknowledgments

The authors would like to thank all those who took part in this study and Alisha Luymes and Laura Runion for assistance in reviewing trainings prior to dissemination.

This study was supported by grant 2 P50 DC00242 and RO1 DC012082-02 from the NIDCD, NIH and the Iowa Lions Foundation.

Abbreviation

CI Cochlear implant

NH	Normal hearing
SNR	Signal-to-Noise Ratio
MIDI	Musical instrument digital interface
ACC	Acoustic Change Complex

References

- American Academy of Audiology. Music Training and Cochlear Implants. 2010. viewed 11 October 2011, <http://www.audiology.org/news/Pages/20100421.aspx>
- Besson M, Schön D, Moreno S, Santos A, Magne C. Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience*. 2007; 25:399–410. [PubMed: 17943015]
- Boothroyd A. Adapting to changed hearing: The potential role of formal training. *Journal of the American Academy of Audiology*. 2010; 21:601–611. [PubMed: 21241648]
- Brown CJ, Etler C, He S, O'Brien S, Erenberg S, Kim J, Dhuldhoya AN, Abbas PJ. The electrically evoked auditory change complex: preliminary results from nucleus cochlear implant users. *Ear and Hearing*. 2008; 29:704–717. [PubMed: 18596644]
- Brown, CJ.; Gfeller, K.; Abbas, P.; Jeon, J.; Driscoll, V.; Mussoi, B.; Tejani, V. Musical training: effects on perception and electrophysiologic measures of discrimination. American Auditory society Scientific and Technology Meeting; Scottsdale Arizona. March 8; 2014.
- Brown CJ, Etler C, He S, O'Brien S, Erenberg S, Kim JR, et al. The electrically evoked auditory change complex: preliminary results from nucleus cochlear implant users. *Ear and Hearing*. 2008; 29:704–717. [PubMed: 18596644]
- Chermak G. Music and auditory training. *The Hearing Journal*. 2010; 63:57–58.
- Chasin M, Russo FA. Hearing AIDS and music, *Trends in Amplification*. 2004; 8 (2):35–47. [PubMed: 15497032]
- Donnelly, PJ.; Limb, C. Music perception in cochlear implant users. In: Niparko, JK., editor. *Cochlear Implants: Principles and Practices*. 2. Philadelphia: Lippincott, Williams & Wilkins; 2009. p. 223
- Drennan WR, Rubinstein JT. Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of Rehabilitation Research Development*. 2008; 45:779–789. [PubMed: 18816426]
- Driscoll V. The effects of training on recognition of musical instruments by adults with cochlear implants. *Seminars in Hearing*. 2012; 33:410–418. [PubMed: 23503992]
- Driscoll V, Oleson J, Jiang D, Gfeller K. Effects of training on recognition of musical instruments presented through cochlear implant simulations. *Journal of the American Academy of Audiology*. 2009; 20:71–82. [PubMed: 19927684]
- Friesen LM, Tremblay KL. Acoustic change complexes recorded in adult cochlear implant listeners. *Ear and Hearing*. 2006; 27:678–85. [PubMed: 17086078]
- Fu Q, Galvin JJ. Perceptual learning and auditory training in cochlear implant recipients. *Trends in Amplification*. 2007; 11:193–205. [PubMed: 17709574]
- Fu Q, Galvin JJ, Wang X, Wu J. Benefits of music training in Mandarin-speaking pediatric cochlear implant users. *Journal of Speech, Language, and Hearing Research*. 2015; 58:163–169.
- Galvin JJ, Eskridge E, Oba S, Fu QJ. Melodic contour identification training in cochlear implant users with and without a competing instrument. *Seminars in hearing*. 2012; 33:399–409.
- Galvin JJ, Fu Q, Nogaki G. Melodic contour identification by cochlear implant listeners. *Ear and Hearing*. 2007; 28:302–319. [PubMed: 17485980]
- Galvin JJ, Fu Q, Shannon RV. Melodic contour identification and music perception by cochlear implant users. *Annals of the New York Academy of Science*. 2009; 1169:518–533.
- Gfeller K. Aural rehabilitation of music listening for adult cochlear implant recipients: Addressing learner characteristics. *Music Therapy Perspectives*. 2001; 19:88–95.

- Gfeller, KE. Music: A human phenomenon and therapeutic tool. In: Davis, WB.; Gfeller, KE.; Thaut, MH., editors. An introduction to music therapy theory and practice. 3. Silver Spring, MD: American Music Therapy Association; 2008. p. 41
- Gfeller KE, Christ A, Knutson JF, Witt SA, Murray KT, Tyler RS. The musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *Journal of the American Academy of Audiology*. 2000; 11:390–406. [PubMed: 10976500]
- Gfeller K, Jiang D, Oleson J, Driscoll V, Knutson J. Temporal stability of music perception and appraisal scores of adult cochlear implant recipients. *Journal of the American Academy of Audiology*. 2010; 21:28–34. [PubMed: 20085197]
- Gfeller K, Jiang D, Oleson JJ, Driscoll V, Olszewski C, Knutson JF, Turner C, Gantz B. The effects of musical and linguistic components in recognition of real-world musical excerpts by cochlear implant recipients and normal-hearing adults. *Journal of Music Therapy*. 2012; 49:68–101. [PubMed: 22803258]
- Gfeller K, Mehr M, Witt S. Aural rehabilitation of music perception and enjoyment of adult cochlear implant users. *Journal of the Academy for Rehabilitative Audiology*. 2001; 34:17–27.
- Gfeller K, Oleson J, Knutson J, Breheny P, Driscoll V, Olszewski C. Multivariate predictors of music perception and appraisal by adult cochlear implant users. *Journal of the American Academy of Audiology*. 2008; 19:120–134. [PubMed: 18669126]
- Gfeller KE, Turner CW, Zhang X, Gantz BJ, Froman RJ, Olszewski CA. Accuracy of cochlear implant recipients on pitch perception, melody recognition and speech reception in noise. *Ear and Hearing*. 2007; 28(3):412–423.10.1097/AUD.0b013e3180479318 [PubMed: 17485990]
- Gfeller K, Witt S, Adamek M, Mehr M, Rogers J, Stordahl J, et al. Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients. *Journal of the American Academy of Audiology*. 2002; 31:132–145. [PubMed: 11936169]
- Gfeller K, Witt S, Stordahl J, Mehr M, Woodworth G. The effects of training on melody recognition and appraisal by adult cochlear implant recipients. *Journal of the Academy of Rehabilitative Audiology*. 2000; 33:115–138.
- Gfeller K, Witt S, Woodworth G, Mehr M, Knutson JF. Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Annals of Otology, Rhinology, and Laryngology*. 2002; 111:349–356.
- Henry BA, Turner CW. The resolution of complex spectral patterns in cochlear implant and normal hearing listeners. *Journal of the Acoustical Society of America*. 2003; 113:2861–2873. [PubMed: 12765402]
- Herholz SC, Zatorre RJ. Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*. 2012; 76:486–502. [PubMed: 23141061]
- Ingvalson EM, Wong PCM. Training to improve language outcomes in cochlear implant recipients. *Frontiers in Psychology*. 2013; 4:1–9. [PubMed: 23382719]
- Kirby BJ, Brown CJ. Effects of nonlinear frequency compression on AAC amplitude and listener Performance. *Ear and Hearing*. 2015
- Kraus N, Skoe E. New directions: Cochlear implants. *Annals of New York Academy of Sciences*. 2009; 1169:516–517.
- Kraus N, Skoe E, Parbery-Clark A, Ashley R. Experience-induced malleability in neural encoding of pitch, timbre, and timing-implications for language and music. *Annals of the New York Academy of Sciences*. 2009; 1169:543–557. [PubMed: 19673837]
- Limb CJ, Roy AT. Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hearing research*. 2014; 308:13–26. [PubMed: 23665130]
- Loebach JL, Pisoni DB. Perceptual learning of spectrally degraded speech and environmental sounds. *Journal of the Acoustical Society of America*. 2008; 123:1126–1139. [PubMed: 18247913]
- Loebach JL, Pisoni DB, Svirsky MA. Transfer of auditory perceptual learning with spectrally reduced speech to speech and non-speech tasks: Implications for cochlear implants. *Ear and Hearing*. 2009; 30:662–674. [PubMed: 19773659]
- Looi V, Gfeller K, Driscoll V. Music appreciation and training for cochlear implant recipients: a review. *Seminars in Hearing*. 2012; 33:307–334. [PubMed: 23459244]

- Looi V, King J, Kelly-Campbell R. A music appreciation training program developed for clinical application with cochlear implant recipients and hearing aid users. *Seminars in Hearing*. 2012; 33:361–380.
- Martin BA, Boothroyd A. Cortical, auditory, evoked potentials in response to changes of spectrum and amplitude. *Journal of the Acoustical Society of America*. 2000; 107:2155–61. [PubMed: 10790041]
- Moore DR, Amitay S. Auditory training: Rules and applications. *Seminars in Hearing*. 2007; 28:99–109.
- Musacchia G, Sams M, Skoe E, Kraus N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States*. 2007; 104:15894–15898.
- Patel, AD. Music, language, and the brain. New York: Oxford university press; 2007.
- Patel A. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*. 2011; 2:142. [PubMed: 21747773]
- Petersen B, Weed E, Sandmann P, Brattico E, Hansen M, Sorensen S, Vuust P. Brain response to musical feature changes in adolescent cochlear implant users. *Frontiers in Human Neuroscience*. 2015; 9:1–14. doi:10.3389/fnhum.2015.00007. [PubMed: 25653611]
- See RL, Driscoll VD, Gfeller K, Kliethermes S, Oleson J. Speech intonation and melodic contour recognition in children with cochlear implants and with normal hearing. *Otology & Neurotology*. 2013; 34:490–498. [PubMed: 23442568]
- Shahin A. Neurophysiological influence of musical training on speech perception. *Frontiers in Psychology*. 2011; 2:126. [PubMed: 21716639]
- Song JH, Skoe E, Wong PC, Kraus N. Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*. 2008; 20:1892–1902. [PubMed: 18370594]
- Spahr AJ, Dorman MF. Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices. *Archives of Otolaryngology Head & Neck Surgery*. 2004; 130:624–628. [PubMed: 15148187]
- Strait DL, Kraus N, Skoe E, Ashley R. Musical experience promotes subcortical efficiency in processing emotional vocal sounds. *Annals of the New York Academy of Sciences*. 2009; 1169:209–213. [PubMed: 19673783]
- Swanson B, Dawson P, McDermott H. Investigating cochlear implant place-pitch perception with the modified melodies test. *Cochlear Implants International*. 2009; 10 (suppl 1):100–104. [PubMed: 19148894]
- Tremblay KL, Kraus N. Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech Language Hearing Research*. 2002; 45:564–572.
- Vandali A, Sly D, Cowan R, van Hoesel R. Training of cochlear implant users to improve pitch perception in the presence of competing place cues. *Ear and Hearing*. 2015; 36:e1–e13.10.1097/AUD.0000000000000109 [PubMed: 25329372]
- Wong PCM, Skoe E, Russo NM, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*. 2007; 10:420–422. [PubMed: 17351633]

Table1**Music Training Components**

Training Components	Response Task	Format	Stimuli characteristics: Timbre, pitch
Pitch Discrimination	Which of three tones (timbre held constant) is a different pitch? (3AFC)	A, FL	9 MIDI timbres; Interval changes of 25 cents to 7 semitones; Hz range: A2-G5
Timbre Discrimination	Which of three tones (pitch held constant) is a different timbre? (3AFC)	A, FL	9 MIDI timbres; Hz range: A2-G5
Melodic Pattern Identification	Identify ascending, descending, Up-down-up, down-up down patterns of 3 tones	A, FL	MIDI piano; 4, 3, 2 semitone pitch intervals; Hz range: C4-C6
Melodic Error Detection**	Identify the presence of an incorrect note in musical phrases from 2 familiar melodies	A/S	MIDI clarinet; Error of 2, 3, 4 semitones; Hz range: F#3-F#5
Song Recognition w/Accompaniment	Familiar song identification: lyrics and melody cues w competing background accompaniment (5 songs)	S	Voice, guitar, piano; accompaniment range=123–392 Hz
Timbre Recognition	Pair image and of, info about instrument with MIDI sound file of 8 instruments;	S	8 MIDI timbres; High, medium, low frequencies
Recognition of recorded excerpts of real-world songs	Song recognition, 4AFC	S	8 excerpts of real-world vocal +instrumental pop or country tunes
Virtual Support Group Suggestions	Read practical tips on music listening from CI users		No auditory stimuli

A=analytic approach, S=Synthetic, FL=fixed level

* adapted from melodic contour tasks published in research by Fu & Galvin. (2007)

** adaptation of Melodic Error Detection test developed by Brett Swanson et al., (2009)