

Music perception in cochlear implant users and its relationship with psychophysical capabilities

Ward R. Drennan, PhD;^{1*} Jay T. Rubinstein, MD, PhD¹⁻²

¹V. M. Bloedel Hearing Research Center, Department of Otolaryngology, University of Washington School of Medicine, Seattle, WA; ²Department of Bioengineering, University of Washington, Seattle, WA

Abstract—This article describes issues concerning music perception with cochlear implants, discusses why music perception is usually poor in cochlear implant users, reviews relevant data, and describes approaches for improving music perception with cochlear implants. Pitch discrimination ability ranges from the ability to hear a one-semitone difference to a two-octave difference. The ability to hear rhythm and tone duration is near normal in implantees. Timbre perception is usually poor, but about two-thirds of listeners can identify instruments in a closed set better than chance. Cochlear implant recipients typically have poor melody perception but are aided with rhythm and lyrics. Without rhythm or lyrics, only about one-third of implantees can identify common melodies in a closed set better than chance. Correlations have been found between music perception ability and speech understanding in noisy environments. Thus, improving music perception might also provide broader clinical benefit. A number of approaches have been proposed to improve music perception with implant users, including encoding fundamental frequency with modulation, “current-steering,” MP3-like processing, and nerve “conditioning.” If successful, these approaches could improve the quality of life for implantees by improving communication and musical and environmental awareness.

Key words: cochlear implant, deafness, hearing, hearing loss, music, perception, psychoacoustics, sound processing, spectral resolution, speech perception, speech processing.

INTRODUCTION

The cochlear implant, introduced commercially in the mid-1980s, has become hugely successful. Over

100,000 people who had profound or severe hearing impairments can now hear thanks to this technological marvel. Speech understanding in most individual users is good to excellent. The clinical impact of the device has been nothing less than extraordinary. The device, however, has some shortcomings. Understandably, the cochlear implant was designed to enable good speech perception when speech is presented in quiet. While successful in delivering speech in quiet, its performance in delivering music and speech in background noise has been less than ideal. Implant users rank music as the second most important acoustic stimulus in their lives next to understanding speech [1], and most implant users find that music does not sound good on their device. “I got speech back but not music,” is a common comment from cochlear implant patients. This article describes some reasons why cochlear implant users do not hear music

Abbreviations: ACE = Advanced Combination Encoder, CAMP = Clinical Assessment of Music Perception, CCIS = Conditioned Continuous Interleaving Sampling, CIS = continuous interleaving sampling, F0 = fundamental frequency, PACE = Psychoacoustic Advanced Combination Encoder, PMMA = (Adapted) Primary Measures of Musical Audiation, SRT = speech reception threshold.

*Address all correspondence to Ward R. Drennan, PhD; V. M. Bloedel Hearing Research Center, University of Washington, Box 357923, CHDD Building Room CD176, Seattle, WA 98195; 206-897-7923; fax: 206-616-1828.

Email: drennan@u.washington.edu

DOI: 10.1682/JRRD.2007.08.0118

well, some results of experiments that evaluate music perception ability in cochlear implantees, the relationship between these abilities and other psychophysical measures of hearing, and ways in which sound processing might be improved to address shortcomings of the cochlear implant with respect to music perception.

To understand why the implant does not encode music well, one must understand a little about how a normal-hearing auditory system encodes music. One of the fundamental elements of music is pitch. Pitch has long been debated in hearing science, back to the 19th century. The debate concerns whether acoustic frequency is encoded spatially, according to place of excitation, or temporally, according to timing of repetition cycles. A complex tone is a set of simultaneously occurring acoustic sine waves usually having a harmonic relationship. The frequency of complex tones is perceived as pitch by the human ear. Helmholtz argued for a place theory of pitch [2]. He hypothesized that the ear acts as a spectrum analyzer and tones of different frequencies are encoded at different places in the ear. This theory was later substantiated by von Békésy's work in the mid-20th century, which showed that the basilar membrane in the cochlea acted roughly like a spectrum analyzer with high frequencies responding best at the basal end and low frequencies responding best at the apical end. Seebeck put forward a temporal theory of pitch [3–4]. He demonstrated empirically that the repetition rate of clicks established a pitch (see Green [5] and Wightman and Green [6] for discussions). At the time, Helmholtz was the dominant scientist and his theory was taken as correct, but given about one hundred years, the scientific community came to understand that both theories had merit [7].

The periodicity element of pitch is particularly necessary when listening to complex tones. Unlike a pure tone, exciting primarily a narrow space on the basilar membrane, a complex tone, like those created by the human voice or musical instruments, comprises numerous harmonics varying in frequency over a wide range. Musical melodies, with even a single instrument, are composed of a series of complex tones. The difference in the spectrum of different complex tones, especially with limited place-pitch resolution, can be quite small. The basis of good perception of complex-tone pitch lies in the repetition rate, which depends on fine-structure temporal encoding. The complex tone has a periodicity rate corresponding to the fundamental frequency. The hair cells act as a half-wave rectifier and respond every cycle of the wave. This

excites the nerve fibers in synchrony with the repetition rate of the acoustic wave, the accuracy of which is good at low frequencies and remains good up to about 2,000 Hz in the normal-hearing system. The degree of synchrony to periodicity declines until it is nearly nonexistent at about 5,000 Hz, above which repetition rate is not encoded [8].

If Helmholtz's theory were wholly correct, cochlear implant users would probably hear music fairly well, but they do not. Sound processing in cochlear implants relies heavily on the Helmholtzian theory in that it works as a place encoder [9–10]. The processors take the acoustic input via a microphone, divide the input in real time into a finite set of frequency channels, extract the temporal envelope of the acoustic wave in each channel, and deliver that temporal envelope into the cochlea with a fixed-rate sequence of biphasic electrical pulses. The fine timing also known as the "fine-structure" of the sound waves is largely lost in the process. Typically, cochlear implant users can only hear repetition rates up to about 300 Hz [11]. Thus, much of the fine-structure that could be used to encode pitch is absent. Implant users must rely heavily on perceiving temporal envelopes (not temporal fine-structure) at specific places. While attempts have been made to encode the periodicity with the electric pulse rate, the fixed-rate method has become dominant in practice, because it has been, to date, clinically the most effective method. Unfortunately, without the ability to extract the periodicity of the waveforms, the pitch contour of a complex-tone melody is extraordinarily muddled [12].

This difficulty is further exacerbated by the spectral resolving power of the implant users' auditory systems. Numerous physiological deficiencies exist in an implant user's auditory system, which can be quite variable and can limit spectral resolving power. In deafness, the auditory nerves deteriorate. Dendrites die back and cell bodies shrink [13–16]. Additional limitations arise from the nature and placement of the implant. The safest and easiest manner for inserting the cochlear implant is through a cochleostomy adjacent to the round window and into the scala tympani. This placement provides good spectral spread because it is near the end point of the auditory nerve where the place map is spatially most spread; however, the electric fields are fairly broad and cannot be focused on a specific place the way an inner hair cell can excite a specific auditory neuron. Electrical stimulation excites a population of nerve fibers. Possibly compounding the problem, the standard electrode configuration

today is monopolar, in which electrical current runs from an electrode in the cochlea to positions outside the cochlea. This method has been the most effective clinically, but unfortunately yields less-than-desirable precision in the place of excitation. Thus, implant users are not only limited by the inability to extract periodicity pitch, but they also perceive pitch based on a degraded spatial representation. This effect has been demonstrated behaviorally. For example, even though implantees get as many as 22 channels of stimulation, they have only been shown to have 3 to 9 “functional” channels [17–19]. For high levels of melody recognition, at least 64 channels are needed [12].

Varying dynamics can add dramatic effects to musical composition; however, dynamic changes in level can confound good pitch perception in cochlear implant users. Dynamic changes are encoded as changes in current level. Such changes in level increase the spread of current, potentially stimulating off-frequency spiral ganglion cells and eliciting an altered pitch [20]. Pitch usually increases in level but not always. From a practical standpoint, A-440 (a sound with fundamental frequency [F0] of 440 Hz) at a quiet level might sound lower in pitch than A-440 at a high level. Timbre is another important aspect of music that is defined as a perceptual quality of a sound that differentiates it from other sounds having the same pitch and loudness. In composition, different timbres are used to create variety in the musical sound, introducing another dimension for musical expression. Even within a single instrument, an artist can generate subtle and artistically pleasing effects by varying the timbre. Timbre is defined physically by the temporal envelope (particularly the onset) and the spectral shape of the acoustic sound [21]. In cochlear implants, the temporal envelopes are encoded fairly well, providing an element of timbre perception to the implantee. The spectral shape is somewhat smeared by the various physiological and engineering limitations, as discussed earlier in terms of the spectral resolving power. These elements limit the ability of implant users to discriminate different musical instrument timbres, but implant users will still have some ability to discriminate timbre given some spectral resolving power and temporal envelope cues.

The dynamic range in electric hearing is highly limited [22]. In normal hearing, the dynamic range is as much as 120 dB. In electric hearing, it can be as little as 10 or 20 dB, owing primarily to the high degree of neural synchrony created by electrical stimulation [23–24] and

the lack of spontaneous activity in the deaf cochlea [16,25–26]. The limited dynamic range in electric hearing has multiple potential negative effects. First, and most directly, it limits the dynamics of music. The implantee will not hear the same dramatic ranges of level that the normal-hearing listener can hear. Second, and more subtly, however, compression of the dynamic range reduces across-spectrum level differences that define the spectral shape. Such differences contribute to the perception of speech—vowels in particular—and to the perception of timbre [21]. Third, the depth of temporal modulations will be compressed. Temporal modulations contribute to the perception of all types of sound. Thus, increasing the dynamic range in electric hearing could improve hearing by providing better resolution of the dynamically varying range of levels in both the spectral and temporal domains.

The final aspect of music perception is rhythm. Coding of the temporal envelope in the implant, which would encode rhythm, is quite good. Shannon, for example, has found that the discrimination of timing events in cochlear implant recipients is nearly normal [22]. Part of the reason for this nearly normal discrimination is the high degree of synchrony between the electrical impulse and the nerve firing described earlier. Acoustic onsets are well defined through the temporal envelope in cochlear implants. A sequence of onsets defines common rhythmic patterns in music. The primary limitation of the implant is the inability to encode the periodicity or fine-structure, which happens in a considerably shorter time than typical musical rhythmic patterns.

ELEMENTS OF MUSIC PERCEPTION IN COCHLEAR IMPLANT USERS

Pitch

The study of pitch perception in cochlear implantees is particularly interesting, because the cochlear implant provides the unique opportunity to independently manipulate temporal and place contributors to pitch perception. Both types of studies have been conducted with implant users. Variation of the electric pulse rate in cochlear implants can elicit a pitch percept, as can changing the place of stimulation by stimulating different electrodes. Both of these approaches are used to investigate the basic function of the pitch perception in implantees by directly manipulating pulse rates and current to individual

electrodes. Other practical studies evaluate the functional ability of implantees to discriminate the frequencies of complex tones.

Cochlear implant users can, based solely on temporal changes, discriminate pitches, hear musical intervals, and even discriminate melodies [27–28]. These results clearly support a temporal mechanism underlying pitch discrimination, but in implant users, the temporal pitch mechanism is only functional at low rates up to about 300 pps [11], with a few exceptions up to about 1,000 pps [29]. Discrimination thresholds based on pulse rate average about 7.3 percent, or just over 1 semitone [30]. Higher frequency and constant pulse rates encode speech better, so modern processors often use constant pulse rates far greater than 600 pps. Thus, implant processing has not been able to practically employ this temporal-pitch sensitivity using pulse rates. Implant users are, however, sensitive to modulation frequency [31–32] that is well encoded by today's high-rate pulsatile stimulation. Modulation frequencies up to 300 Hz also elicit a pitch percept, especially with large modulation depths, e.g., 100 percent.

Townshend et al. found a broad range of place-pitch sensitivity in three subjects with the use of monopolar stimulation [33]. In this type of stimulation, commonly used today, the electrical current runs from an electrode within the cochlea to electrodes outside the cochlea. While one subject was sensitive to a change from one electrode to the next, another could barely detect a pitch difference between any electrodes. Tong and Clark also found a broad range of sensitivity for bipolar configurations where current runs from one electrode to the next within the cochlea [29]. Their listeners had good sensitivity ($d' = 1$) for separations ranging from one to five electrodes. Nelson et al. completed a thorough study of electrode discrimination ability using more cochlear implant listeners ($n = 14$) and monopolar stimulation [34]. They found results similar to Tong and Clark's. Some people had extremely good place-pitch discrimination ability with just one electrode (0.75 mm) separation required for good discrimination, and others required three or four electrodes of separation of electrical excitation to achieve good performance ($d' = 1$).

Of practical interest is how well cochlear implant listeners can discriminate complex tones such as those produced by a voice or a musical instrument. As noted earlier, this task is particularly challenging for implant users for two primary reasons. First, the place of excitation is obscured because the multiple sinusoidal components

of complex sounds excite a broad range of electrodes, and second, the temporal periodicity is obscured because the modern processors use constant-rate pulse trains. However, low-rate periodicities can be encoded as amplitude modulation frequencies, which can contribute to F0 discrimination.

In testing the ability of 49 implant users to discriminate the F0 of a complex sound, Gfeller et al. found that the ability to discriminate a change in pitch direction was widely variable [35]. Some listeners could discriminate one semitone reliably, whereas others required an F0 difference as much as two octaves to detect the difference. Mean performance was 7.56 ± 5.18 standard deviation of semitones. Using continuous interleaving sampling (CIS) processing [36], Guerts and Wouters observed somewhat better performance in four listeners, with most users discriminating F0 within several tones, with the exception of one person who could not discriminate an octave [37]. Nimmons et al. investigated F0 discrimination ability using the University of Washington Clinical Assessment of Music Perception (CAMP) [38]. They found that complex-tone pitch direction discrimination ranged from less than 1 semitone in a few implant users to as much as 12 semitones in others. Nearly all their listeners had F0 discrimination ability between one and six semitones.

Rhythm

Another critical element of music is rhythm. Recognizing melodies based solely on rhythm with no pitch information is possible. The neural system of cochlear implant users actually locks more tightly to timing information than the normal-hearing system [24]. Additionally, cochlear implant users have behaviorally been shown to hear temporal changes such as gaps and amplitude modulation as well as normal-hearing listeners [22,39]. In cochlear implant users, rhythmic differences are encoded as temporal gaps, amplitude modulations, or both. Rhythm discrimination is generally good but not quite as good in cochlear implant listeners as in normal-hearing listeners [40–43].

Gfeller et al. used the Adapted Primary Measures of Musical Audiation (PMMA) test to evaluate rhythmic pattern discrimination and a six-pulse task in 17 listeners [40]. In the PMMA test, listeners discriminated two patterns that differed in their rhythmic pattern by note duration. In the 6-pulse task, listeners had to identify a change in the temporal location or short interval among four long intervals. The overall durations of the stimuli were

decreased (i.e., tempo increased) until the listeners could no longer hear the changing location of the short interval. In the PMMA rhythmic pattern test, both normal-hearing and implant listeners scored about 84 percent correct. In the 6-pulse task, implant users could hear the changes for an average pattern duration of about 1,070 ms, whereas normal-hearing listeners could hear the changes at a significantly shorter duration of 607 ms.

Kong et al. investigated tempo discrimination ability and complex rhythm discrimination ability using four listeners [42]. They found that tempo discrimination was near normal in implant users. Both implantees and normal-hearing listeners could hear tempo change of four to six beats per minute. Complex rhythm discrimination, however, was not as good as for normal-hearing listeners. In a study of three cochlear implant listeners, one listener was nearly as good as normal, scoring near 95 percent correct. Another two listeners were typically performing about 20 percent less than normal. Thus, rhythm discrimination ability was good but, on average, not as good as normal-hearing listeners.

Timbre

Timbre is an essential element of the aesthetic value of music. Different instruments add musical “color” to performances and greatly enhance variety in musical sounds. The composer Ravel, for example, made impressive use of different timbres by using different instruments repeating the same melody many times in his composition *Bolero*. Timbre is encoded via the temporal envelope (onset characteristics in particular) and by the spectral shape of sound [21]. While the temporal envelopes are fairly well preserved in cochlear implant processing, the spectral information is reduced relative to normal-hearing listeners. Consequently, timbre recognition in cochlear implant users is better than chance but not nearly as good as for normal-hearing listeners.

Gfeller et al. studied timbre perception in 51 implant users with live recordings of eight different musical instruments [44]. They reported a mean performance of 47 percent correct, whereas normal-hearing listeners correctly identified timbre 91 percent of the time on the same test. McDermott reported that implantees scored 44 percent versus 97 percent for normal-hearing listeners [43]. McDermott also presented a confusion matrix that suggested that percussive instruments, i.e., those with distinctive temporal envelopes, were identified more readily than wind or string instruments. Confusions did

not seem to occur within instrument families, i.e., among woodwinds, brass, or strings, but rather across instrument families. Flute, for example, was often confused with trumpet, and the organ was often confused with violin. Using the CAMP, Nimmons et al. reported similar results with eight live-recorded musical instruments [38]. The mean timbre recognition for eight listeners was 49 percent correct, with a range of 21 to 54 percent correct. Their confusion analysis also indicated that the more percussive guitar and piano were easier to identify than wind or string instruments. Nimmons et al. and Gfeller et al. also observed confusions that were not consistent with musical instrument families, perhaps reflecting some other common characteristic or acoustic quality that was confused. For example, select instruments from different families might have similar temporal envelopes.

Melody

Melody perception in cochlear implant users is generally extremely poor, although in “real-world” melodies, lyrics and rhythms can help users identify melodies. Gfeller et al. found that implant users were not nearly as good as normal-hearing listeners at identifying real-world melodies [45]. The real-world melodies were categorized by musical genre. A group of 79 listeners was tested on identifying 15 songs in each genre. Performance with country and pop music averaged about 20 and 17 percent, respectively, while performance for the lyric-less classical music averaged about 10 percent correct. Kong et al. presented 12 familiar songs with and without rhythm to 6 cochlear implant listeners [42]. All notes had the same duration when no rhythm was presented, forcing the users to base a decision solely on the pitch contour. With rhythm, performance averaged from 50 to 60 percent; without rhythm, all but one subject was at chance levels, with performance averaging 10 to 15 percent. In a similar study, Gfeller et al. presented implant users with 12 familiar melodies [35]. Some melodies were highly rhythmic and others had limited rhythmic cues. Eighteen implant users averaged about 20 percent correct with the rhythmic melodies and 10 percent with the arrhythmic melodies. This contrasts with normal-hearing performance on the same test of 90 percent for rhythmic melodies and 77 percent for the melodies with limited rhythmic cues.

Nimmons et al. employed a rhythmless melody test as part of the CAMP that used isochronous melodies in which long tones were repeated in a continuous eighth-note

pattern [38]. They reported a mean score of 23 percent correct for 8 listeners with 12 common melodies. This included an outlier at 81 percent correct. Galvin et al. conducted a study of melodic contour discrimination in which they used simple pitch contours rather than a familiar melody [46]. Using nine different contours, average performance increased from 32 to 64 percent as the interval between notes increased from one semitone to five semitones. With the broad, five-semitone separation, performance ranged from 14 percent correct (chance levels) to 98 percent correct. They also conducted a familiar melody recognition test and got a mean performance of 60 percent with rhythmic but only 28 percent without rhythmic cues. Galvin et al. further observed that with sustained training up to about 50 days, melodic contour identification improved about 20 percent for the originally good performers and 30 to 50 percent for those who had poor performance initially. Taken as a whole, the results suggest that melody recognition in implantees is generally poor and highly dependent on rhythmic cues; however, melodic contour identification can be improved with training.

RELATIONSHIPS WITH OTHER PSYCHO-PHYSICAL TASKS

In recent studies at the University of Washington and University of Iowa, correlations have been observed between speech perception, music perception, and spectral and temporal discrimination abilities in cochlear implantees. The spectral ripple test, developed at the University of Iowa, is a good clinical test to evaluate the ability of implant users to resolve the acoustic spectrum [47–48]. This test evaluates listeners' abilities to resolve the acoustic spectrum by asking listeners to differentiate different "ripples" in the shape of the acoustic spectrum. In recent testing of 30 implantees using the CAMP test and the spectral ripple test, researchers found that spectral ripple discrimination ability correlates significantly with melody recognition ($r = 0.52$, $p = 0.003$), timbre recognition ($r = 0.64$, $p < 0.001$), and pitch direction discrimination ability ($r = -0.46$, $p = 0.01$) [38,49].

The Schroeder-phase test has been used to evaluate the ability of birds and humans to hear changes in the temporal fine-structure of acoustic waves [50]. This test was also used to evaluate temporal sensitivity in implantees [51–52]. Implant sound processors, such as the Cochlear

Corporation (Lane Cove NSW, Australia) Advanced Combination Encoder (ACE) or the Advanced Bionics Corporation (Valencia, California) HiRes[®], transform the Schroeder temporal fine-structure changes into a rapid series of temporal envelope packets that sweep either from apex to base or from base to apex in the cochlea for positive and negative Schroeder phase, respectively. Implant users were asked to discriminate the difference between positive and negative Schroeder phase for 50, 100, 200, and 400 Hz complexes. Each one of these frequencies yields successively faster sweeps. For 25 implantees, average Schroeder-phase discrimination ability correlated with pitch direction discrimination ($r = -0.4$, $p = 0.047$).

Won et al. assessed the ability of 28 implantees to evaluate speech in noise using adaptive tracking with spondees in speech-shaped random noise [49]. The approach yielded a speech reception threshold (SRT) that was the signal-to-noise ratio at which implantees could accurately detect 50 percent of the words [53]. Significant correlations were found between SRTs and melody ($r = -0.52$, $p = 0.005$), timbre ($r = -0.58$, $p = 0.001$) and pitch direction discrimination ($r = 0.57$, $p = 0.0017$). Gfeller et al. found similar correlations between melody identification ability and word recognition in quiet with 33 subjects ($r = 0.65$, $p < 0.001$) [35]. These data strongly suggest that improving music perception in implant recipients will also yield benefit in other practical tasks such as speech recognition in quiet and in noise. Thus, improving music perception ability in implant recipients will likely provide clinical benefit in multiple domains and could dramatically enhance the quality of life for implant users.

IMPROVING MUSIC PERCEPTION WITH COCHLEAR IMPLANTS

A number of laboratories have implemented several approaches that are intended to improve music perception with cochlear implants. These include "current steering," MP3-like processing, 100-percent amplitude modulation across channels at the F0, the combination of acoustic and electric stimulation, and nerve conditioning.

Current steering uses simultaneous activation of neighboring electrodes that are weighted appropriately to match a spectral shape. This method has been shown to increase the number of pitch percepts in cochlear implant

users [54–55]. Current steering is believed to improve spectral resolution, and in theory, the number of pitch percepts available with current steering is sufficiently high to enable much better music perception ability. Potentially, more than 64 possible pitch percepts exist with current steering. This method is available presently using the Advanced Bionics with processing called “HiRes Fidelity 120™.” This processing could aid music perception, although any objective clinical benefits, outside of increasing the number of potential pitch percepts, have yet to be documented.

Another approach, “MP3000,” is being tested in a clinical trial by Cochlear Corporation. This method uses psychophysical masking to limit information transfer of acoustical information that is masked. MP3000 is a modification of the ACE strategy commonly used with Cochlear Corporation implants. ACE is an “ n of m ” MP3000-type strategy that picks the maximum n spectral peaks of m channels for delivery to the implantee. The new method, formerly known as Psychoacoustic Advanced Combination Encoder (PACE), picks the largest amplitude component for transmission but determines the masking pattern of this component and then selects the next largest acoustic component of the stimulus, taking into account the masking and nonlinear summation of excitation that would be produced in a normal-hearing auditory system. In this way, only the most perceptually salient components of the stimulus will be delivered to the implant users rather than the largest components. Such an approach could improve spectral resolution and thereby enhance music perception. Tested acutely, using a sentence-in-noise test, Nogueira et al. demonstrated with eight listeners that a four-channel PACE algorithm ($4 = n$ in “ n of m ”) exceeded the performance of ACE by 17 percent [56]. No significant difference was found with eight channels, although only five listeners were tested in this condition and a trend appeared in which PACE yielded slightly better performance. Combined with psychophysical results demonstrating relationships between speech perception and music perception [49], this result suggests that MP3000 might yield better music perception.

Laneau et al. employed an approach called “F0mod” in which the F0 was extracted from acoustic waves and used to modulate all channels with 100 percent amplitude modulations [57]. This approach provided a clear temporal cue for the perception of frequency. Laneau et al. showed that F0mod produced much better complex-tone pitch

discrimination ability, better melody recognition, and better estimations of pitch intervals in cochlear implant users. The results suggest that better incorporation of temporal pitch cues could enhance music perception abilities in cochlear implant recipients.

In a subset of the severely hearing-impaired population with some residual hearing, the combination of acoustic and electric hearing can be used to achieve better music perception than electric hearing alone. A hearing aid can be used in combination with a long-electrode array on the opposite side or with a short-electrode array on the same side to combine acoustic and electric hearing. If some low-frequency residual hearing still exists, the apical hair cells can provide essential low-frequency fine-structure information, improving the ability of these patients to hear music. Also, when some residual hearing is present, use of a hearing aid is sometimes recommended to delay additional nerve atrophy that could result from continued auditory deprivation [58]. Kong et al. demonstrated that melody recognition can be improved as much as 20 to 30 percent when acoustic residual hearing on one aided side is combined with electric hearing on the contralateral side [59]. Gfeller et al. tested “hybrid” patients who used both acoustic hearing and electric hearing in the same ear [60]. Given some residual hearing in the apical region, a short-electrode array was implanted in the basal region combined with a hearing aid for low-frequency hearing. Melody recognition for a group of subjects using the hybrid approach was 40 percent better than for a group with the traditional long-electrode cochlear implant, and without lyrics, performance was 50 percent better. Timbre recognition showed similar results, with the hybrid group identifying instruments about 20 percent more accurately than the electric-only group. While these effects are dramatic, the hybrid group most likely has better nerve survival, given their residual hearing. That is, their hearing loss is not as severe as those without any residual hearing. Even so, the hybrid approach offers excellent opportunities for appropriate candidates. As of January 2008, the hybrid was still in the clinical trial phase.

Rubinstein et al. proposed an innovative approach to sound processing that could be combined with current sound-processing approaches [61]. This processing is based on a physiological model and is known as “conditioning.” The auditory nerve in a deaf ear lies dormant if it receives no input [16,25–26]. The auditory nerve in a normal-hearing ear has spontaneous activity resulting from stochastic

hair cell activity even when it receives no acoustic input [62–63]. This spontaneous activity effectively keeps the nerve ready to fire when an acoustic event activates it. In response to electrical stimulation, the auditory nerve fires with a high degree of synchrony across the spectrum [23–24], causing an extremely limited dynamic range for electric hearing. The normal range of about 120 dB is limited to 10 to 30 dB in electric hearing [22]. By use of a “conditioner,” a low-level, constant amplitude high-rate pulsatile stimulation, “pseudospontaneous” activity can be created in the cat auditory nerve that reduces across-channel synchrony and increases the electric dynamic range [64]. Conditioning has also been shown physiologically to produce temporal firing patterns in electric hearing that are more like those seen in normal hearing [65–67]. While the auditory nerves of implanted humans deteriorate as a result of inactivity and might not respond as predictably as the nerves in the animal studies, conditioning could possibly generate a more normal or more natural nerve response in humans.

Conditioned processing could improve hearing in a number of ways. The increased dynamic range could yield better spectral encoding by providing better resolution in the amplitude domain [68]. Likewise, the temporal envelope might be better defined with improved resolution in the amplitude domain. This could yield better discrimination of pitch via better encoding of the F0 of complex sounds using F0 amplitude modulations. It could also produce better speech discrimination by having an improved resolution of the temporal envelope. Additionally, improved temporal firing patterns could yield better encoding of pitch via temporal mechanisms. Thus, conditioning could produce better pitch and melody recognition.

A Conditioned Continuous Interleaved Sampling (CCIS) processing strategy has been implemented for normal daily wear in about 30 subjects in Leiden, the Netherlands; Iowa City, Iowa; and Seattle, Washington. This processing uses a low-level conditioning stimulus at about 10 percent of the dynamic range in combination with traditional CIS processing [3]. The initial results suggest that this type of processing yields better speech understanding in quiet or in noise in about half the subjects. The implementation of CCIS is, however, limited at present because it can only be implemented with 8 channels on the Advanced Bionics Corporation CII BTE (Behind-The-Ear), whereas nonconditioning strategies using the same device have 16 processing channels. Only five sub-

jects have been tested to date on music perception with the CAMP. While many implantees report that music sounds better with CCIS, only a few have demonstrated objective musical benefit using CCIS. Making an acute comparison of eight-channel CCIS versus eight-channel CIS, two of five participants showed improvement on the timbre test of 17 and 21 percent correct. One participant was actually 16 percent worse on the timbre test using CCIS than using CIS. The remaining two participants had the same performance with both strategies. Only one of the five participants showed improved pitch perception with CCIS. For that person, fundamental frequency difference limens were 2.7 semitones better with CCIS averaged across 4 frequencies. No differences were seen with melody recognition testing. With melodies, all five participants were at or near chance levels regardless of the strategy. Given the current limitation on the quantity of data and the practical limitations in the implementation of CCIS, more research needed to adequately address the effectiveness of conditioning on music perception.

CONCLUSIONS

Electric sound processing approaches are not likely to provide “normal” music perception, but efforts to improve music perception in implantees are of the utmost importance. Improved music perception could have a highly positive impact on the lives of implantees, not only by providing potential enhanced enjoyment of music, but also by improving their overall hearing and improving speech understanding in quiet and in noise. One advantage of electrical stimulation is that it is quite good at delivering one of the most fundamental elements of music: rhythm. Rhythm can provide a sense of musical pleasure to a listener on its own, particularly to those deaf at birth or early in life with no previous experience with melodies. Electrical stimulation, however, lacks the ability to deliver detailed spectral information that helps define the pitch as well as the different colors or timbres of music. Electrical stimulation also lacks the ability to deliver temporal fine-structure, which helps normal-hearing listeners hear exact pitches and segregate the different instruments as well as the tones of a chord. Scientists and engineers will endeavor to improve electrical delivery of these perceptual attributes.

ACKNOWLEDGMENTS

This material was based on work supported in part by the National Institutes of Health, grants R01-DC007525 and P30-DC004661 and a subcontract of P50-DC00242. Additional funding was provided by Advanced Bionics and Cochlear Corporation.

Dr. Jay T. Rubinstein is a paid consultant for and receives research funding from Advanced Bionics and Cochlear Corporation, two manufacturers of cochlear implant devices. Advanced Bionics and Cochlear Corporation did not play any role in study design, data collection, analysis, or interpretation. Neither company played any role in the decision to write or submit this article for publication.

REFERENCES

- Gfeller K, Christ A, Knutson JF, Witt S, Murray KT, Tyler RS. Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *J Am Acad Audiol*. 2000;11(7):390–406. [\[PMID: 10976500\]](#)
- Helmholtz HLF. *Die Lehre von den Tonempfindungen als physiologische Grundlage fuer die Theorie der Musik*. 1st ed. Brunswick (Germany): Vieweg-Verlag; 1863. Trans: *On the sensation of tone*. 2nd English ed. New York (NY): Dover Publications; 1954.
- Seebeck A. Beobachtungen uber einige Bedingungen der Entstehung von Tönen = Observations about some conditions of the origin of tone. *Annalen der Physik und Chemie*. 1841;53:417–36. German, English.
- Seebeck A. *Über die Sirene = About the siren*. *Annalen der Physik und Chemie*. 1843;60:449–81. German, English.
- Green DM. *An introduction to hearing*. Hillsdale (NJ): Lawrence Erlbaum Associates; 1976.
- Wightman FL, Green DM. The perception of pitch. *Am Sci*. 1974;62(2):208–15. [\[PMID: 4815736\]](#)
- Licklider JC. Auditory frequency analysis. In: Cherry C, editor. *Information theory*. New York (NY): Academic Press; 1956. p. 253–68.
- Johnson DH. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am*. 1980;68(4):1115–22. [\[PMID: 7419827\]](#)
- Drennan WR, Rubinstein JT. Sound processors in cochlear implants. In: Waltzmann SB, Roland JT, editors. *Cochlear implants*. 2nd ed. New York (NY): Thieme; 2006. p. 40–47.
- Loizou PC. Mimicking the human ear. *IEEE Signal Process Mag*. 1998;15(5):101–30.
- Zeng FG. Trends in cochlear implants. *Trends Amplif*. 2004;8(1):1–34. [\[PMID: 15247993\]](#)
- Smith ZM, Delgutte B, Oxenham AJ. Chimaeric sounds reveal dichotomies in auditory perception. *Nature*. 2002;416(6876):87–90. [\[PMID: 11882898\]](#)
- Nadol JB Jr, Young YS, Glynn RJ. Survival of spiral ganglion cells in profound sensorineural hearing loss: Implications for cochlear implantation. *Ann Otol Rhinol Laryngol*. 1989;98(6):411–16. [\[PMID: 2729822\]](#)
- Nadol JB Jr. Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation. *Otolaryngol Head Neck Surg*. 1997;117(3 Pt 1):220–28. [\[PMID: 9334769\]](#)
- Nadol JB Jr, Xu WZ. Diameter of the cochlear nerve in deaf humans: Implications for cochlear implantation. *Ann Otol Rhinol Laryngol*. 1992;101(12):988–93. [\[PMID: 1463299\]](#)
- Shepherd RK, Javel E. Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status. *Hear Res*. 1997;108(1–2):112–44. [\[PMID: 9213127\]](#)
- Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am*. 2001;110(2):1150–63. [\[PMID: 11519582\]](#)
- Dorman MF, Loizou PC, Rainey D. Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *J Acoust Soc Am*. 1997;102(4):2403–11. [\[PMID: 9348698\]](#)
- Fishman KE, Shannon RV, Slattery WH. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *J Speech Lang Hear Res*. 1997;40(5):1201–15. [\[PMID: 9328890\]](#)
- Arnoldner C, Kaider A, Hamzavi J. The role of intensity upon pitch perception in cochlear implant recipients. *Laryngoscope*. 2006;116(10):1760–65. [\[PMID: 17003738\]](#)
- Handel S. Timbre perception and auditory object formation. In: Moore BC, editor. *Hearing*. 2nd ed. San Diego (CA): Academic Press; 1995. p. 425–61.
- Shannon R. Psychophysics. In: Tyler RS, editor. *Cochlear implants: Audiological foundations*. San Diego (CA): Singular Publishing, Inc; 1993. p. 357–89.
- Moxon EC. Neural and mechanical responses to electrical stimulation of the cat's inner ear [thesis]. Cambridge (MA): Massachusetts Institute of Technology; 1971.
- Van den Honert C, Stypulkowski PH. Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings. *Hear Res*. 1984;14(3):225–43. [\[PMID: 6480511\]](#)
- Kiang NY, Moxon EC, Levine RA. Auditory nerve activity in cats with normal and abnormal cochleas. In: Wolstenholme GE, Knight J, editors. *Sensorineural hearing loss*. London (England): Churchill; 1970. p. 241–68.

26. Liberman MC, Dodds LW. Single-neuron labeling and chronic cochlear pathology. II. Stereocilia damage and alterations of spontaneous discharge rates. *Hear Res.* 1984; 16(1):43–53. [\[PMID: 6511672\]](#)
27. Pijl S, Schwarz DW. Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes. *J Acoust Soc Am.* 1995;98(2 Pt 1):886–95. [\[PMID: 7642827\]](#)
28. McDermott HJ, McKay CM. Musical pitch perception with electrical stimulation of the cochlea. *J Acoust Soc Am.* 1997;101(3):1622–31. [\[PMID: 9069629\]](#)
29. Tong YC, Clark GM. Absolute identification of electric pulse rates and electrode positions by cochlear implant patients. *J Acoust Soc Am.* 1985;77(5):1881–88. [\[PMID: 3839004\]](#)
30. Moore BC, Carlyon RP. Perception of pitch by people with cochlear hearing loss and by cochlear implant users. In: Plack CJ, Oxenham AJ, Fay RR, editors. *Pitch: Neural coding and perception.* New York (NY): Springer; 2005. p. 234–77.
31. McKay CM, McDermott HJ, Clark GM. Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. *J Acoust Soc Am.* 1994;96(5 Pt 1): 2664–73. [\[PMID: 7983272\]](#)
32. McKay CM. Psychophysics and electrical stimulation. In: Zeng FG, Popper AN, Fay RR, editors. *Springer handbook of auditory research: Cochlear implants, auditory prostheses, and electric hearing.* New York (NY): Springer-Verlag; 2004. p. 286–333.
33. Townshend B, Cotter N, Van Compernelle D, White RL. Pitch perception by cochlear implant subjects. *J Acoust Soc Am.* 1987;82(1):106–15. [\[PMID: 3624633\]](#)
34. Nelson DA, Van Tasell DJ, Schroder AC, Soli S, Levine S. Electrode ranking of “place pitch” and speech recognition in electrical hearing. *J Acoust Soc Am.* 1995;98(4):1987–99. [\[PMID: 7593921\]](#)
35. Gfeller K, Turner C, Mehr M, Woodworth G, Fearn R, Knutson JF, Witt S, Stordahl J. Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants Int.* 2002;3(1):29–53.
36. Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM. Better speech recognition with cochlear implants. *Nature.* 1991;352(6332):236–38. [\[PMID: 1857418\]](#)
37. Guerts L, Wouters J. Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *J Acoust Soc Am.* 2001;109(2):713–26. [\[PMID: 11248975\]](#)
38. Nimmons GL, Kang RS, Drennan WR, Longnion J, Ruffin C, Worman T, Yueh B, Rubenstein JT. Clinical assessment of music perception in cochlear implant listeners. *Otol Nuerotol.* 2007;29(2):149–55. [\[PMID: 18309572\]](#)
39. Shannon RV. Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hear Res.* 1983;11(2):157–89. [\[PMID: 6619003\]](#)
40. Gfeller K, Woodworth G, Robin DA, Witt S, Knutson JF. Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. *Ear Hear.* 1997;18(3):252–60. [\[PMID: 9201460\]](#)
41. Leal MC, Shin YJ, Laborde ML, Calmels MN, Verges S, Lugardon S, Andrieu S, Deguine O, Fraysse B. Music perception in adult cochlear implant recipients. *Acta Otolaryngol.* 2003;123(7):826–35. [\[PMID: 14575398\]](#)
42. Kong YY, Cruz R, Jones JA, Zeng FG. Music perception with temporal cues in acoustic and electric hearing. *Ear Hear.* 2004;25(2):173–85. [\[PMID: 15064662\]](#)
43. McDermott HJ. Music perception with cochlear implants: A review. *Trends Amplif.* 2004;8(2):49–82. [\[PMID: 15497033\]](#)
44. Gfeller K, Witt S, Woodworth G, Mehr MA, Knutson J. Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Ann Otol Rhinol Laryngol.* 2002;111(4):349–56. [\[PMID: 11991588\]](#)
45. Gfeller K, Olszewski C, Rychener M, Sena K, Knutson JF, Witt S, Macpherson B. Recognition of “real-world” musical excerpts by cochlear implant recipients and normal-hearing adults. *Ear Hear.* 2005;26(3):237–50. [\[PMID: 15937406\]](#)
46. Galvin JJ 3rd, Fu QJ, Nogaki G. Melodic contour identification by cochlear implant listeners. *Ear Hear.* 2007;28(3): 302–19. [\[PMID: 17485980\]](#)
47. Henry BA, Turner CW, Behrens A. Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired and cochlear implant listeners. *J Acoust Soc Am.* 2005;118(2):1111–21. [\[PMID: 16158665\]](#)
48. Won JH, Drennan WR, Rubinstein JT. Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *J Assoc Res Otolaryngol.* 2007;8(3):384–92. [\[PMID: 17587137\]](#)
49. Won JH, Drennan WR, Kang R, Longnion J, Rubinstein JT, editors. Relationships among music perception, speech perception in noise, Schroeder phase and spectral discrimination ability in cochlear implant users. In: *Proceedings of the Conference on Implantable Auditory Prostheses; 2007 Jul 15–20; Lake Tahoe, CA.* Los Angeles (CA): House Ear Institute; 2007.
50. Dooling RJ, Leek MR, Gleich O, Dent ML. Auditory temporal resolution in birds: Discrimination of harmonic complexes. *J Acoust Soc Am.* 2002;112(2):748–59. [\[PMID: 12186054\]](#)
51. Longnion J, Ruffin C, Drennan WR, Rubinstein JT, editors. Discrimination of Schroeder-phase harmonic complexes by cochlear implant users [abstract]. In: *Proceedings of the Thirtieth Midwinter Meeting of the Association for Research in Otolaryngology; 2007; Denver, CO.*

52. Drennan WR, Longnion JK, Ruffin C, Rubinstein JT. Discrimination of schroeder-phase harmonic complexes by normal-hearing and cochlear-implant listeners. *J Assoc Res Otolarygol.* 2008;9(1):138–49. [\[PMID: 18066624\]](#)
53. Turner CW, Gantz BJ, Vidal C, Behrens A, Henry BA. Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing. *J Acoust Soc Am.* 2004;115(4):1729–35. [\[PMID: 15101651\]](#)
54. Firszt JB, Koch DB, Downing M, Litvak L. Current steering creates additional pitch percepts in adult cochlear implant recipients. *Otol Neurotol.* 2007;28(5):629–36. [\[PMID: 17667771\]](#)
55. Koch DB, Downing M, Osberger MJ, Litvak L. Using current steering to increase spectral resolution in CII and HiRes 90K users. *Ear Hear.* 2007;28(2 Suppl):38S–41S. [\[PMID: 17496643\]](#)
56. Nogueira W, Buchner A, Lenarz T, Edler B. A psychoacoustic “NofM”-type of speech coding strategy for cochlear implants. *EURASIP J Appl Signal Process.* 2005;18:3044–59.
57. Laneau J, Wouters J, Moonen M. Improved music perception with explicit pitch coding in cochlear implants. *Audiol Neurotol.* 2006;11(1):38–52. [\[PMID: 16219993\]](#)
58. Silman S, Silverman CA, Emmer MB, Gelfand SA. Adult-onset auditory deprivation. *J Am Acad Audiol.* 1992;3(6):390–96. [\[PMID: 1486201\]](#)
59. Kong YY, Stickney GS, Zeng FG. Speech and melody recognition in binaurally combined acoustic and electric hearing. *J Acoust Soc Am.* 2005;117(3 Pt 1):1351–61. [\[PMID: 15807023\]](#)
60. Gantz BJ, Turner C, Gfeller KE. Acoustic plus electric speech processing: Preliminary results of a multicenter clinical trial of the Iowa/Nuclear Hybrid implant. *Audiol Neurotol.* 2006;11 Suppl 1:63–68. [\[PMID: 17063013\]](#)
61. Rubinstein JT, Wilson BS, Finley CC, Abbas PJ. Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation. *Hear Res.* 1999;127(1–2):108–18. [\[PMID: 9925022\]](#)
62. Kiang NY, Moxon EC. Physiological considerations in artificial stimulation of the inner ear. *Ann Otol Rhinol Laryngol.* 1972;81(5):714–30. [\[PMID: 4651114\]](#)
63. Liberman MC. Auditory-nerve response from cats raised in a low-noise chamber. *J Acoust Soc Am.* 1978;63(2):442–55. [\[PMID: 670542\]](#)
64. Hong RS, Rubinstein JT. High-rate conditioning pulse trains in cochlear implants: Dynamic range measures with sinusoidal stimuli. *J Acoust Soc Am.* 2003;114(6 Pt 1):3327–42. [\[PMID: 14714813\]](#)
65. Litvak L, Delgutte B, Eddington D. Improved neural representation of vowels in electric stimulation using desynchronizing pulse trains. *J Acoust Soc Am.* 2003;114(4 Pt 1):2099–2111. [\[PMID: 14587608\]](#)
66. Litvak LM, Delgutte B, Eddington DK. Improved temporal coding of sinusoids in electric stimulation of the auditory nerve using desynchronizing pulse trains. *J Acoust Soc Am.* 2003;114(4 Pt 1):2079–98. [\[PMID: 14587607\]](#)
67. Litvak LM, Smith ZM, Delgutte B, Eddington DK. Desynchronization of electrically evoked auditory-nerve activity by high-frequency pulse trains of long duration. *J Acoust Soc Am.* 2003;114(4 Pt 1):2066–78. [\[PMID: 14587606\]](#)
68. Hong RS, Rubinstein JT. Conditioner pulse trains in cochlear implants: Effects on loudness growth. *Otol Neurotol.* 2006;27(4):50–56. [\[PMID: 16371847\]](#)

Submitted for publication August 8, 2007. Accepted in revised form March 18, 2008.