

New music for the Bionic Ear: An assessment of the enjoyment of six new works composed for cochlear implant recipients

Hamish Innes-Brown,^{*1} Agnes Au,^{**2} Catherine Stevens,^{^3} Emery Schubert,^{§4} Jeremy Marozeau,^{*5}

^{*} *The Bionics Institute, Melbourne, Australia*

[#] *Department of Audiology and Speech Pathology, The University of Melbourne, Australia*

[^] *MARCS Institute, University of Western Sydney, Australia*

[§] *School of English, Media and Performing Arts, University of New South Wales, Australia*

¹*hinnes-brown@bionicsinstitute.org*, ²*agnes.au@unimelb.edu.au*, ³*kj.stevens@uws.edu.au*,
⁴*e.schubert@unsw.edu.au*, ⁵*jmarozeau@bionicsinstitute.org*

ABSTRACT

The enjoyment of music is still difficult for many cochlear implant users. This study aimed to assess cognitive, engagement, and technical responses to new music composed specifically for CI users. From 407 concertgoers who completed a questionnaire, responses from groups of normally-hearing listeners (NH, n = 44) and CI users (n = 44), matched in age and musical ability, were compared to determine whether specially-commissioned works would elicit similar responses from both groups. No significant group differences were found on measures of interest, enjoyment and musicality, whereas ratings of understanding and instrument localization and recognition were significantly lower from CI users. Overall, ratings of the music were typically higher for percussion pieces. The concert successfully elicited similar responses from both groups in terms of interest, enjoyment and musicality, although technical aspects, such as understanding, localisation, and instrument identification continue to be problematic for CI users.

I. INTRODUCTION

INTERIOR DESIGN: Music for the Bionic Ear generated new musical works designed for hearing via electrical stimulation of the auditory nerve via a cochlear implant. There are nearly 200,000 cochlear implant users worldwide. Most implant users can understand speech in quiet environments (Helms et al., 2004), and children, when implanted early and provided with appropriate training and support, can integrate with their peers in the classroom and at play (Dettman, Pinder, Briggs, Dowell, & Leigh, 2007). However, there are two related situations where cochlear implant users still have difficulty. The first is the understanding of speech in noisy environments, or with multiple talkers (Loizou et al., 2009); and the second is the perception and enjoyment of music (McDermott, 2004). *INTERIOR DESIGN: Music for the Bionic Ear* is a complement to this body of research, and brought the unique experience and skills of composers to bear on the problem. Rather than attempt to improve the signal processing and engineering of the device itself, the composers involved worked to make new music, specifically tailored for listening through a cochlear implant.

This paper will firstly provide a high-level review of some of the possible reasons why some cochlear implant users might not enjoy music, then concentrate on the development and operation of the project, and finally will present results from an audience survey taken during two performances of the works in February 2011.

A. Sound, cochlear implants and music perception

Hearing via a cochlear implant is quite different to natural hearing – many of the biological systems that underlie the perception of pitch, timbre, and other perceptual aspects of sound are bypassed. To understand why music might not be enjoyable when heard through a cochlear implant, it is necessary to understand a little about how normal hearing works, and which parts of the auditory system are replaced by a cochlear implant.

The cochlea is a spiral structure with a shape similar to the Nautilus shells sometimes found washed up on beaches, and its job is to convert mechanical vibrations within the cochlea into electrical pulses in the auditory nerve. The cochlea is embedded in the temporal bone, and in a healthy ear contains rows of hair cells lined up along its length. These hair cells stimulate auditory nerves when they are moved by vibrations in the basilar membrane, in which they are mounted. The basilar membrane has mechanical properties causing it to resonate at different frequencies along its length. The hair cells are thus set in motion at different points along the membrane depending on the frequency of the sound. If hair cells close to the middle ear are vibrated, a high-pitched sound is heard, and the pitch gradually gets lower as hair cells deeper into the cochlear are vibrated. There are around 3500 of these hair cells along the length of the basilar membrane, and in people with profound ‘sensorineural’ hearing loss, it is these hair cells that are mostly damaged.

A cochlear implant largely replaces the function of the outer, middle, and most of the inner ear – up to the level of the auditory nerve. It consists of two main parts. First, the sound processor is worn externally, and hooks behind the ear. It contains microphones, batteries, and a miniaturized computer system that converts the acoustic signal received at the microphones into a series of electric pulses according to a programmable software algorithm called a ‘strategy.’ Implanted in the mastoid bone behind the ear is the implant itself. It receives power, as well as the electrical signals from the sound processor via a wireless link through the skin. At the end of the implant is a very fine linear array of up to 22 electrodes, which is inserted about half-way into the spiral-shaped cochlea. These electrodes stimulate the auditory nerve, thus replacing the function of the hair cells that are lost or damaged in sensorineural deafness.

The strategy embedded in the sound processor determines which combinations of electrodes to stimulate according to the acoustic signal received by the microphone. The strategy most

commonly used divides the incoming sound signal into as many frequency bands as there are electrodes (22 in the Cochlear Ltd Nucleus devices), selects a small number of the bands with the highest amplitude (typically the 8 highest of the total 22 available), and then stimulates those electrodes at a current level related to the smoothed amplitude in each band.

If a high-frequency pure tone is played, about 8 KHz for example, the first electrode, closest to the middle ear, is stimulated. If the frequency of the tone is gradually decreased, electrodes further and further into the cochlear are stimulated. The listener hears something akin to a high-pitched sound that gradually decreases in pitch.

There are a few complications to this basic description that are relevant to this discussion. In a healthy ear a pure tone stimulates a limited region of the cochlea, which is connected to a proportion of the 30,000 auditory nerves. Therefore, each auditory nerve is stimulated by a specific range of frequencies. However, in a CI the filters that separate the 22 bands overlap substantially. Depending on intensity, a pure tone can thus stimulate more than one electrode. Furthermore, each electrode is located at a distance from the nerves that it stimulates, and the current needed to stimulate the nerves spreads widely within the cochlea. This results in the excitation of a large number of auditory nerves tuned to other frequencies.

B. Speech

Speech signals convey semantic meaning through a rapid succession of vowel and consonant sounds. Vowel sounds (such as *i e æ ʌ ɒ ɔ u*) are produced without significant constrictions in the vocal tract, and are generally “voiced” – that is the vocal cords vibrate and produce a harmonic sound. Vowels in English and other non-tonal languages generally have a fairly consistent voicing frequency (F_0), with a unique pattern of harmonics called formants, labelled F_1 , F_2 , F_3 etc. Depending on the vowel sound produced, the first formant in American English varies between 300-770 Hz, and the second between 900-2300 (Hillenbrand, Getty, Clark, & Wheeler, 1995). Most vowel sounds can be distinguished by the first and second formants alone (Pols, van der Kamp, & Plomp, 1969).

In a CI, the steady first formant activates one or more of the lowest electrodes, and a number of the higher electrodes are activated by the higher formants – with a different pattern activated for each vowel sound. Thus the CI user receives a fairly unique pattern of electrode activation for each vowel sound. Figure 1 shows the F_1 and F_2 frequencies for some of the cardinal vowel sounds in Australian English with a male speaker (Cox, 1996), overlaid on a grid representing the edges of the default CI frequency bands. Crucially, speech can be understood with a relatively small number of vowel sounds, so that despite the problems of overlapping filter bands and current spread, there is enough frequency resolution using 22 electrodes for many CI users to successfully distinguish between many of the vowels (Blamey, Dowell, Brown, Clark, & Seligman, 1987; Eddington, 1980).

Consonants on the other hand can be either voiced or unvoiced, and are generally produced by forcing air through a constriction in the vocal tract. Examples are [p], where the constriction is at the lips, [d] which is produced by the front of the tongue, or [h], a slight constriction in the throat. Whereas vowels are comparatively simple acoustically, with a unique

pattern of formants distinguishing between most vowel sounds, consonants can vary in a large number of ways. Depending on how they are produced in the vocal tract, the voice onset time, degree of voicing, length of voicing, and the overall amplitude envelope are all variables that contribute to the linguistic meaning of consonant phonemes. However, compared to vowel sounds, which are distinguishable mostly on the basis of spectral or harmonic information, consonants are mostly distinguishable on the basis of how the overall amplitude varies in time (Diehl, Lotto, & Holt, 2004). There are obvious exceptions to this rule – whereas ‘tap’ and ‘tan’ differ in time-varying features of the final phoneme (among other features), ‘tap’ and ‘tat’ are less obviously different. The rate of stimulation pulses in CIs can vary from around 200 Hz up to 1200 Hz. At these rates, gross temporal cues can be transmitted fairly well. Thus, despite the complex acoustic nature of consonant sounds, many of the time-based cues used to distinguish between consonant sounds are successfully transmitted to the listener (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995).

C. Music.

As we have seen in the previous sections, speech signals generally consist of a relatively small number of vowel sounds, which are distinguishable based on a unique pattern of stimulated electrodes, and larger number of consonants, which are distinguishable by onset times and other time-varying aspects of the sound. Generally speaking, enough perceptual cues are transmitted by the implant such that a sufficient number of the phonetic elements of speech can be distinguished from one another, leading to relatively efficient transmission of speech information.

Most music shares these same basic features, with spectral parameters encoding pitch, melody, and tonal aspects of timbre, and time-varying parameters encoding rhythm, and

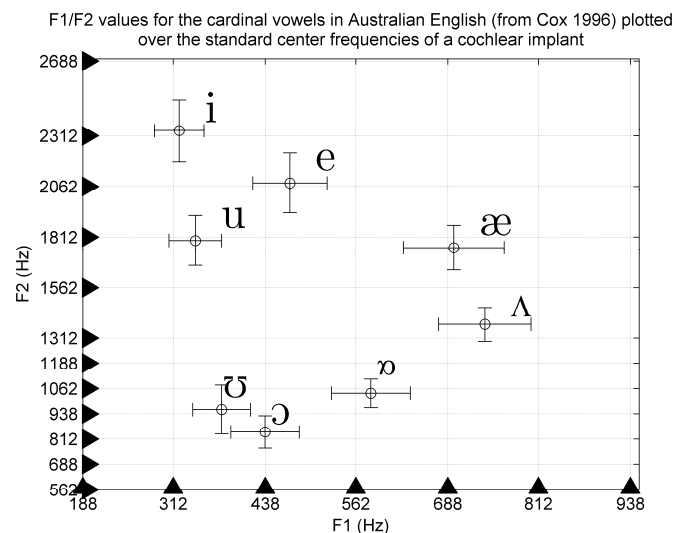


Figure 1 Eight cardinal vowel sounds are plotted according to the frequency of their first and second formants (F_1 , F_2) measured by Cox (1996) in male Australians. The grid overlaid corresponds with the boundaries of the lowest six (on the F_1 axis) and eleven (on the F_2 axis) electrode frequency bands specified in a default CI map (the exact boundaries shift slightly for different individual fittings).

impulsiveness aspects of timbre. However, musical signals are acoustically more complex than speech. The frequency map for the entire electrode array only covers a portion of the upper half of the keys on a standard piano, for example. As we have seen, the signal processing employed in most standard CIs destroys many of the acoustic parameters in the signal, only passing the smoothed amplitude envelopes of a series of band pass filters. This has a number of effects on music perception.

1) Pitch

The perception of pitch is based largely on the fundamental frequency (F0) of an acoustic signal. It is not completely clear how pitch is coded in the auditory system, but research so far points to the conjunction of three physiological cues. First, as described in section 1C, different auditory nerves are stimulated depending on the frequency of the acoustic signals. Therefore, frequency information can be transmitted to the brain by detecting which auditory nerves have been activated. This cue is called *place coding*. Second, the basilar membrane within the cochlea resonates, and therefore triggers the auditory nerves at a rate related to the input frequency (at least up to about 1-4 kHz). This temporal pattern of neural firing can also convey pitch information. This is called *temporal coding*. Third, as the high frequencies excite a portion of the membrane located at the entrance of the membrane, and the low frequencies a portion at the end of the membrane, the delay of excitation will be different according to the frequency – the low frequencies will be delayed by the time needed to travel along the cochlea. Therefore, pitch information can also be conveyed through the timing of activation of the nerves (the high frequencies will arrive first). This is called *phase coding*.

In current sound processing strategies, pitch information is for the most part conveyed via place coding, as different electrodes are activated according to the frequency. However, as stated above, only 22 electrodes are present, so the frequency resolution is limited. It might be possible in the future to introduce more electrodes; however, due to the spread of current, it is unclear if this will improve the frequency resolution. In most current CI recipients, the pulse rate is fixed to 900 Hz, therefore no temporal cues above less than half this frequency (around 300 Hz) can be accurately transmitted. It is possible to increase the pulse rate, however this does not appear to improve pitch perception (Vandali, Whitford, Plant, & Clark, 2000), but does decrease battery life. Finally, in most current strategies the phase delay is not implemented, so recipients cannot benefit from this cue. Experimental strategies have been tested to determine whether the addition of a phase delay will improve speech perception. Results have found small but significant improvement in speech in noise (Bailet et al., 2001; Taft, Grayden, & Burkitt, 2010).

In summary, CIs only partially convey two out of the three main pitch cues. This explains their poor results in pitch discrimination tasks. A study has shown that most CI recipients could not reliably identify a pitch direction change below three semi-tones or that only 20% could identify a well know melody without rhythm cues (Gfeller et al., 2007).

2) Timbre

Timbre is a complex perceptual quality that can be further decomposed into multiple dimensions such as brightness and impulsiveness, which are in turn correlated with the spectral

centroid and attack time respectively (Marozeau, de Cheveigne, McAdams, & Winsberg, 2003). As with the perception of speech, the CI effectively transmits gross amplitude envelope variations, such as attack time or tremolo but less effectively transmits fine spectral cues. In a recent experiment, Kang et al. (2009) asked 42 CI listeners to identify the instrument played in 3 second recordings among a set of 12 possible instruments. Participants were able to correctly identify the instrument 45% of the time (compared to 87% of the time for normal hearing listeners). It is interesting to note for example that CI listeners often confused the flute and cello because they have similar amplitude envelope fluctuations despite having different spectral content.

3) Consonance/Dissonance

Dissonance, and its eventual resolution into consonant intervals, is a very common tool used by composers to shape the listeners emotional experience. As with timbre, the sensory perception of consonance and dissonance is driven by fine details in the spectral content of the acoustic signal. When a consonant dyad is played (such as a fifth), the harmonics of each note are either widely spaced or exactly matching. This spacing allows neurons responsive to a number of harmonically related pitches as well as each fundamental to resolve each harmonic. In the case of dissonant intervals such as the minor 2nd, the harmonics are much more closely spaced, and are not all resolved. This causes a time-varying modulation of the neurons involved, and a sensation of ‘roughness’ or dissonance results (Tramo, Cariani, Delgutte, & Braida, 2003). For listeners with a CI, the electrode spacing interferes with this process, and it is unclear if they perceive the same magnitude of pitch interval. For example, in normal hearing listeners, an octave is clearly heard with the same chroma, and is perceived as highly consonant. For listeners with a CI, it is likely that they perceive a different interval, and a different chroma. Therefore, when a composer attempts to resolve a melody, this may be perceived as dissonant and unresolved by a CI recipient.

4) Rhythm

On the positive side, many studies have showed that listeners with a CI can detect rhythm differences just as well as normal hearing listeners. This is due once more to the accurate reproduction of the amplitude envelope pattern. Therefore, they can recognize melodies based mainly on rhythmical cues (Cooper, Tobey, & Loizou, 2008).

5) The effect of visual and other sensory cues on music enjoyment.

Prior to the advent of recording and playback devices in the last 100 or so years, music was experienced as a live and multi-sensory event, with visual, tactile, and possibly even olfactory contributions (Thompson, Graham, & Russo, 2005). The perception of sound sensations such as pitch, timbre, consonance/dissonance, and location are largely driven by parameters of the acoustic signal. However, there are some aspects of auditory perception that can be influenced by signals in other sensory modalities, such as vision and touch.

The power of visual cues to improve auditory perception has long been known, particularly in the case of speech perception in background noise. When a speaker’s lip and facial movements are visible, an improvement in performance

equivalent to increasing the signal-to-noise ratio by up to 15 dB has been observed (Sumbly & Pollack, 1954). Visual stimuli can also affect perception in other auditory tasks. For instance, presentation of a visual stimulus can increase the perceived loudness of white noise (Odgaard, Arieh, & Marks, 2003, 2004), and discriminations of pitch and loudness improve when presentation of a concurrent visual stimulus matches the features of the sound (Marks, Ben-Artzi, & Lakatos, 2003).

There is now a large body of literature in neuroscience describing how congruent audio-visual stimuli of many different types are recognised faster and detected more accurately at near-threshold levels than either the visual or auditory stimulus alone (for a recent review see Alais, Newell, & Pascal, 2010). Visual information has also been shown to influence stream segregation in normally-hearing listeners (Marozeau, Innes-Brown, Grayden, Burkitt, & Blamey, 2010) as well as CI users (Innes-Brown, Marozeau, & Blamey, 2011).

Visual information has also been found to influence aspects of music appreciation, such as measures of tension and phrasing (Vines, Krumhansl, Wanderley, & Levitin, 2006), skin conductance responses during music listening (Chapados & Levitin, 2008), and bowing vs plucking judgements for stringed instruments (Saldaña & Rosenblum, 1993). Gestures made by the performers have also been shown to influence the perceived duration of notes played by percussion instruments (Schutz & Lipscomb, 2007), and ratings of expressiveness and interest in marimba players are higher when the performers use 'projected' vs 'deadpan' performance styles (Broughton & Stevens, 2009).

D. Summary of data on music perception of CI users

As we have seen above, CI users may not perceive pitch as the composer intends, may not experience the sensations of consonance and dissonance in the expected way, and may have trouble differentiating between different instruments. In all, these are serious problems for listening to and enjoying music. On the other hand, vocals and rhythm are perceived relatively well, and visual or other sensory input may serve to increase the enjoyment of music, particularly when performed live.

II. Development of the project

Six composers were chosen to write new works for the project. As detailed in the previous sections, many of the usual musical tools that composers use in their work (melody, harmony) are not reliably transmitted by CIs, so it was important that the composers involved were able to work outside the usual conventions. For this reason, the composers chosen were from the experimental, contemporary classical domain. The composers chosen were Natasha Anderson, who combines an interest in the performance of medieval and baroque music with contemporary electro-acoustic composition; pianist and co-founder of the Golden Fur ensemble James Rushford; Rohan Drape, whose work takes in computer music, instrumental composition and installation; percussion specialist Eugene Ughetti; and Ben Harper, whose interest in microtonal tuning systems made him an ideal candidate for a project that could involve alternate tunings. Robin Fox's interests include audio-visual synchrony and multi-channel audio diffusion.

III. Tools and methods - Composition

E. Papers

After the final list of composers was confirmed, the first discussions concerned the motivations of the project and the technical operation of CIs. After this initial meeting, there was a period of reading and email discussions, where a series of review and research articles were circulated to the composers. These included general review articles (Donnelly & Limb, 2009; Gfeller et al., 2000; Gfeller & Knutson, 2003; Gfeller et al., 2005; McDermott, 2004; McDermott & McKay, 1997), articles on the technical aspects of how CIs operate (Clark, 2009; Middlebrooks, Bierer, & Snyder, 2005; Wilson & Dorman, 2008; Zeng, 2004), and more specific articles addressing issues raised by some of the composers, such as the effect of visual cues (Boltz, Ebendorf, & Field, 2009; Chapados & Levitin, 2008), gestures (Saldaña & Rosenblum, 1993), tactile cues (Balliet, Mosher, & Leahy, 2001; Candia, Rosset-Llobet, Elbert, & Pascual-Leone, 2005), and melody segregation (Innes-Brown, Marozeau, Grayden, Burkitt, & Blamey, 2010; Marozeau et al., 2010).

F. Meetings

After the initial email phase, several meetings between the composers and two scientists (authors JM and HIB) took place at the Bionics Institute. The scientists began with a crash course on the auditory system and the CI. The purpose of these meetings was to consolidate the information given in the reading phase, and give the composers some feedback on their initial thoughts on how to approach the composition.

Following these meetings, the composers spent a period of time in their studios working on auditory and compositional experiments. In the third phase the size of the meetings grew again, with a group of CI users attending to give further feedback on the sounds and compositional experiments. This last series of meetings was vitally important for the success of the project. Driven by the prior meetings, the composers had specific theories concerning what types of sounds or musical elements might be best perceived by the CI users. They brought along various equipment (laptops, speakers, instruments) and conducted sound-based experiments with the CI users using elements of their in-progress compositions. The results from these experiments then drove the final compositions.

G. Sonifications

As well as the interactions with CI users, the composers had an additional tool – a CI acoustic simulator, combined with a simple musical sequencer. This simulator/sequencer was a piece of software written by Robin Fox and the BI staff. It is very difficult to understand and to reproduce the actual perception of sound through an implant. However, it is possible to *sonify* the electrical output of the sound processor. By doing so the composers could hear which musical cues were transmitted and which one were discarded by the sound-processing algorithms. The sonification software operated in real time, and was based on the operation of the standard sound processing strategy, with 22 band-pass filters followed by an envelope extractor. The output of this envelope was finally multiplied with narrow-band noise, filtered with the same parameters as the signal band-pass filters. Connected to the sonification section was a sequencer, which allowed the

composers to test a variety of tunings and note amplitude envelopes. The software could also operate on a direct audio input, allowing those composers who were not using computer-based audio to test sounds from any other source.

The auditorium in which the works were performed featured an 11.1 channel audio diffusion system. Speakers were placed on stage and suspended above the seating area in order to create 11 audio channels through which composers could create spatialized soundscapes. Pre-recorded audio was played through this system using 11-channel audio files, and the signals from microphones mounted on live instruments on the stage could also be routed selectively through any audio channel. Each composer combined aspects of their own artistic practice with the knowledge obtained from the learning program to develop their composition. The six resulting pieces varied widely in their approach and style as briefly reviewed below.

Piece 1 - Variations: The composer explored the contours and intervals created by using dynamics, duration, pitch, and repeated variations on a simple theme. Instrumentation: pre-recorded piano with live clarinet, viola and cello.

Piece 2 – Percussion/Vibraphone: Percussion instruments were chosen based on feedback that CI users found them to be easy instruments to distinguish. A bowed vibraphone was used to add a controlled pulse to the sound produced. Synthesiser bursts were also used to create sound textures that could be interpreted by the CI user. The live performance enabled listeners to make use of visual cues provided by the movements of the performers, while pre-recorded material was used to create spatialized sound from the 11.1 channel audio diffusion system in the auditorium. Instrumentation: live percussion and bowed vibraphone, spatialized diffusion of pre-recorded piano, cello, and synthesised sound.

Piece 3 – Spoken Word: Based on the well-established speech processing capabilities of the CI, this piece primarily used fragments of spoken phrases superimposed over short electronic melodies that emulated the cadence and prosody of the speech fragments. Instrumentation: pre-recorded voice with spatially-diffused synthesised keyboard/vibraphone using tuning system based on CI sound-processor channel frequencies.

Piece 4 - Pitch: The composer created musical texture by using multiple lines of melody in a trio for cello, viola, and tape-recorder. Extended technique and preparation of live instruments introduced differences in the instruments’ timbre and attack, sustain and release times. Instrumentation: live cello and viola, pre-recorded processed tape recordings using the above instruments.

Piece 5 - Electronic: Pulse, rhythm and tone were separately explored in this three-part piece. Study (1): spatialisation of a constant pulse was achieved via the 11.1 channel diffusion sound system. Study (2): rhythmic patterns were created using the centre frequencies of the 22 filters present in a CI. Study (3): chords were generated by gradually introducing single tones through different channels of the 11.1 channel sound system. Finally, screens with visual effects were used throughout the piece to provide synchronised audio-visual cues.

Piece 6 – Percussion: Voice and percussion instruments were exclusively used in this piece, taking advantage of the implant’s ability to convey amplitude envelope fluctuations.

Familiar rhythms were overlaid and varied. Instrumentation: large percussion ensemble with three players.

The aim of this study was to determine whether CI users and NH listeners might, for the first time, report similar ratings on cognitive, engagement and technical aspects of the music, as measured using subjective rating-scale items. If CI users are sensitive to the measured dimensions of the commissioned works then there should be no difference between NH listeners and CI users in their ratings for each piece of music. As CIs transmit impulsiveness cues relatively well, we additionally predicted that CI users would assign relatively positive ratings to pieces with percussion instruments compared to those without.

IV. Tools and methods – concert survey data

H. Participants

In the lead up to the two concerts, invitations were sent to implantees through the cochlear implant Clinic at the Royal Victorian Eye and Ear Hospital in Melbourne, and details of the concerts were made available on the Bionics Institute website and the Arts Centre Melbourne website. Participants in this study were audience members who attended the concerts; a mix of hearing aid users, CI users and NH listeners. Questionnaires were distributed to the 588 people who attended; of those, 407 participants returned a completed questionnaire at the end of the performances. Due to the substantially larger sample size of NH participants ($n = 301$) compared to the CI users ($n = 44$), a subsample of 44 NH participants was selected to be part of the analyses, and the two groups were matched on the variables of age range (*Median range* = 35–50 years) and musical ability ($M_{NH} = 1.7$; $M_{CI} = 1.6$)¹ as closely as possible. An additional hearing aids group was excluded from the analyses due to the small sample size ($n = 13$).

TABLE 1. Summary of demographic data for subsample of NH listeners and CI users.

	NH Group	CI group
N (females)	44 (22)	44 (27)
Median age group (years)	51-70	51-70
Musical ability	1.7	1.6
Unilateral CI	NA	34
Bilateral CI	NA	10

I. Materials

An audience response questionnaire was developed to collect quantitative, qualitative and demographic data, the design of which was based on the Audience Response Tool (Glass, 2005, 2006; Stevens, Glass, Schubert, Chen, & Winskel, 2007). Biographical data were collected for age, gender, musical ability, hearing impairment, music enjoyment before and after impairment, and type of hearing amplification used. For each of the six pieces of music, 16 items were scored on a seven-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). In the present study, six of these items were chosen for analysis, measuring cognitive response,

¹ Musical ability was self-rated on a scale ranging from 1 (no musical ability) to 5 (performance-level musical ability).

engagement and technical aspects of the music, such as localisation and timbre recognition. An open-ended question asking participants to record their thoughts and reactions was also included for each piece of music.

J. Procedure

Two concerts were held at the Arts Centre Melbourne on 13 February, 2011, both containing the same material performed in the same order. At the beginning of each concert, members of the audience were provided with the questionnaire and instructed to complete the relevant section immediately after each piece. The questionnaires were collected at the end of each concert. Data from both concerts were combined for analysis. The full questionnaire can be downloaded: <http://www.bionicsinstitute.org/interiordesign/Survey.pdf>

V. Results

The mean ratings for all 16 questionnaire items for CI users (grey lines) and NH listeners (black lines) are shown in Figure 2. As explained below, no further analysis was performed using the mean ratings. They are provided to show that overall mean ratings for most questions were similar for both CI users and the NH listeners.

K. Statistical analysis of selected items

Statistical analyses were carried out using SPSS 17. Six items from the questionnaire were chosen for analysis. These items were selected in order to explore three main areas of interest: *cognitive response to the music* (as measured by items 1 and 2, addressing interest and understanding), *engagement with the music* (measured by items 7 and 8, asking whether the concert was musical and enjoyable), and *technical aspects of the music* (measured by items 15 and 16, addressing localisation and timbre recognition ability). As the rating data was ordinal, and not normally-distributed, non-parametric analyses were performed. Boxplots of the median, inter-quartile range, and range of responses for each of the 6 items are shown in Figure 3. In order to test for the overall effect of group (CI users vs NH listeners), Mann-Whitney tests were performed on the mean responses across pieces for each of the six items analysed. To test for the effect of each piece, Friedman ANOVAs were run for each of the six questionnaire items analysed. Significant main effects were followed up by post-hoc Wilcoxon signed rank tests. In order to test the hypothesis that pieces with percussion would have higher ratings than those without, only the eight combinations including the two percussion pieces were tested. A Bonferroni correction was applied, so these pairwise tests were reported as significant when $p < .006$. The open-ended questions will require further thematic analyses which are beyond the scope of this paper.

6) Cognitive response to the music

Item 1 Interest: 'The piece was very interesting': Figure 3A shows the median ratings for the 6 pieces and the two groups. There was no significant effect of group, however the main effect of piece was found to be significant, $\chi^2(5) = 106.2$, $p < .001$. Post hoc tests revealed that the ratings for interest were significantly higher for the two percussion pieces (2 and 6) than each of the other four pieces.

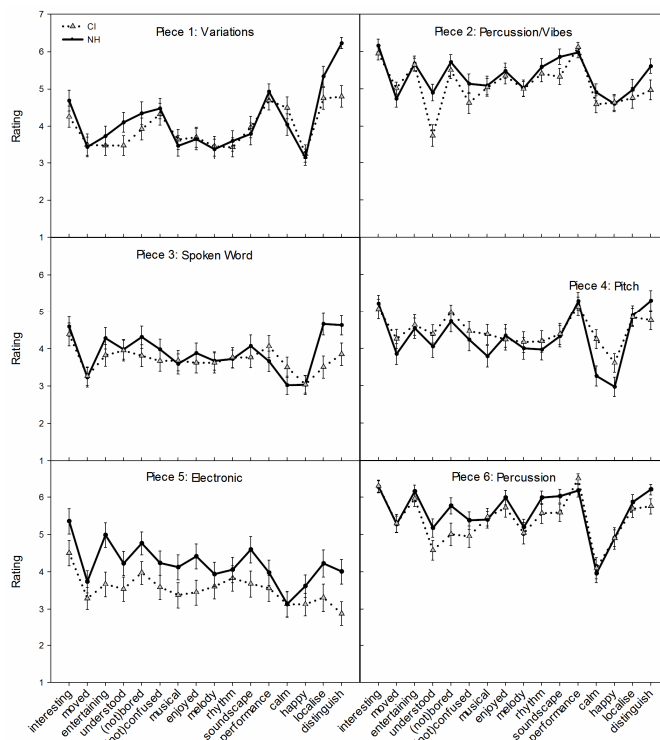


Figure 2. Mean ratings from the 16 survey items for each piece.

Item 4 Not-understanding: 'I did not understand the piece':

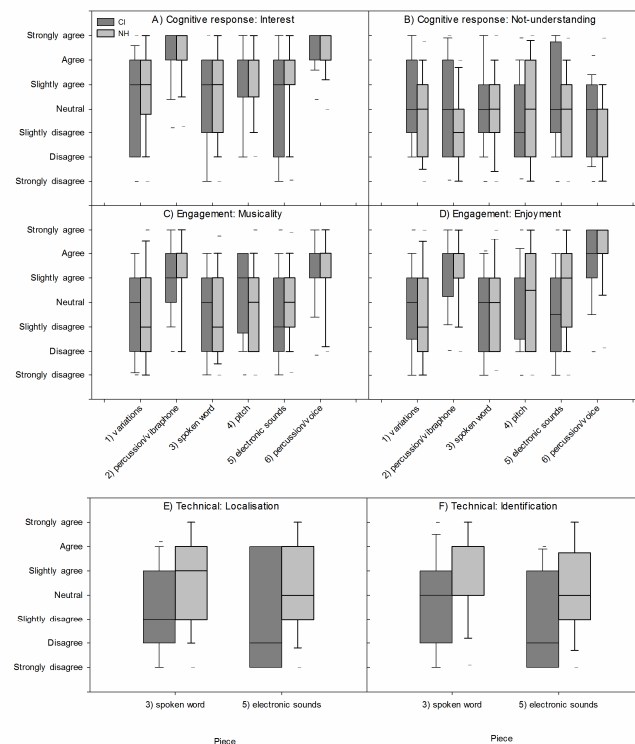


Figure 3. Boxplots showing the distribution of responses from cochlear implant users (dark grey – CI) and normally-hearing listeners (light grey – NH) for the six items analysed, separated by piece. Boxes show the median, and 25th and 75th percentiles. Whiskers show the 10th and 90th percentiles. Small horizontal lines indicate the minimum and maximum values. A) Item 1 – Interest, B) Item 4 – Not-understanding, C) Item 7 – Musicality, D) Item 8 – Enjoyment, E) Item 15 – Localisation, F) Item 16 – Identification.

In order to make sure that the participant read each questions carefully, some negative assertions were added. Therefore positive responses were located on the negative side of the scale (*Strongly disagree*). Significant effects were found for group, $U = 738.5, p = .05$, and piece, $\chi^2(5) = 25.8, p < .001$. The NH group (*Mean Rank* = 39.1) had a better understanding (lower not-understanding ratings) of the music than the CI group overall (*Mean Rank* = 49.0). Post hoc tests revealed that the ratings for interest were significantly higher for the two percussion pieces (2 and 6) than each of the other four pieces (Figure 1B).

7) Engagement with the music

Item 7 Musicality: 'I found the piece very musical': The main effect of piece was found to be significant, $\chi^2(5) = 65.4, p < .001$, but not the main effect of group. Post hoc tests revealed that the ratings for interest were significantly higher for the two percussion pieces (2 and 6) than each of the other four pieces (Figure 1C).

Item 8 Enjoyment: 'The piece was very enjoyable': The main effect of piece was found to be significant, $\chi^2(5) = 84.0, p < .001$, but not the main effect of group, or the group x piece interaction. Post hoc tests revealed that the ratings for interest were significantly higher for the two percussion pieces (2 and 6) than each of the other four pieces (Figure 1D).

8) Technical aspects of the music

Item 15 Localisation: 'I can tell where all the sounds were coming from': The electronic piece and spoken word piece used only sampled recordings, while the other four pieces all featured live instruments on stage. As this item was a measure of the ability of participants to localise, it was not useful to include pieces with and without the visual cues generated by the instruments on stage. Thus, only the electronic piece and the spoken word piece were included in the analysis. There was a significant main effect of group, $U = 589.0, p = .02$, indicating that the NH group (*Mean Rank* = 47.3) reported to be able to better localise sounds than the CI group (*Mean Rank* = 35.1). There was no significant main effect of piece for this item. It is important to note, however, that responses related to localisation may have been affected by the location of each respondent in relation to the instruments and elements of the sound diffusion system, a factor which was not able to be controlled for.

Item 16 Timbre recognition 'I can distinguish different instruments': Again, only the electronic and spoken word pieces were included to avoid confounds related to the presence of instruments on stage. There were significant main effects of group, $U = 580.0, p = .04$. The NH group ratings (*Mean Rank* = 45.5) were significantly higher than the CI group (*Mean Rank* = 34.7).

VI. Discussion

For items measuring interest, musicality, and engagement, there was no difference in the ratings made by NH listeners and CI users. Where the main effect of group was found to be significant, however, the median rating of NH listeners was invariably more positive than that of the CI group. Significant differences for piece were found across all three areas of cognitive response, engagement, and technical aspects; and the ratings from both groups in these areas were typically higher

for the pieces involving percussion when all six pieces were part of the analysis.

L. Cognitive response to the concert

Measures of interest and understanding were used to explore this category. For the item measuring interest, there was no significant difference in the mean rating provided by CI users and NH listeners, suggesting that CI users were just as interested in the music presented as the NH listeners. This result is encouraging, and indicates that CI users were able to perceive the music sufficiently well to engage their interest. However, despite equivalent levels of interest, the group effect for the item measuring understanding suggests that inherent differences between the two groups enabled the NH listeners to have a better understanding of the music than the CI users. This could be due to limitations of the CI preventing it from transmitting musical features such as pitch and timbre that would help facilitate better understanding. Another reason could be attributed to the musical preference of the audience as a confounding variable. While the CI users were invited to attend the concert, many NH participants may have chosen to attend based on their existing interest in the music to be performed. It could thus be reasoned that those NH listeners who attended were perhaps more experienced listeners of contemporary music and were better able to understand the music than were CI users.

The significant effect of piece for both interest and understanding in this category indicates that some pieces were more successful than others in facilitating cognitive appreciation of the music. In particular, the ratings for interest and understanding were higher for the percussion/voice piece and the percussion/vibraphone piece. This is an interesting finding in the context of CIs because from what is known of the CI, the amplitude envelope fluctuations that convey rhythm are preserved quite well, as are the distinctive impulsiveness characteristics of percussion instruments. Taken together, it is not surprising that the two percussion pieces were found to be the most interesting and the best understood by the CI group.

M. Engagement with the music

This category was measured based on ratings of how *musical* and *enjoyable* each piece was. The non-significant effect of group for these two items indicates that NH participants and CI users had similar levels of engagement with the music. This is a promising result for one of the main aims of this study, which was to develop new music that could be appreciated in the same way by both NH listeners and CI recipients. Whether the participants liked or disliked each piece, there were no significant differences between CI users and NH listeners in their ratings of enjoyment and musicality for each piece of music.

The significant effect of piece for these two items shows that pieces differed in how they engaged the audience. Ratings for the musicality and enjoyment items were higher for the percussion pieces than the others. This is consistent with general understanding of the CI's strengths in relation to retaining amplitude envelope cues, and lends support to the experimental hypothesis that percussion instruments would be well perceived by CI users.

N. Technical aspects of the music

Instrument localisation and timbre recognition were used to explore the technical aspects of music. The significant effect of group for both items in this category indicates that NH listeners report they were better able to perceive the more technical elements of localisation and timbre. Localisation depends on the processing of differences in level and timing between the two ears. This result is therefore not surprising, considering that the CI users in the current study were mostly unilaterally implanted (Table 1).

Timbre recognition tasks are generally performed less accurately by implantees, owing to the implant's imprecise transmission of spectral content. The significant main effect of group for the timbre item suggests that implantees were not able to distinguish between instrument sounds in the same way as the NH listeners. In general, the instrument identification ratings were lower for the electronic piece than the piece using spoken word, although this effect was not statistically significant. This could be attributed to the type of instruments represented in the sampled recordings: whereas the spoken word piece used conventional instruments, such as synthesised keyboard notes, the electronic piece employed potentially unfamiliar computer-generated sounds that may have been difficult to recognise in a conventional sense.

O. Further research

The works presented at the concert were written by composers of contemporary experimental music. These particular composers were chosen as they were able to work in an experimental manner – by iteratively proposing and testing works with CI users during the composition process. Although this unique process resulted in works which were interpreted similarly by CI users and NH listeners, the resulting works also contained musical structures and sounds which may have been unfamiliar. Throughout the composition process, the composers learnt that many musical 'building blocks,' such as harmony and melody based on small intervals could not be expected to be perceived in similar ways by CI users and NH listeners. Thus, the works were based around alternative sonic structures, such as radical shifts in timbre and dynamics, repetition of melody, the use of vocal elements, alternative tuning systems and synthesised sounds, and substantial use of rhythm. Some of these structures, such as rhythm, may have been familiar to both groups in the audience, whereas others, such as the alternative tuning systems, melodies with unusually large intervals, and synthesised sounds, may have been unfamiliar to the majority of the audience.

In addition, the works were by definition new – none of the audience had previously heard any of the works. Familiarity has long been known to affect judgements of enjoyment and the emotional response to music (Schubert, 2010). Repetition of unfamiliar works can lead to increases in the enjoyment of individual pieces with initially low ratings (Mull, 1957; Peretz, Gaudreau, & Bonnel, 1998), and familiar pop songs elicit greater activity in emotion-related limbic brain areas than unfamiliar songs (Pereira et al., 2011). Thus, the generation of new works and deliberate avoidance of often-used musical structures resulted in highly unfamiliar music. In the current study, the two works based heavily on rhythm generally had the highest ratings of enjoyment, possibly reflecting the audiences' familiarity with the rhythmic musical structures. In future research, comparing responses from CI users and NH

listeners to repeated presentations of the works over several weeks (as in the study by Mull (1957)) would help disentangle the effects of familiarity from the hypothesised effect of the specifically-composed music itself. Future research may also benefit from a comparison of the works from the current study with other non CI-specific works from the same composers.

P. Conclusion

The aim of the current study was to investigate the reception of new music that was intended to be interpreted and appreciated by both NH listeners and CI recipients. The results indicate that at least in terms of engaging with the audience, this was a success. CI users gave higher ratings on measures of interest, engagement and musicality for the pieces with percussion instruments. Overall, however, NH participants typically rated all items higher than CI users, and the effects of group were large when significant, in particular for localisation and instrument identification. This suggests that, for now, CI technology is still unable to deliver a complete musical experience to CI users. More novel methods of circumventing the limitations of CI technology may form the basis for future studies of this kind, with particular focus on preserving the technical elements of music. The findings also suggest that while technological changes are necessary to improve musical experiences for CI users, composers and performers also have an important role to play.

ACKNOWLEDGMENT

Thanks are due to Robin Fox, Rohan Drape, Ben Harper, Natasha Anderson, James Rushford, and Eugene Ughetti - the composers who contributed much of their time and effort to the project. Musical director Robin Fox was also involved in the project development. Many thanks to all the participants of this study, particularly to those CI users who volunteered their time to give feedback to the composers in the sound-testing phase. Special thanks to Dr Tom Francart, Mr Kyle Slater, Dr Dean Freestone, Ms Aimee Clague, and Ms Rebecca Argent who helped with wrangling surveys and data collection on the night of the concert. Prof Richard Dowell had significant input into statistical design and an earlier version of the manuscript. The authors gratefully acknowledge the financial support of the Music Board of the Australia Council for the Arts, Arts Victoria, the Arts Centre Melbourne, Blamey & Saunders Hearing, Arts Access Australia, the Cochlear Foundation, and the Bionics Institute. The Bionics Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastructure Support Program. This study was part of AA's final year research project for the completion of her Master of Clinical Audiology at the University of Melbourne, supervised by Dr Jeremy Marozeau and Prof Richard Dowell. HIB's attendance was funded by the Harold Mitchell Travelling Fellowship 2012.

REFERENCES

- Alais, D., Newell, F., & Pascal, M. (2010). Multisensory Processing in Review: from Physiology to Behaviour. *Seeing and Perceiving*, 23, 3-38.
- Baillet, S., Riera, J., Marin, G., Mangin, J., Aubert, J., & Garnero, L. (2001). Evaluation of inverse methods and head models for

- EEG source localisation using a human skull phantom. *Physics in Medicine & Biology*, 46(1), 77-96.
- Balliet, S., Mosher, J. C., & Leahy, R. M. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 18(6), 14-30.
- Blamey, P. J., Dowell, R. C., Brown, A. M., Clark, G. M., & Seligman, P. M. (1987). Vowel and consonant recognition of cochlear implant patients using formant-estimating speech processors. *Journal of the Acoustical Society of America*, 82(1), 48-57.
- Boltz, M., Ebendorf, B., & Field, B. (2009). AUDIOVISUAL INTERACTIONS: THE IMPACT OF VISUAL INFORMATION ON MUSIC PERCEPTION AND MEMORY. *Music Perception*, 27(1), 43-43.
- Broughton, M., & Stevens, C. (2009). Music, movement and marimba: an investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*, 37(2), 137-153.
- Candia, V., Rosset-Llobet, J., Elbert, T., & Pascual-Leone, A. (2005). Changing the Brain through Therapy for Musicians' Hand Dystonia. *Annals of the New York Academy of Sciences*, 1060(The Neurosciences and Music II: From Perception to Performance), 335-342.
- Chapados, C., & Levitin, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition*, 108(3), 639-651.
- Clark, G. (2009). The multi-channel cochlear implant: Past, present and future perspectives. *Cochlear Implants International*, 10(S1), 2-13.
- Cooper, W. B., Tobey, E., & Loizou, P. C. (2008). Music Perception by Cochlear Implant and Normal Hearing Listeners as Measured by the Montreal Battery for Evaluation of Amusia. *Ear and hearing*, 29(4), 618-626-618-626.
- Cox, F. (1996). Vowel change in Australian English. *Phonetica*, 56(1-2), 1-27.
- Dettman, S. J., Pinder, D., Briggs, R. J. S., Dowell, R. C., & Leigh, J. R. (2007). Communication development in children who receive the cochlear implant younger than 12 months: Risks versus benefits. *Ear and Hearing*, 28(2), 11-18S.
- Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual review of psychology*, 55, 149-179.
- Donnelly, P. J., & Limb, C. J. (2009). Music Perception in Cochlear Implant Users. In J. K. Nipakp (Ed.), *Cochlear Implants. Principles and Practices*. 2nd ed. Philadelphia: Lippincott Williams & Wilkens.
- Eddington, D. K. (1980). Speech discrimination in deaf subjects with cochlear implants. *Journal of the Acoustical Society of America*, 68(3), 885-891.
- Gfeller, K., Christ, A., Knutson, J. F., Witt, S., Murray, K. T., & Tyler, R. S. (2000). Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *Journal of the American Academy of Audiology*, 11(7), 390-406.
- Gfeller, K., & Knutson, J. F. (2003). Music to the impaired or implanted ear. Psychosocial implications for aural rehabilitation. Retrieved July 2011, from <http://www.asha.org/publications/leader/archives/2003/q2/f030429a.htm>
- Gfeller, K., Olszewski, C., Rychener, M., Sena, K., Knutson, J. F., Witt, S., & Macpherson, B. (2005). Recognition of "Real-World" Musical Excerpts by Cochlear Implant Recipients and Normal-Hearing Adults. *Ear and Hearing*, 26(3), 237-237.
- Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., & Olszewski, C. (2007). Accuracy of Cochlear Implant Recipients on Pitch Perception, Melody Recognition, and Speech Reception in Noise. *Ear and Hearing*, 28(3), 412-423.
- Glass, R. (2005). Observer response to contemporary dance. In R. Grove, Stevens, C., & McKechniw, S. (Ed.), *Thinking in four dimensions: Creativity and cognition in contemporary dance* (pp. 107-121). Melbourne: Melbourne University Press.
- Glass, R. (2006). *The Audience Response Tool (A.R.T.): The impact of choreographic intention, information and dance expertise on psychological reactions to contemporary dance.*, University of Western Sydney, Sydney.
- Helms, J., Weichbold, V., Baumann, U., von Specht, H., Schon, F., Muller, J., . . . D'Haese, P. (2004). Analysis of ceiling effects occurring with speech recognition tests in adult cochlear-implanted patients. *ORL; Journal of Oto-Rhino-Laryngology and Its Related Specialties*, 66(3), 130-135.
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J Acoust Soc Am*, 97(5 Pt 1), 3099-3111.
- Innes-Brown, H., Marozeau, J., & Blamey, P. (2011). The effect of visual cues on difficulty ratings for segregation of musical streams in listeners with impaired hearing. *PLoS one*, 6(12), e29327.
- Innes-Brown, H., Marozeau, J., Grayden, D. B., Burkitt, A. N., & Blamey, P. (2010). Visual cues can improve musical stream segregation for cochlear implant users. *Clinical EEG and Neuroscience*, 41(2), 108.
- Kang, R., Nimmons, G. L., Drennan, W., Longnion, J., Ruffin, C., Nie, K., . . . Rubinstein, J. (2009). Development and Validation of the University of Washington Clinical Assessment of Music Perception Test. *Ear and hearing*, 30(4), 411-418.
- Loizou, P. C., Hu, Y., Litovsky, R., Yu, G., Peters, R., Lake, J., & Roland, P. (2009). Speech recognition by bilateral cochlear implant users in a cocktail-party setting. *Journal of the Acoustical Society of America*, 125(1), 372-383.

- Marks, L. E., Ben-Artzi, E., & Lakatos, S. (2003). Cross-modal interactions in auditory and visual discrimination. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, 50(1-2), 125-145.
- Marozeau, J., de Cheveigne, A., McAdams, S., & Winsberg, S. (2003). The dependency of timbre on fundamental frequency. *Journal of the Acoustical Society of America*, 114(5), 2946-2957.
- Marozeau, J., Innes-Brown, H., Grayden, D. B., Burkitt, A. N., & Blamey, P. J. (2010). The effect of visual cues on auditory stream segregation in musicians and non-musicians. *PLoS One*, 5(6), e11297.
- McDermott, H. J. (2004). Music Perception with Cochlear Implants: A Review. *Trends in Amplification*, 8(2), 49-82.
- McDermott, H. J., & McKay, C. M. (1997). Musical pitch perception with electrical stimulation of the cochlea. *Journal of the Acoustical Society of America*, 101(3), 1622-1631.
- Middlebrooks, J. C., Bierer, J. A., & Snyder, R. L. (2005). Cochlear implants: the view from the brain. *Current Opinion in Neurobiology*, 15(4), 488-493.
- Mull, H. K. (1957). The effect of repetition upon the enjoyment of modern music. *Journal of Psychology*, 43, 155-162.
- Odgaard, E. C., Arieh, Y., & Marks, L. E. (2003). Cross-modal enhancement of perceived brightness: sensory interaction versus response bias. *Perception & Psychophysics*, 65(1), 123-132.
- Odgaard, E. C., Arieh, Y., & Marks, L. E. (2004). Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. *Cognitive, Affective and Behavioral Neuroscience*, 4(2), 127-132.
- Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: familiarity matters. *PLoS one*, 6(11), e27241.
- Peretz, I., Gaudreau, D., & Bonnel, A. M. (1998). Exposure effects on music preference and recognition. *Memory & cognition*, 26(5), 884-902.
- Pols, L. C., van der Kamp, L. J., & Plomp, R. (1969). Perceptual and physical space of vowel sounds. *J Acoust Soc Am*, 46(2), 458-467.
- Saldaña, H. M., & Rosenblum, L. D. (1993). Visual influences on auditory pluck and bow judgments. *Perception & Psychophysics*, 54(3), 406-416.
- Schubert, E. (2010). Affective, evaluative and collative responses to hated and loved music. *Psychology of Aesthetics, Creativity, and the Arts*, 4(1), 36-46.
- Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *PERCEPTION-LONDON*, 36(6), 888-888.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303-304.
- Stevens, C. J., Glass, R., Schubert, E., Chen, J., & Winskel, H. (2007). Methods for measuring audience reactions. *Proceedings of the inaugural International Conference on Music Communication Science*(5-7 December).
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26, 212-215.
- Taft, D. A., Grayden, D. B., & Burkitt, A. N. (2010). Across-frequency delays based on the cochlear traveling wave: enhanced speech presentation for cochlear implants. *IEEE transactions on bio-medical engineering*, 57(3), 596-606.
- Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, 156(1-4), 203-227.
- Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2003). Neurobiology of harmony perception. In I. Peretz & R. Zatorre (Eds.), *The Cognitive Neuroscience of Music*. New York: Oxford University Press.
- Vandali, A. E., Whitford, L. A., Plant, K. L., & Clark, G. M. (2000). Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear and hearing*, 21(6), 608-624.
- Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101(1), 80-113.
- Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research*, 242(1-2), 3-21.
- Zeng, F.-G. (2004). Trends in Cochlear Implants. *Trends in Amplification*, 8(1), 1-34.