Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

What cochlear implants can tell us about pitch perception in normal hearing

Brett A. Swanson (1), Hugh J. McDermott (2)

(1) Cochlear Ltd, Sydney, Australia
(2) Bionic Ear Institute, Melbourne, Australia

PACS: 43.64.Me, 43.66.Ba, 43.66.Hg

ABSTRACT

Temporal models of pitch perception postulate that pitch depends on auditory nerve inter-spike intervals. Such models were assessed with cochlear implant recipients by stimulating via a single electrode with two distinct pulse timing patterns. Pitch-ranking results using on-off modulated pulse trains were not consistent with a published autocorrelation model, but were consistent with an alternative model that analyses first-order inter-spike intervals. Place of stimulation cues were investigated by choosing a fixed pulse rate, high enough to avoid temporal cues (1800 pps), and then varying the electrodes that were activated. Pitch-ranking results suggested that when a group of neighbouring electrodes was activated, the percept depended on the centroid of the stimulation pattern. This is similar to a model of brightness (timbre) perception in normal hearing. Cochlear implant recipients performed a melody perception test with melodies presented by (i) varying the pulse rate on a single electrode (temporal cues only), and (ii) varying the centroid of the stimulation pattern (place cues only). Scores were worse than those of normal-hearing subjects listening to tones containing only resolved harmonics, but similar to those obtained with unresolved harmonics. Such tones contain only temporal cues to pitch and evoke a weak pitch sensation. Given the dissimilarities in the neural firing patterns, it is surprising that place cues alone and temporal cues alone can each provide useful information about melodic pitch.

INTRODUCTION

Researchers are often forced to resort to esoteric acoustic stimuli to probe the finer nuances of pitch perception in normal hearing. Cochlear implants offer an alternative research tool, where place and temporal cues to pitch can be manipulated completely independently, allowing pitch perception models to be tested in ways that are not possible with normal acoustic hearing.

TEMPORAL PITCH

Modern temporal models for pitch perception in normal hearing propose that pitch is dependent on auditory nerve interspike intervals [1]. In one type of temporal model, pitch depends on the intervals between each nerve spike and every other spike (i.e., all-order intervals), which is equivalent to an autocorrelation of the spike train [2],[3]. However, some psychophysical results are inconsistent with the autocorrelation model, and an alternative model has been proposed that utilizes the intervals between successive spikes only; i.e., first-order intervals [4],[5].

These models were assessed with cochlear implant recipients by stimulating the auditory nerve electrically via a single apical electrode with two distinct patterns of pulse timing, as shown in Figure 1. The Single Pulse per Period (SPP) sequence was a pulse train with a constant pulse rate. The Multiple Pulse per Period (MPP) sequence was an on-off modulated pulse train. These stimuli were used in earlier studies [6],[7]

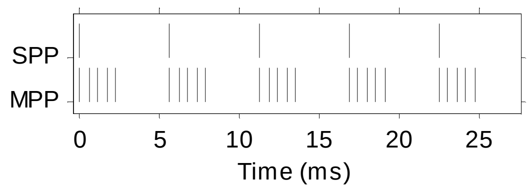

Figure 1. Single Pulse per Period (SPP) and Multiple Pulse per Period (MPP) sequences, both with a repetition frequency of 178 Hz (a period of 5.63 ms). Each pulse is represented by a vertical line. The MPP sequence has a pulse rate within each burst of 1776 Hz, and an inter-burst gap of 3.38 ms.

Method

Seven Nucleus 24 cochlear implant recipients participated in an experiment that aimed to identify a SPP sequence that matched the pitch of a MPP sequence. In each trial, the reference stimulus was a fixed 178 Hz MPP sequence and the comparison stimulus was an SPP sequence, with pulse rate in the range 125 to 400 pps. Each stimulus was 500 ms duration. The subject was asked whether the pitch of the SPP sequence was higher or lower than that of the MPP sequence. After several blocks of trials, the set of SPP frequencies was manually adjusted for each subject to straddle the apparent 50% point of the psychometric function (i.e., where the pitches matched).

Results

The SPP sequence that would theoretically best match the pitch of the 178 Hz MPP sequence was found by fitting a psychometric function to the experimental data, and locating the 50% intercept. Figure 2 shows that, for each subject, the pitch-matched SPP sequence had a period (i.e., inter-pulse interval) that was between the period and the inter-burst gap of the MPP sequence. When the periods were equal, the SPP sequence generally had a lower pitch than the MPP sequence.

These results are consistent with the results of Busby & Clark [7], who found that numerical pitch estimates for MPP sequences (with frequencies in the range 71 to 250 Hz) were either significantly higher than or equal to those for corresponding SPP sequences in six Nucleus 22 recipients. There were no cases where a SPP sequence had a higher pitch than the MPP sequence with equal period.

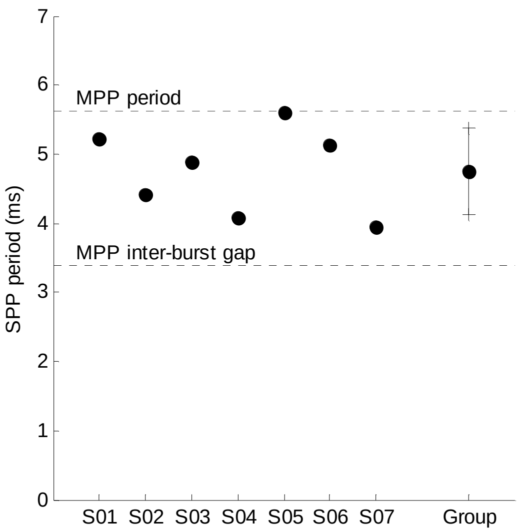

Figure 2. Period of the SPP sequence that matched the pitch of the 178 Hz MPP sequence for seven cochlear implant recipients. The group mean is shown with error bars indicating the standard deviation. The dashed horizontal lines indicate the period and the inter-burst gap of the MPP sequence.

Applying temporal pitch models

The neural responses to the SPP and MPP sequences were simulated following the method of Bruce et al [8]. The loss of inner hair cells results in negligible levels of spontaneous activity in the auditory nerve, so that spikes only occur in response to stimulation pulses. The probability of a nerve firing in response to the *n*th pulse within a train of pulses is modelled by:

$$p(n) = \Phi\left(\frac{I_{stim} - r(n)\, I_{thresh}}{\sigma}\right) \quad (1)$$

where I_{stim} is the stimulus current, I_{thresh} is the threshold current for a single isolated pulse (at which the firing probability is 50%), $r(n)$ is a refractory function, σ is the standard deviation of the threshold noise, and Φ is the cumulative normal distribution. The refractory function $r(n)$ represents a threshold shift, and depends on the time since the last spike. It is infinite during the absolute refractory time (approximately 1 ms), and falls exponentially to one during the relative refractory period.

For a SPP sequence, all inter-spike intervals are multiples of the pulse period, and both the autocorrelation and the first-order interval model predict a pitch equal to the pulse rate.

For a MPP sequence, if the inter-burst gap is long enough for the nerves to fully recover, then the compound neural response to each burst of pulses will be similar. For the 178 Hz MPP sequence, the inter-burst gap was 3.38 ms. Measurements in Nucleus 24 cochlear implant recipients using neural response telemetry (NRT) indicate that the electrically-evoked compound action potential has typically recovered to 80% of its amplitude by this time [9], so the first pulse in each burst will have a relatively high probability of causing a spike. The simulated response to the MPP sequence is shown in Figure 3. The compound response for a population of 1000 fibres is shown in Figure 3c. Following Cariani and Delgutte [2], the autocorrelation of each individual fibre response was calculated, and then summed across the population of fibres (Figure 4). The summed autocorrelation has a peak at the fundamental period (5.63 ms), and therefore this model predicts a pitch equal to the fundamental frequency. The same result is obtained if the autocorrelation of the compound response is taken. This is not consistent with the experimental results. For the MPP sequence, the pulse rate within the burst of 1776 pulses per second (inter-pulse interval of 563 µs) is too high for any single neuron to fire on every pulse (Figure 3b), and therefore the first-order interval statistics will depend on refractory effects. The distribution of first-order intervals will be broader than the SPP case, and contain shorter intervals. For example, a nerve may fire on only the first pulse in two successive bursts, giving an inter-spike interval equal to the fundamental period. Alternatively, a nerve may fire on the last pulse in one burst and the first pulse in the next burst, giving an inter-spike interval equal to the inter-burst gap. Thus the first-order interval model is consistent with the results in Figure 2**,** where all subjects matched the pitch of the MPP sequence to that of a SPP sequence with an interval between the fundamental period and the inter-burst gap. The across-subject variation may be due partly to differing neural refractoriness; e.g., subject S05 matched to the fundamental period, perhaps implying that his nerve fibres only fired once for each burst.

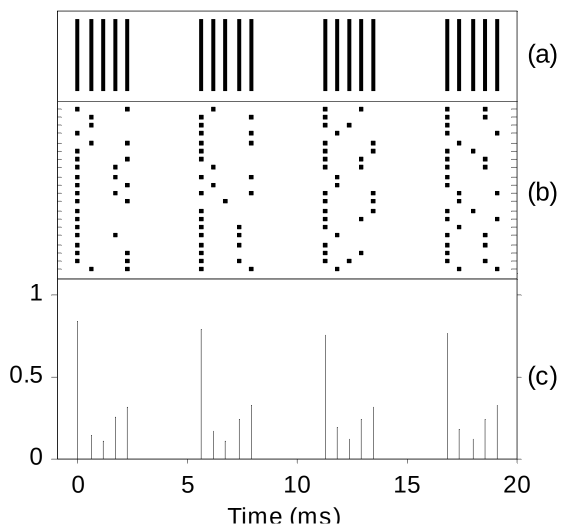

Figure 3. Simulated response to 178 Hz MPP sequence. (a) Stimulation pulses. (b) Responses of 20 individual fibres, with the spikes for each fibre shown as a horizontal row of dots. (c) Spike probability, calculated from the response of 1000 fibres.

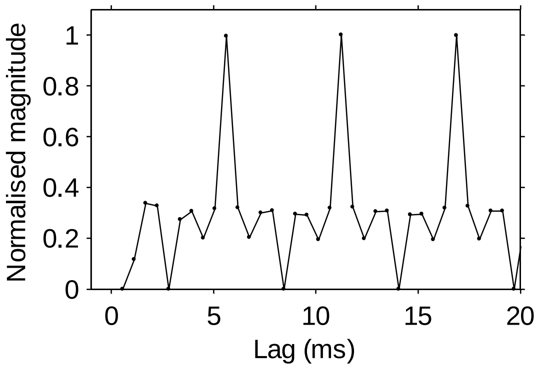

Figure 4. Autocorrelation of individual neural firing times, summed across a population of 1000 fibres, for the 178 Hz MPP sequence.

PLACE PITCH

Early psychophysical studies with multiple-channel cochlear implants showed that the percepts produced by stimulating individual electrodes could be ranked in a generally tonotopic order, corresponding to the location of the intracochlear electrodes. Apical electrodes have lower place-pitch than basal electrodes. Having 22 electrodes could imply that only 22 distinct place-pitch sensations can be produced; however, intermediate place-pitch percepts can be created by sequential stimulation of adjacent channels [10],[11].

A quantitative model of cochlear implant place-pitch perception [12] proposes that the place pitch of a stimulus on a group of electrodes is determined by the centroid *c* (or "centre of gravity") of the stimulation pattern, calculated as:

$$c = \frac{\sum_k k\, a(k)}{\sum_k a(k)} \qquad (2)$$

where *k* is the electrode number, and *a(k)* is the amplitude of stimulation on that electrode. The ability to discriminate between two stimuli on the basis of place pitch should depend on the distance between the two centroids; i.e., the sensitivity index *d′* for a pitch-ranking task using place-pitch cues alone can be modelled by:

$$d' = m c' \qquad (3)$$

where *c′* denotes the difference between the centroids of the two stimuli, and *m* is a constant for each subject characterising their usage of the place-pitch cue. Thus the proportion-correct score *p* should be related to the centroid difference as:

$$p = \Phi\left(\frac{m\,c'}{\sqrt{2}}\right) \qquad (4)$$

where Φ is the cumulative normal distribution.

Method

Six Nucleus 24 cochlear implant recipients participated in this experiment. A set of pure tones was constructed, spanning a two-octave range from 99 to 397 Hz. Each tone was 500 ms in duration, and was turned on and off with 50 ms raised-cosine ramps to minimise transients in the filter outputs. In each trial, a subject heard a pair of tones that had a frequency interval of two, four, or six semitones, and was asked whether the pitch was rising or falling.

All subjects used the ACE sound-coding strategy with a pulse rate of 1800 pps. A pure tone can produce stimulation on multiple electrodes, due to the broad, overlapping frequency responses of the filters (Figure 5). In a trial where the two tones have a large frequency interval, a different set of electrodes will be activated, providing a strong place-pitch cue. For small intervals, the ratios of the currents on adjacent electrodes may provide a finer place-pitch cue. The stimulation had no temporal modulation because quadrature envelope detection was used [13].

In calculating Equation 2, Laneau et al. [12] used the filter envelope amplitudes, but this fails to take into account the mapping from amplitude to stimulus current. In the sound processor, non-linear compression is applied, and any amplitude values that are below a base-level are discarded. Figure 5 (top panel) shows the compressed amplitude response of the first four filters of the ACE strategy. The bottom panel shows the centroid of the compressed amplitudes. Each step corresponds to the activation of an additional electrode. According to this model, the change in pitch when an additional electrode is activated is large compared to the change in pitch as the current varies within a group of electrodes.

In analysing the experimental data, the current levels on each electrode were determined, and a perceptual model of loudness [14] was applied to calculate the loudness profile and hence the loudness centroid of the stimulation pattern produced by each pure tone. The loudness centroid difference for each pair of tones was then calculated.

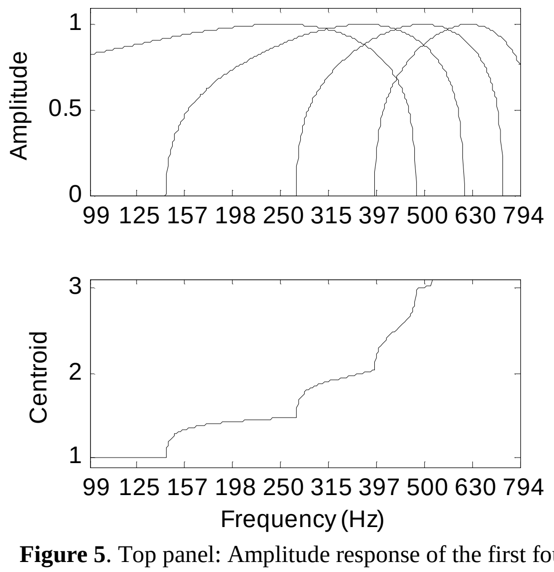

Figure 5. Top panel: Amplitude response of the first four filters of the ACE processing strategy, after non-linear compression. Bottom panel: centroid of the resulting stimulation pattern for a pure tone as a function of tone frequency. The centroid varies from the first to the third electrode.

Results

Two subjects scored at chance levels even with six-semitone intervals, so their data were excluded from further analysis. Results for the remaining four subjects are shown in Figure 6. Each percent correct score for a pair of tones is plotted against the calculated loudness centroid difference for that pair of tones. The scatter plot demonstrates that the data are approximated very well by a psychometric function according to Equation 4. Figure 7 shows the scores for subject S02 as a function of tone frequency. The most striking feature is the non-monotonic dependence of the scores on frequency, which is well-predicted by the model. Higher scores were obtained when the two tones stimulated a different set of electrodes; i.e., straddled a step in the centroid profile (Figure 5).

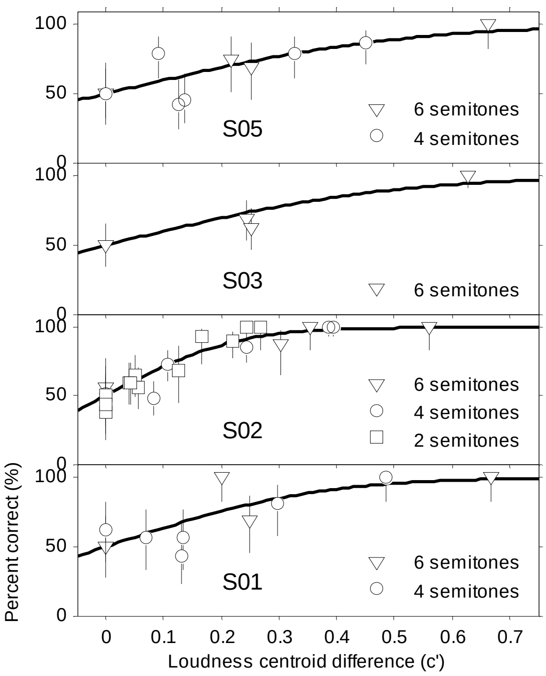

Figure 6. Scatter plot showing the pure-tone pitch ranking scores of four cochlear implant recipients against the loudness centroid difference. The black curves show the psychometric function that best fits each subject's data.

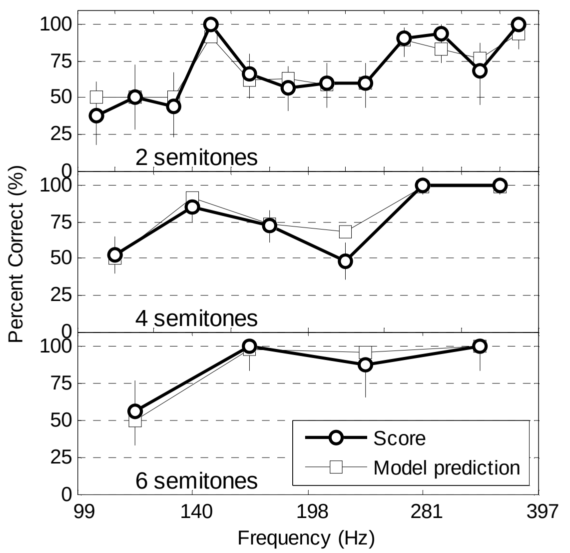

Figure 7. Pitch ranking pure tones with ACE for subject S02: percent correct score (thick line) and loudness centroid model prediction (thin line). Each panel shows the results for one frequency interval (2, 4, and 6 semitones). Each percent correct score is positioned on the abscissa midway between the frequencies of the two tones that were ranked.

Discussion

Timbre is the perceptual quality that distinguishes between two tones that have the same pitch, loudness and duration [15],[1]. The "brightness" of a harmonic tone depends on the centroid of the spectral profile [16],[15]. This is very similar to the centroid model of cochlear implant place pitch. McDermott [17] and Moore and Carlyon [18] speculated that cochlear implant place-pitch could be more akin to brightness than to pitch. As brightness can also be ordered on a low-to-high scale, it would allow high scores in a pitch-ranking procedure.

An operational definition of pitch is that variations in pitch can convey a melody. To differentiate between brightness and pitch, a test that uses melody perception is required.

MELODY PERCEPTION

Method

The Modified Melodies test measures pitch perception in a melodic context [13]. In each trial, the opening phrase of a familiar melody (Old MacDonald or Twinkle Twinkle Little Star) was presented twice. In one of the presentations, randomly selected, the pitch was deliberately modified. The rhythm was unchanged. The subject was asked to select the un-modified melody (a two-alternative forced-choice task). No feedback was given.

The Modified Melodies test supports several different types of pitch modification. In the Nudge modification, one note of the melody was shifted away from its correct pitch by a specified number of semitones. Blocks of trials were performed with shifts in the range 0.5 to 7 semitones. Scores were plotted as a function of shift, then a psychometric function was fitted, and the threshold (in semitones) was defined as the shift that would produce a score of 75% correct.

To investigate melody perception with only place cues to pitch, the melodies were played using pure tones [19]. Seven cochlear implant recipients were tested in the "Place C5" condition, with pure tones in the octave starting at 523 Hz. Three of these recipients were also tested in the "Place C3" condition, with pure tones in the octave starting at 131 Hz. To investigate melody perception with only temporal cues to pitch, the same three recipients were tested in the "Rate C3" condition, with melodies presented by stimulating on a single apical electrode with a varying pulse rate. The pulse rate was equal to the fundamental frequency of each note, in the octave starting at 131 Hz [20].

For comparison, six normal-hearing subjects were also tested. In the "Resolved" condition, the notes were in the octave starting at C4 (262 Hz), and each complex tone contained only low-numbered harmonics which would be resolved in the normal auditory system. In the "Unresolved" condition, the notes were in the octave starting at C3 (131 Hz), and each tone comprised only harmonics 20 to 24, which would not be resolved. The unresolved condition provides only temporal cues to pitch.

Results

For each normal-hearing subject, scores with unresolved harmonics were significantly worse than scores with resolved harmonics. The thresholds are shown in Figure 8. The subjects reported anecdotally that the melodies could be perceived with unresolved harmonics, but they had a weak, harsh, or "buzzy" pitch quality.

The thresholds for the cochlear implant recipients are shown in Figure 9. Subjects S02, S08, and S11 scored at chance level for the Place C5 condition, so the corresponding threshold is arbitrarily plotted at 7 semitones. No cochlear implant recipient obtained a threshold as low as that obtained by normal hearing subjects with resolved harmonics. However, subjects S01, S06, and S10 obtained some thresholds that were in the same range as normal hearing subjects with unresolved harmonics.

The results for the Rate C3 condition support the hypothesis that temporal cues in isolation can provide melodic pitch. This was also found by Pijl & Schwarz [21], who reported that three cochlear implant recipients were able to recognise melodies and judge musical intervals when the notes were presented by varying the pulse rate on a single electrode.

The results for the Place conditions support the hypothesis that place cues in isolation can provide melodic pitch. The possibility that subjects were recognising patterns of brightness changes cannot be completely ruled out. However, anecdotal reports from subjects suggested that they were hearing a melody based on pitch changes.

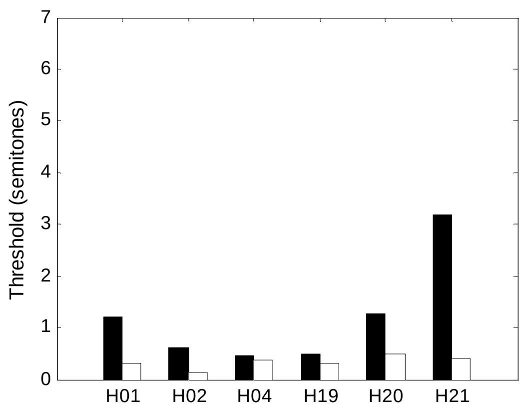

Figure 8. Modified Melodies Nudge thresholds for five normal-hearing subjects. Stimulus conditions: Unresolved harmonics (Black); Resolved harmonics (White).

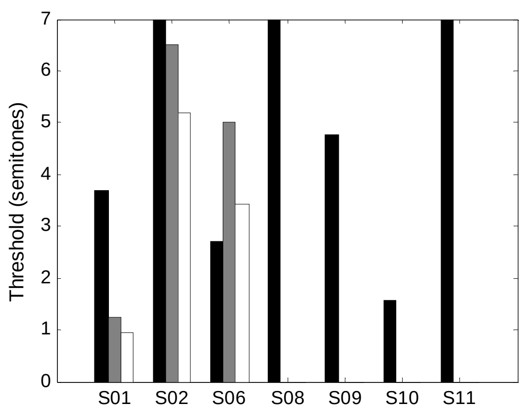

Figure 9. Modified Melodies Nudge thresholds for seven cochlear implant recipients. Stimulus conditions: Place C5 (Black); Place C3 (Grey); Rate C3 (White).

DISCUSSION

Earlier research has shown that cochlear implant place and temporal cues form independent perceptual dimensions [22],[23]. It is clear that there are two independent attributes at the neural level: the electrode determines which nerves fire, and the pulse timing determines the firing times. It would be surprising if these two very different neural attributes produced the same sensation; instead the expectation is that place and rate should be qualitatively different. Yet recipients are willing to label both of these perceptual attributes as "pitch", and the experimental results suggest that each by itself can evoke a melody. No present model of pitch perception is able to explain this paradox. It has been hypothesised that a strong pitch sensation requires a specific phase relationship between the nerve firing times across a local region of the cochlea [24]. This distinctive spatio-temporal excitation pattern is produced by a resolved harmonic. However, stimuli that do not produce the ideal spatio-temporal pattern can still evoke a weak pitch sensation that allows above-chance performance on pitch and melody tasks.

REFERENCES

[1] B. Moore, *An introduction to the psychology of hearing*, London: Academic Press, 1997.

[2] P. Cariani and B. Delgutte, "Neural correlates of the pitch of complex tones. I. Pitch and pitch salience," *Journal of Neurophysiology*, vol. 76, 1996, pp. 1698-1716.

[3] R. Meddis and L. O'Mard, "A unitary model of pitch perception," *The Journal of the Acoustical Society of America*, vol. 102, 1997, pp. 1811-1820.

[4] C. Kaernbach and L. Demany, "Psychophysical evidence against the autocorrelation theory of auditory temporal processing," *The Journal of the Acoustical Society of America*, vol. 104, 1998, pp. 2298-2306.

[5] R. Carlyon, A. van Wieringen, C. Long, J. Deeks, and J. Wouters, "Temporal pitch mechanisms in acoustic and electric hearing," *The Journal of the Acoustical Society of America*, vol. 112, Aug. 2002, pp. 621-633.

[6] Y. Tong, P. Blamey, R. Dowell, and G. Clark, "Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant," *The Journal of the Acoustical Society of America*, vol. 74, Jul. 1983, pp. 73-80.

[7] P. Busby and G. Clark, "Pitch and loudness estimation for single and multiple pulse per period electric pulse rates by cochlear implant patients," *The Journal of the Acoustical Society of America*, vol. 101, Mar. 1997, pp. 1687-1695.

[8] I. Bruce, L. Irlicht, M. White, S. O'Leary, S. Dynes, E. Javel, and G. Clark, "A stochastic model of the electrically stimulated auditory nerve: pulse-train response," *IEEE Transactions on Biomedical Engineering*, vol. 46, Jun. 1999, pp. 630-637.

[9] P. Abbas, C. Brown, J. Shallop, J. Firszt, M. Hughes, S. Hong, and S. Staller, "Summary of results using the Nucleus CI24M implant to record the electrically evoked compound action potential," *Ear & Hearing*, vol. 20, Feb. 1999, pp. 45-59.

[10] H. McDermott and C. McKay, "Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea," *The Journal of the Acoustical Society of America*, vol. 96, 1994, pp. 155-162.

[11] B. Kwon and C. van den Honert, "Dual-electrode pitch discrimination with sequential interleaved stimulation by cochlear implant users," *The Journal of the Acoustical Society of America*, vol. 120, 2006, pp.

EL1-EL6.

[12] J. Laneau, J. Wouters, and M. Moonen, “Relative contributions of temporal and place pitch cues to fundamental frequency discrimination in cochlear implantees,” *The Journal of the Acoustical Society of America*, vol. 116, 2004, pp. 3606-3619.

[13] B. Swanson, “Pitch perception with cochlear implants,” PhD thesis, The University of Melbourne, 2008.

[14] C. McKay, K. Henshall, R. Farrell, and H. McDermott, “A practical method of predicting the loudness of complex electrical stimuli,” *The Journal of the Acoustical Society of America*, vol. 113, Apr. 2003, pp. 2054-2063.

[15] R. Plomp, *Aspects of tone sensation*, London: Academic Press, 1976.

[16] W. Lichte, “Attributes of complex tones,” *Journal of Experimental Psychology*, vol. 28, Jun. 1941, pp. 455-480.

[17] H. McDermott, “Music perception with cochlear implants: a review,” *Trends in Amplification*, vol. 8, 2004, pp. 49-82.

[18] B. Moore and R. Carlyon, “Perception of pitch by people with cochlear hearing loss and by cochlear implant users,” *Pitch: neural coding and perception*, C. Plack, A. Oxenham, R. Fay, and A. Popper, Eds., New York: Springer, 2005, pp. 234 - 277.

[19] B. Swanson, P. Dawson, and H. McDermott, “Investigating cochlear implant place-pitch perception with the Modified Melodies test,” *Cochlear Implants International*, vol. 10, 2009, pp. 100-104.

[20] B. Swanson, “Inadequacy of Temporal Fine Structure for Pitch Perception,” *7th Asia Pacific Symposium on Cochlear Implants and Related Sciences*, Singapore: Medimond, 2009, pp. 183-188.

[21] S. Pijl and D. Schwarz, “Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes,” *The Journal of the Acoustical Society of America*, vol. 98, 1995, pp. 886-895.

[22] Y. Tong, G. Clark, P. Blamey, P. Busby, and R. Dowell, “Psychophysical studies for two multiple-channel cochlear implant patients,” *The Journal of the Acoustical Society of America*, vol. 71, Jan. 1982, pp. 153-60.

[23] C. McKay, H. McDermott, and R. Carlyon, “Place and temporal cues in pitch perception: are they truly independent?,” *Acoustics Research Letters Online*, vol. 1, Jul. 2000, pp. 25-30.

[24] G. Loeb, M. White, and M. Merzenich, “Spatial cross-correlation,” *Biological Cybernetics*, vol. 47, Jul. 1983, pp. 149-163.