

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Cochlear Implant Stimulation Rates and Speech Perception

Komal Arora, Richard Dowell and Pam Dawson

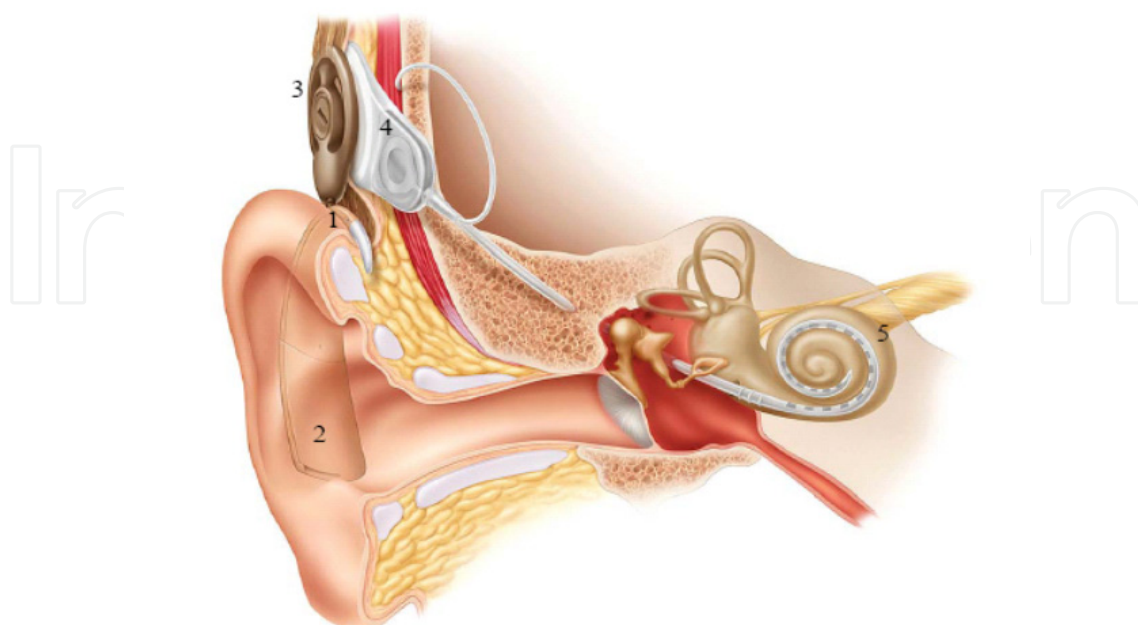
Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/49992>

## 1. Introduction

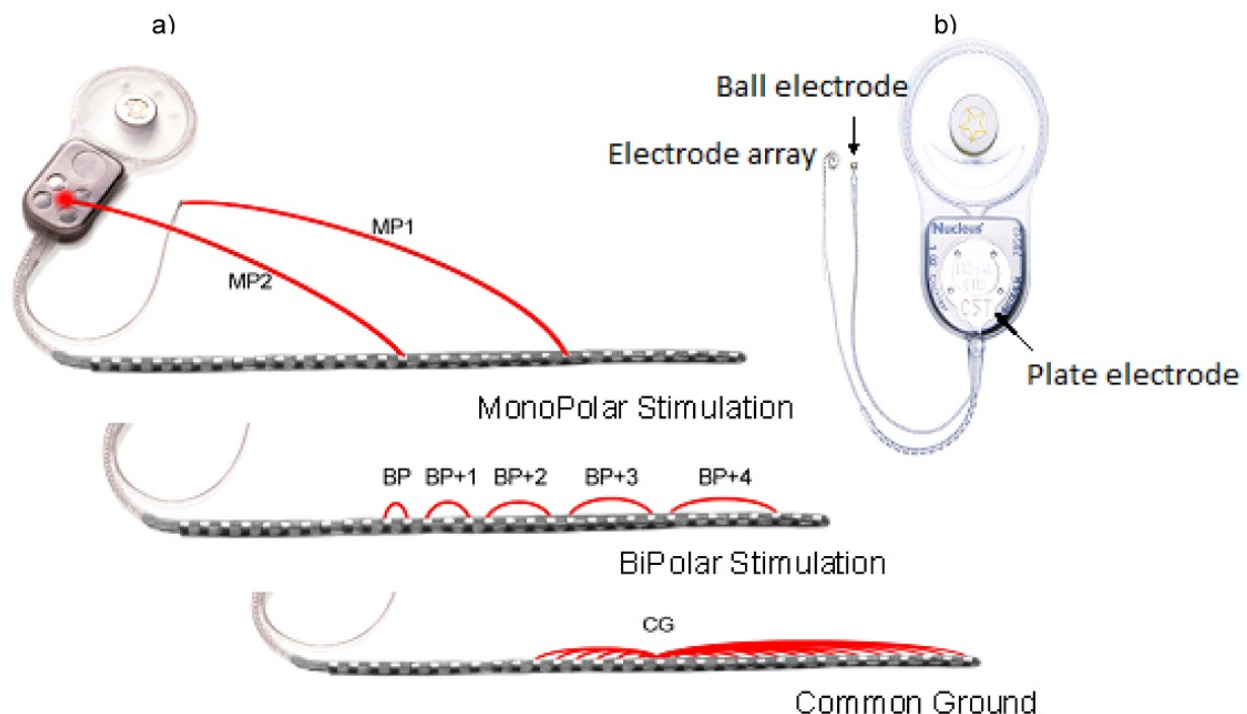
### 1.1. Cochlear implant system

For individuals with severe to profound hearing losses, due to disease or damage to the inner ear, acoustic stimulation (via hearing aids) may not provide sufficient information for adequate speech perception. In such cases direct electrical stimulation of the auditory nerve by surgically implanted electrodes has been beneficial in restoring useful hearing. This chapter will provide a general overview regarding sound perception through electrical stimulation using multi channel cochlear implants.



**Figure 1.** Cochlear implant system (Cowan, 2007)

The multiple channel cochlear implants consist of 1) a microphone which picks up sounds from the environment, 2) a sound processor which converts the analog sounds into a digitally coded signal, 3) a transmitting coil which transmits this information to the 4) receiver stimulator which decodes the radio frequency signals transmitted from the sound processor into the electrical stimuli responsible for auditory nerve stimulation via 5) an electrode array which provides multiple sites of stimulation within the cochlea (figure 1).



**Figure 2.** a) Stimulation modes. b) Extracochlear electrodes (Seligman, 2007).

Most of the current cochlear implant systems use *intracochlear* and *extracochlear* electrodes. Three different modes of current stimulation have been used in cochlear implant systems – Monopolar, bipolar and common ground (figure 2a). In monopolar stimulation, current is passed between one active intracochlear electrode and the extracochlear electrodes (which provide the return current path) placed either as a ball electrode under the temporalis muscle (MP1) or a plate electrode on the receiver casing (MP2) (figure 2b). When both of these extracochlear electrodes act as return electrodes in parallel, it is called MP1+2 configuration. In bipolar stimulation, current flows between an active and a return electrode within the cochlea; whereas in common ground stimulation, current flows from one electrode within the cochlea to all other intracochlear electrodes.

## 1.2. Speech processing in cochlear implants

The basic function of speech processing is to extract the speech information present in the acoustic signals and process it to produce a decoded version of the electrical stimulation signals that are transmitted as radio frequency (RF) signals to the receiver stimulator. The receiver stimulator converts this information into electrical stimulation patterns. A number

of speech processing strategies are available in current CI systems. This presents a choice of fitting parameters when setting up a system for individual patients. These parameters include the rate of stimulation, number of channels to be activated, mode of stimulation, electrical pulse width etc. These parameters along with the amplitude mapping of all available electrodes define a “map” for an individual cochlear implantee. Amplitude mapping involves measuring for each active electrode, the user’s threshold (T) level that is the level at which he/ she can just hear the stimulus and the maximum comfortable (C) level, that is, the level which produces a loud but comfortable sensation. These maps are loaded in the client’s sound processor. An individual cochlear implantee’s speech perception outcomes may differ based on the type of strategy he/she is using (Pasanisi et al., 2002; Psarros et al., 2002; Skinner et al., 2002a, b; Plant et al., 2002)

The current Nucleus® cochlear implants employ filter bank strategies which analyse the incoming signal using a bank or band-pass filters. The earliest filter bank strategy implemented in Nucleus cochlear implants called Spectral Maxima Sound processor (SMSP) (McDermott et al., 1992) did not extract features from the speech waveform. Instead, the incoming speech signal was sent to 16 band pass filters with centre frequencies from 250 Hz to 5400 Hz. The output from the six channels with the greatest amplitude (maxima) was compressed to fit the patient’s electrical dynamic range. The resultant output was then sent to the six selected electrodes at a rate of 250 pulses per channel (pps/ch). It is beyond the scope of this chapter to provide detailed information on all available speech coding strategies. A more comprehensive review of speech processing strategies is provided in Loizou (1998). A brief review of the most commonly currently used speech processing strategy in Nucleus devices is provided in section 1.2.3. Figure 3 shows the overall signal flow in the most commonly used filterbank strategies: Continuous Interleaved Sampling (CIS), Spectral Peak (SPEAK) and Advanced combination Encoders (ACE™).



**Figure 3.** Audio signal path used in current filterbank strategies (Swanson, 2007).

### 1.2.1. Front end

The input signal is at first picked up by the microphone and undergoes front end processing. The front end includes a *preamplifier*, *pre-emphasis*, *automatic gain control* and *sensitivity control*. The preamplifier amplifies the weak signals so that they can be easily handled by the rest of the signal processing. After that a pre-emphasis of 6dB/octave in frequency is applied to the signal. This again increases the gain for high frequency speech sounds. The automatic gain control helps reduce distortion or overloading. In Nucleus systems, the AGC has infinite compression, that is, the signal is at first amplified linearly to a certain level and compressed if it reaches beyond that fixed point. This type of

compression controls only the amplitude without distorting the temporal pattern of the speech waveform. The last part of the front end processing is automatic sensitivity control (ASC). Sensitivity refers to the effective gain of the sound processor and affects the minimum acoustic signal strength required to produce stimulation. At a higher sensitivity setting less signal strength is needed. On the other hand, at very low sensitivity settings, higher sound pressure levels are needed to stimulate threshold or comfortable levels. The sensitivity setting determines when the AGC will start acting and is aligned to C-level stimulation. This is automated in the current Nucleus cochlear implant devices and is called as ASC. The front end processing discussed so far is similar in SPEAK, CIS or ACE.

### *1.2.2. Filterbank*

After the front end processing, the signal passes through a series of partially overlapping band pass filters (each passing a different frequency range) where the signal is analyzed in terms of frequency and amplitude. The filterbank splits the audio signal into number of frequency bands simulating the auditory filter mechanism in normal hearing. The filter bands in the ACE strategy in current Nucleus processors are spaced linearly from 188 to 1312 Hz and thereafter logarithmically up to 7938 Hz. Each filter band is allocated to one intracochlear electrode in the implant system according to the tonotopic relationship between frequency and place in the cochlea.

Different speech coding strategies differ in the number of filter bands they use. For example the Nucleus implementation of CIS strategy has a small number (4 to 12) of wide bands, and SPEAK and ACE strategies have a large number (20 to 22) of narrow bands. The maximum number of bands is determined by the number of electrodes that are available in the particular implant system.

### *1.2.3. Sampling and selection*

To produce a satisfactory digital representation of a signal, the acoustic signal needs to be sampled sufficiently rapidly. According to the Nyquist theorem, the sampling rate should be twice the frequency to be represented. Aliasing errors can occur at lower rates. In the SPrint™ and Freedom™ processors, the filterbank uses a mathematical algorithm known as Fast Fourier Transform (FFT). This technique splits the spectrum into discrete lines or “bins” spaced at regular intervals in frequency. The sound is sampled at a regular rate in blocks, typically of 128 samples. The sampling rate for the Nucleus processors (SPrint™ and Freedom™ BTE) that use FFT is 16 kHz.

On the other hand, the ESPrit™ processors use analog switched capacitor filters. For smaller behind the ear units such as ESPrit™ and ESPrit™ 3G processors (used in the experiments discussed in this chapter, sections 3 and 4) switched capacitor filters were used because they were most power efficient at that time. These switched capacitor filters sample low frequencies at a rate of 19 kHz and high frequencies at a rate of 78 kHz. This filtered signal is further rectified to extract the envelopes. In the ESPrit™ series of processors filter outputs are measured using peak detectors. These peak detectors in ESPrit™ processors are matched

to the time response of the filters as each filter in the filterbank has different time response. The Med-El implants use Hilbert transform for measuring filter outputs. Hilbert transform gives amplitude response equal to bandpass filter response.

After the low pass filtering in SPrint™ / Freedom™ and peak detection in ESPrit™ processors, the outputs are further analyzed for the amplitude maxima. Each analysis window selects N maxima amplitudes (depending on the strategy) from filterbank outputs. The rate at which a set of N amplitude maxima are selected is referred to as the update rate. In Freedom™/ SPrint™ processors this rate is fixed at 760 Hz. However, in the ESPrit™ series of processors, it varies. For the high level sounds, the update rate is 4 kHz and for low level sounds, it is 1 kHz. This is also the rate at which new information is generated by the sound processor. In Med-El sound processors, using CIS strategy, new data is available from the sound processor at the stimulation rate.

The following sections describe how sampling and selection is done for each strategy.

#### *1.2.3.1. Spectral Peak (SPEAK) strategy*

The Spectral Peak (SPEAK) strategy is a derivative of the SMSP strategy (McDermott, et al., 1992) where the number of channels was increased from 16 to 20 with the center frequencies ranging from 250-10,000 Hz. The frequency boundaries of the filters could be varied. It also provided flexibility to choose the number of maxima from one to ten with an average of six. The selected electrodes were stimulated at a fixed rate that varied between 180 and 300 pps/ch. The stimulation rate was varied depending upon the number of maxima selected. In case of limited spectral content, lesser maxima were selected and the stimulation rate increased which provided more temporal information and hence may have compensated for reduced spectral cues. Similarly when more maxima were selected, the stimulation rate was reduced (Loizou, 1998). This strategy was implemented in the Spectra sound processor (Seligman and Mc Dermott, 1995) and was next incorporated in the Nucleus 24™ series of sound processors. However, the SPrint™ and Freedom™ systems (in Nucleus 24 series) used a fixed analysis rate of 250 Hz. Nucleus 24 also allowed for higher rates which are covered in the description of ACE strategy (section 1.2.3.3). A typical SPEAK strategy in the Nucleus devices consists of 250 pps/ch stimulation rate with the selection of six or eight maxima out of 20 channels.

#### *1.2.3.2. Continuous Interleaved Sampling (CIS) strategy*

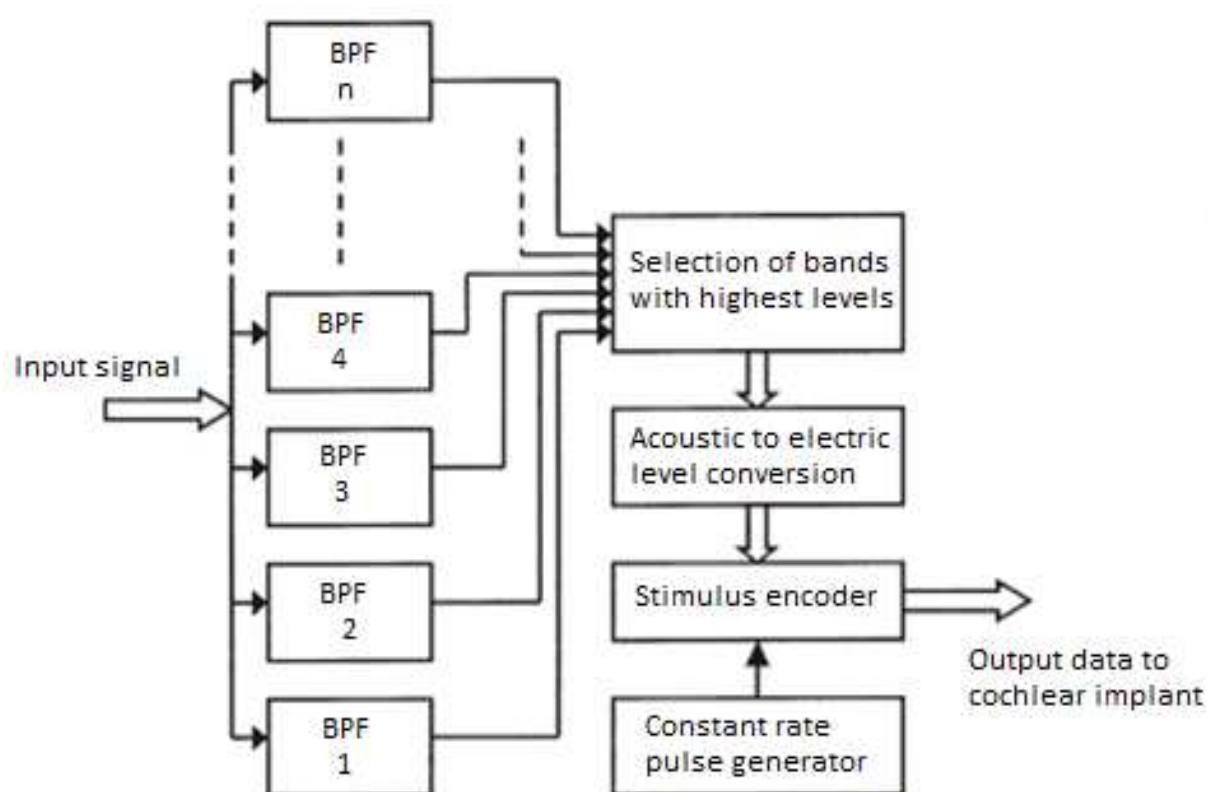
The CIS strategy (Wilson et al., 1991) was developed for the Ineraid implant. In this strategy, the filterbank has six frequency bands. The envelope of the waveform is estimated at the output of each filter. These envelope signals are sampled at a fixed rate. The envelope outputs are finally compressed to fit the dynamic range of electric hearing and then used to modulate biphasic pulses. The filter output in terms of electrical pulses is delivered to six fixed intracochlear electrodes. The key feature of the CIS strategy is the use of higher stimulation rates to provide better representation of temporal information. The variations in the pulse amplitude can track the rapid changes in the speech signal. This is possible due to the shorter pulses with minimal delays/inter-pulse interval (Wilson et al., 1993). The



possible benefits of using high stimulation rates are discussed further in the Section 2. The stimulation rate used in the CIS strategy is generally 800 pps/ch or higher. A modified version (CIS+) of the CIS strategy is used these days, more typically in Med-El implants. CIS+ uses a Hilbert Transformation (Stark and Tuteur, 1979) to represent the amplitude envelope of the filter outputs. This transformation tracks the acoustic signal more closely for a more accurate representation of its temporal dynamics compared to the other techniques used to represent the amplitude envelope in other implant systems like “wave rectification”, “low-pass filtering” or “fast Fourier transform” (Med-El, 2008).

### 1.2.3.3. Advanced combination encoder (ACE™) strategy

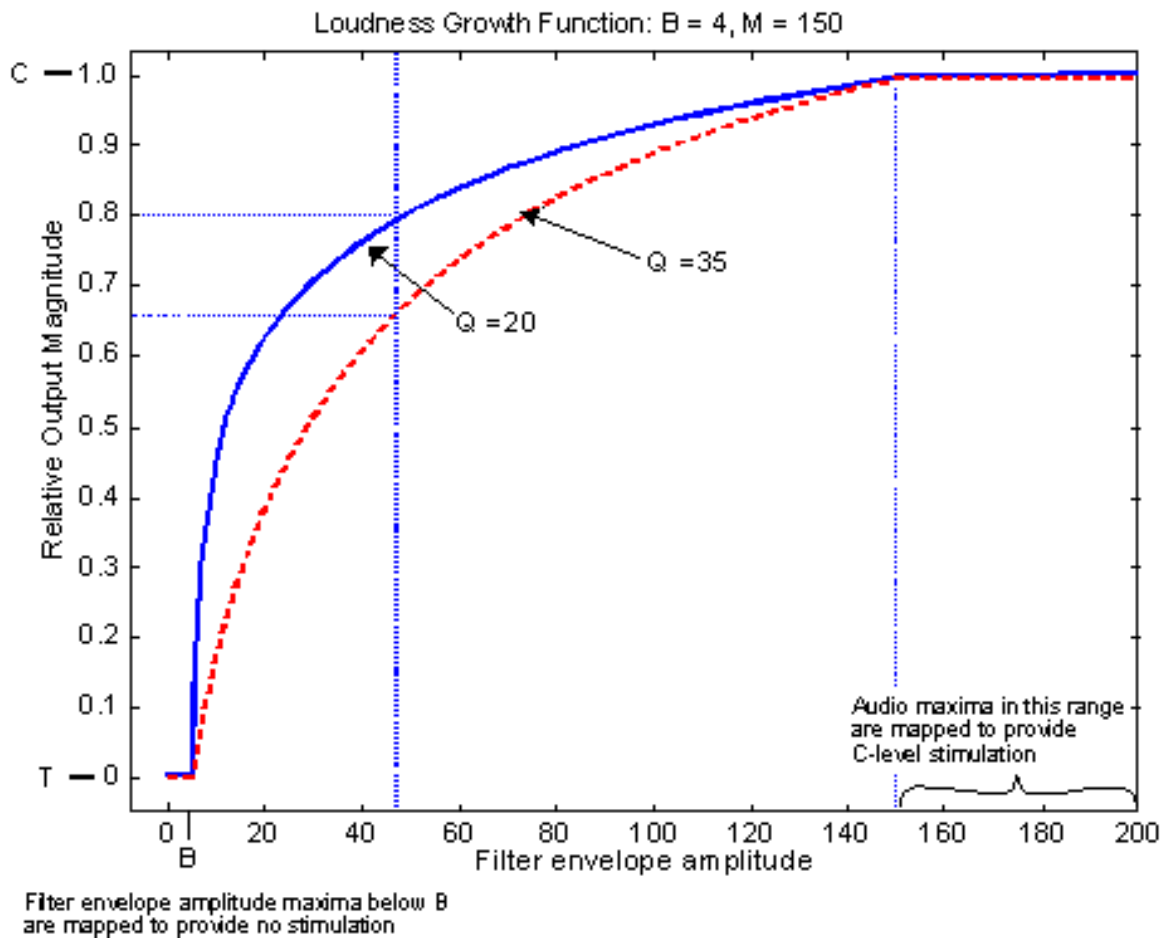
The ACE™ strategy (Vandali et al., 2000) is similar to the SPEAK strategy except that it combines the higher stimulation rate feature of the CIS strategy. Selection of electrodes is similar to the SPEAK strategy; however it can also be programmed to stimulate fixed electrode sites as in CIS. Thus this strategy attempts to provide the combined benefits of good spectral and temporal information for speech. Figure 4 depicts the schematic diagram of the ACE strategy. The filter bands in the ACE strategy are spaced linearly from 188 to 1312 Hz and thereafter logarithmically up to 7938 Hz.



**Figure 4.** Schematic diagram of the ACE strategy (McDermott, 2004).

The output of the filters is low pass filtered with an envelope cut off frequency between 200 and 400 Hz (In SPrint™ and Freedom™ processors). The ESPr™ series of processors use peak detectors (discussed in section 1.2.3) rather than low pass filtering. After the envelopes

from each of the band pass filters are extracted, a subset of filters with the highest amplitudes is selected. These are called maxima which can number from 1 to 20. The range of stimulation rates in ACE can be selected from a range of 250 to 3500 pps/ch.



**Figure 5.** Loudness growth function in a Nucleus cochlear implant system. The x-axis shows the input level (dB SPL) and the y-axis shows the current level. The steeper loudness curve has smaller  $Q$  (Swanson, 2007).

### 1.3. Amplitude mapping

The next stage is amplitude mapping where energy in each filter band decides the amplitude of the pulses delivered to the corresponding electrodes. An amplitude conversion function is applied to convert acoustic levels into electrical levels. This function is described in terms of T-SPL and C-SPL. T-SPL refers to the sound pressure level (SPL) required to stimulate at threshold level (around 25 dB) and C-SPL refers to the sound pressure level required to stimulate at maximum comfortable level (65 dB). For speech at normal conversational levels, there are hardly any speech sounds below 25 dB SPL. A standard 40 dB input dynamic range (IDR) that is the difference between the T-SPL and C-SPL is mapped on to an electrical or output dynamic range which depends on the threshold and maximum comfortable electrical levels (typically about 8 dB). Thus, the output from the filters is compressed to fit the electrical dynamic range. The compression is described by the



loudness growth function. The parameter  $Q$  (steepness factor) controls the steepness of the loudness growth curve (figure 5). Nucleus ESPrit™ processors operate on the input DR of 30 dB. IDR can be increased up to 50 dB in current Nucleus sound processors.

#### 1.4. Radio frequency (RF) encoder

The data from the sound processor is transmitted to the receiver stimulator. The transmitted code is made up of a digital data stream and transmitted by pulsing the RF carrier. The receiver-stimulator decodes the current specific information (e.g. electrode selection, current level, mode of stimulation etc.) and converts this information into electrical stimulus pulses.

The biphasic electrical current pulses are then delivered non-simultaneously to the corresponding electrodes at a fixed rate. The non-simultaneous presentation (one electrode is stimulated at one time) of current pulses is used to avoid channel interaction (Wilson et al., 1998). The range of stimulation rates available in current Nucleus devices range from 250 pps/ch to 3500 pps/ch. In the current study the Nucleus ESPrit™ 3G processor was used which has a maximum stimulation rate of 2400 pps/ch. However, it is unlikely that these higher stimulation rates provide any additional temporal information unless the processor update rate is at least equivalent (760 Hz in SPrint™ and Freedom™ processors and 1000 - 4000 Hz in ESPrit™ processors). From the signal processing point of view, the ESPrit™ series of processors have the potential to add new temporal information with high rates because of their higher update rate (1-4 kHz).

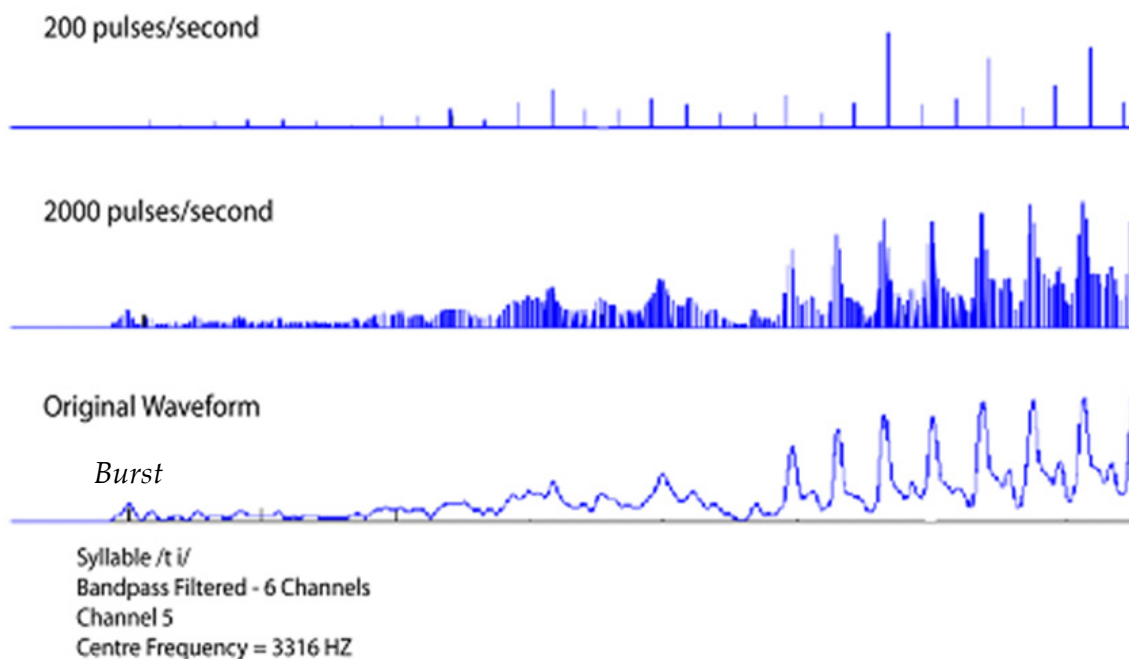
## 2. Electrical hearing

As discussed in the sections above, various speech coding strategies and choice of fitting parameters are available in current CI systems. Studies have demonstrated that different strategies and/or parameter choices can provide benefits to individual patients but there is no clear method for determining these for a particular individual. The current literature in this area shows a lack of consistency in outcomes, particularly, when the electrical stimulation rate is varied. There could be some underlying physiological or psychological correlates behind it. For example the outcomes may be related to the temporal processing abilities of CI users. This section will review some of the existing literature pertaining to the stimulation rate effects on the performance of the cochlear implant subjects.

### 2.1. Stimulation rate effects on speech perception

The stimulus signals delivered in existing CI systems are generally derived by sampling the temporal envelope of each channel at some constant (analysis) rate and using its intensity to control the stimulation current level delivered to the corresponding electrode site (again at a constant stimulation rate which is typically equal to the analysis rate). The range of stimulation rates employed in devices varies extensively amongst systems from low (<500 pps/ch), to moderate (500-1000 pps/ch), through to high (>1000 pps/ch). Figure 6 shows the effect of stimulation rate in processing the syllable /ti/. The bottom most waveform is the

original speech envelope of channel 5 for the syllable /ti/. As seen in the 200 pps/ch stimulation rate condition, pulses are spaced relatively far apart, so this sort of processing may not be able to extract all of the important temporal fine structure of the original waveform. When a higher pulse rate is used, the pulses are placed more closely and they can carry the temporal fine structure more precisely (Loizou et al., 2000). From a signal processing point of view this seems reasonable, however in practice, perceptual performance of CI users is often not improved when using higher rates.



**Figure 6.** The pulsatile waveforms for channel 5 of the syllable /ti/ with stimulation rates of 200 pps/ ch and 2000 pps/ch. The syllable /ti/ was band pass filtered into six channels and the output was rectified and sampled at the rates indicated in this figure. The bottom panel shows the speech envelope of channel 5 for syllable /ti/ (modified from Loizou et al., 2000).

When considering the appropriate rate to employ for coding of F0 temporal information, Nyquist's theorem states that the rate must be at least twice the highest frequency to be represented. However, according to McKay et al (1994), the stimulation rate for CI systems should be at least four times the highest frequency to be represented. This suggests that rates of >1200 pps per channel are needed to effectively code the voice pitch range up to 300 Hz. On the other hand studies examining neural responses to electrical stimulation in animals have shown that at rates above >800 pps/ch, there is poorer phase locking and less effective entrainment of neurons due to refractory effects being more dominant (Parkins, 1989; Dynes & Delgutte, 1992). It is therefore simplistic to assume that a higher stimulation rate alone will necessarily result in more effective transfer of temporal information in the auditory system.

A number of studies explored the effect of stimulation rate on speech perception in CI users. Results for some of the previous studies using the continuous interleaved sampling (CIS) speech coding strategy and the MED-El implant showed benefits for moderate and high stimulation rates (Loizou et al, 2000; Keifer et al, 2000; Verchuur, 2005; Nie et al, 2006). However, other studies using the CIS strategy did not show a benefit for high rates (Plant et al, 2002; Friesen et al, 2005). The comparison of these studies is complicated by the use of different implant systems. Studies using the Nucleus devices with 22 intracochlear electrodes and the ACE strategy did not show a conclusive benefit for higher rates (Vandali et al, 2000; Holden et al, 2002; Weber et al, 2007; Plant et al, 2007). Again, there are some limitations in these studies due to the specific hardware used. The higher stimulation rates tested by Vandali et al (2000) and Holden et al (2002) probably did not add any extra temporal information due to the limited analysis rate of 760 Hz employed in the SPrint™ processor used in those studies. Many of these studies reported large individual variability among subjects. Although the recent study by Plant et al (2007) found no significant group mean differences between higher rate and lower rate programs, five of the 15 subjects obtained significantly better scores with higher rates (2400 pps/ch & 10 maxima, or 3500 pps/ch & 9 maxima) compared to lower rates (1200 pps/ch & 10 maxima, or 1200 pps/ch & 12 maxima) for speech tests conducted in quiet or noise. Only two subjects obtained significant benefits in both tests using the higher set of rates, and the results were not conclusive because significant learning effects were observed in the study. Likewise, in the study by Weber et al (2007), group speech perception scores in quiet and noise did not demonstrate a significant difference between stimulation rates of 500, 1200, and 3500 pps/ch using the ACE strategy. However, some variability in individual scores was observed for six of the 12 subjects for the sentences in noise test.

Reports on subjects' preferences for particular stimulation rates with Nucleus devices have shown results in favor of low to moderate stimulation rates. In the study done by Vandali et al (2000), 250 and 807 pps/ch rates were preferred over 1615 pps/ch. Similarly, Balkany et al (2007) reported preferences for slower set of rates (500 to 1200 pps/ch, ACE strategy) for 37 of the 55 subjects, compared to faster set of rates (1800 to 3500 pps/ ch, ACE RE strategy). Authors also reported that the rate preference by individual subjects tended towards the slower rates within each of the two sets of stimulation rates. Similarly, in a clinical trial conducted in North America and Europe by Cochlear Ltd (2007) on subject selection of stimulation rate with the Nucleus Freedom system, there was a preference for stimulation rates of 1200 pps/ch or lower. Speech perception test results also showed improved performance with stimulation rates of 1200 pps/ch or lower compared to a higher set of rates (1800, 2400, and 3500 pps/ch).

## 2.2. Perception of complex sounds in electric hearing

Both spectral and temporal information in acoustic signals encoded by the auditory system are important for speech perception in normal hearing. Spectral and temporal information are coded by the site (or place) and timing of neural activity along the basilar membrane respectively. For speech sounds, (broad) spectral information such as, the frequency and

bandwidth of vowel formants, is encoded via place along the tonotopic axis of the cochlea. Fine spectral structure is also encoded, such as the frequency of the fundamental (F0) and lower-order harmonics of the fundamental for voiced sounds (Plomp, 1967; Houtsma, 1990).

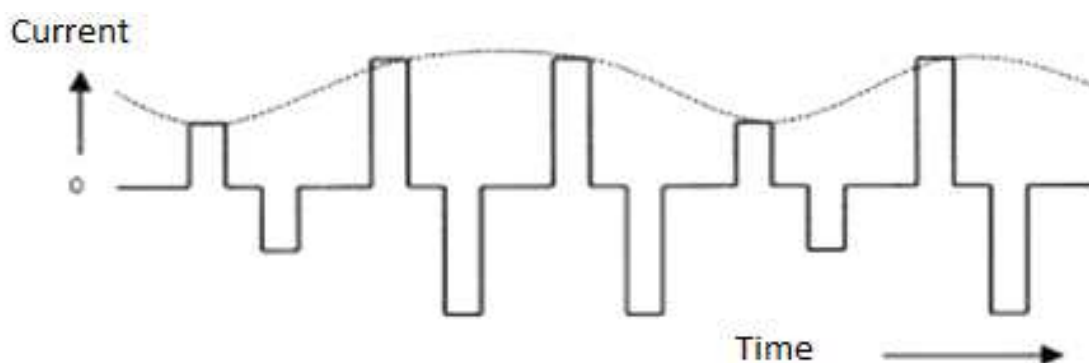
Temporal properties of speech encoded in the auditory system comprise low frequency envelope cues (<50 Hz) which provide information about phonetic features of speech, higher frequency envelope information (>50 Hz), such as F0 periodicity information in auditory filters in which vowel harmonics are not resolved, and most importantly fine temporal structure (Rosen, 1992). The perceived quality or timbre of complex sounds is mostly attributable to the spectral shape. For example, each vowel has specific formant frequencies and patterning of these formant frequencies helps in determining the vowel quality and vowel identity (Moore, 2003b).

The frequency coding in cochlear implants takes place in two ways: a) spectral information presented via the distribution of energy on multiple electrodes along the cochlea, b) temporal information which is mainly presented via the amplitude envelopes of the electrical stimulation pulses. These two ways of coding and the spectral shape coding are described in the following sub sections.

### *2.2.1. Spectral/ Place pitch coding*

In current multichannel cochlear implants, the pitch sensations produced at various electrodes vary depending on the position of electrodes at different sites in the cochlea (Simmons, 1966; Pialoux et al., 1976; Clark et al., 1978; Tong et al., 1979, 1982; Burian et al., 1979). An important finding of some of these studies was that the subjects' description of pitch being sharp or dull depended on whether the higher or lower frequency regions excited according to the tonotopic organization of the cochlea (Clark et al., 1978; Tong et al., 1979, 1982; Donaldson and Nelson, 2000; Busby and Clark, 2000). However, place coding is relatively crude due to the limited numbers of electrodes (up to 22) in current cochlear implant systems compared to approximately 15,000 receptor hair cells in the normal cochlea, degeneration of auditory nerve fibres innervating the cochlea, and the fact that electrode arrays do not access the full length of the cochlea (Ketten et al., 1998; Baskent and Shannon, 2004). That is, the most apical electrodes do not stimulate the most apical site in the cochlea. In addition, the tonotopicity seen in the CI is not perfect in all subjects. Some subjects cannot discriminate between the pitches across different electrodes and/or a more basal electrode may sound lower in pitch than a more apical one (so called pitch reversal) (Busby et al., 1994; Cohen et al., 1996; Nelson et al., 1995; Donaldson and Nelson, 2000). Spatial separation between electrodes has also been found to affect CI users' pitch perception ability. Increased separation improved pitch ranking performance of CI subjects (Nelson et al., 1995; Tong and Clark, 1985). In addition, most CI listeners cannot make full use of the spectral information available through their implants. Friesen et al. (2001) reported no significant improvement in speech perception (in quiet and noise) as the number of electrodes was increased beyond eight. The above mentioned place anomalies could affect the speech perception of cochlear implantees as spectral pitch is an important attribute in speech. Therefore, the use of

temporal information to assist speech perception for cochlear implant users may be quite important.



**Figure 7.** An amplitude modulated current pulse train (McDermott, 2004).

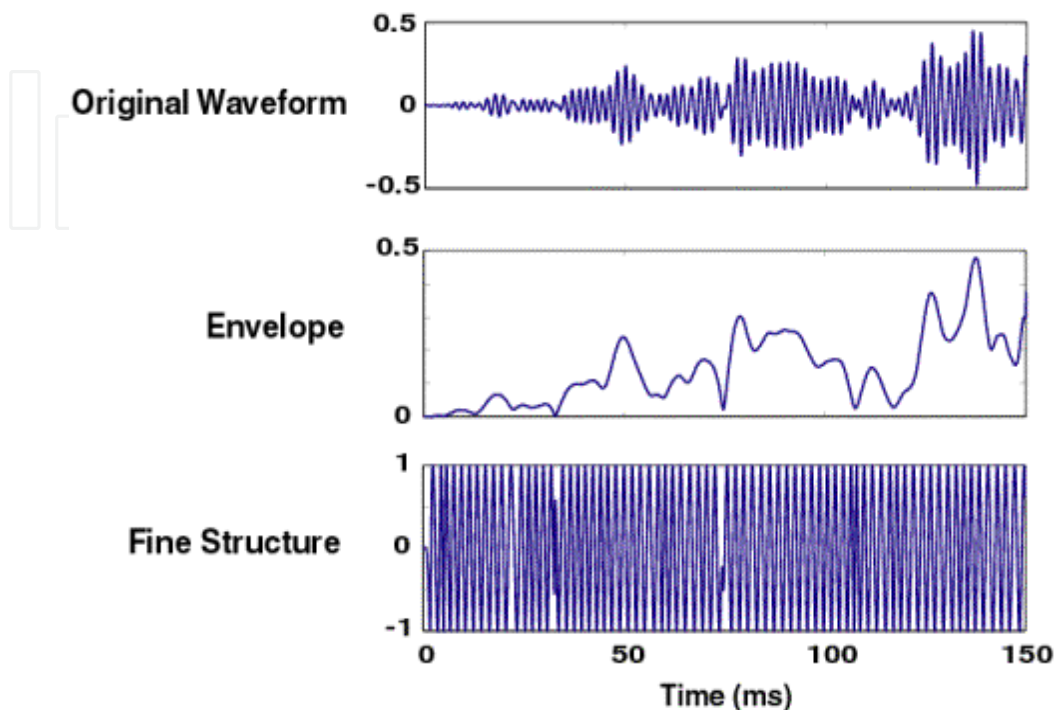
### 2.2.2. Temporal pitch coding

In current fixed rate speech processing strategies, the amplitude of the pulse trains on each electrode is modulated by the amplitude envelope of the acoustic signal (figure 7). The amplitude of pulse trains on each electrode depends on their corresponding amplitudes estimated from the filter bank. These amplitude variations over time are responsible for the subjective temporal pitch in cochlear implant users. In these filter bank strategies, the speech information is low pass filtered within each frequency band using a filter frequency between 200- 400 Hz with the elimination of most of the fine temporal structure. Thus, temporal information is mainly presented via the amplitude envelopes of the electrical stimulation pulses. The slower overall variations in amplitude over time modulating high-frequency carrier refer to the envelope of a waveform. The rapid variations in amplitude over time with average rate close to the carrier signal are described as the fine structure portion of a waveform. An example of a filtered speech waveform with its envelope and fine structure composition is depicted in figure 8.

As mentioned earlier, temporal information provides envelope, periodicity and fine structure information in speech (Rosen, 1992). Low frequency temporal envelope information is sufficient for segmental speech information (Fu and Shannon, 2000; Shannon et al., 1995; Xu et al., 2002). These studies show that 16-20 Hz temporal information is sufficient when adequate spectral cues are available and in case of poor spectral cues subjects may be forced to use envelope frequencies up to 50 Hz. On the other hand, higher frequency temporal cues may assist CI listeners to perceive the finer fundamental frequency (F0) temporal cues. Fu et al. (2004) found that for normal hearing subjects and cochlear implant recipients, voice gender recognition improved when the speech envelope low pass filter frequency was increased from 20- 320 Hz. It is possible that if a listener can attend to the voice pitch of the target speaker, he can separate the speech from competing noise thereby improving the overall speech perception (Brokx and Nöteboom, 1982; Assmann and Summerfield, 1990). A need for higher resolution spectral information (for instance, an increased number of channels) is likely to be needed for improving speech perception in



noise (Fu et al., 1998; Dorman et al., 1998). If this increased spectral resolution can not be obtained with current electrode technology, better coding of periodicity cues may provide another avenue for improving performance for CI users.



**Figure 8.** The envelope and fine structure components of a filtered speech signal (<http://research.meei.harvard.edu/chimera/motivation.html>).

Psychophysical studies of electrical stimulation in the human auditory system indicate that temporal pitch information up to 300 Hz is probably available to CI users (Eddington et al., 1978; Tong et al., 1979; Shannon, 1983; Moore & Carlyon, 2005). These studies used steady pulse trains (with varied rate of stimulation) delivered to single electrode sites. For very low pulse rates (<50 Hz), the signal is reported to be perceived as buzz-like sound and for rates above 300 Hz, a little change in perceived pitch is reported. This ability of CI users varies with a few being able to discriminate rate increases up to 1000 Hz (Fearn and Wolfe, 2000; Townshend et al., 1987).

### 2.2.3. Coding of spectral shape/timbre

Spectral shape is important in the perception of complex sounds, including dynamic characteristics; however, in cochlear implant users it is particularly important. Spectral shape discrimination can help in identifying vowels in the speech and many other types of sound even when other acoustic cues are absent (McDermott, 2004). In the normal auditory system spectral shape is represented by the relative response across filters. There are different ways to code the relative level across filters such as relative firing rates of neurons as a function of characteristic frequency (place coding); relative amount of phase locking between neurons and the different frequency components (temporal coding); and by the



level dependent phase changes on the basilar membrane. In implant systems, spectral shape, like frequency coding is coded by filtering the signal into several frequency bands and the relative magnitude of the electric signal across electrode channels. The coding of spectral shape cannot be as precise as that in the normal ear due to the relatively small number of effective channels in current cochlear implant systems. In addition, the detailed temporal information relating to formant frequencies is not conveyed due to the inability to effectively code temporal information above about 300 Hz. One approach to improving the representation of the temporal envelope and higher frequency periodicity cues is to increase the low pass filter frequency applied to the amplitude envelope and/or to use higher stimulation rates. However, results so far do not conclusively show benefit for higher rates (section 2.1). Furthermore, it is also not clear how effectively CI listeners can resolve these temporal modulations.

In normal hearing and hearing impaired subjects, temporal resolution can be characterized by the temporal modulation transfer function (TMTF) which relates the threshold for detecting changes in the amplitude of a sound to the rapidity of changes/modulation frequency (Bacon and Viemeister, 1985; Burns and Viemeister, 1976; Moore and Glasberg, 2001). In this task, modulation detection can be measured for a series of modulation frequencies. The stimuli used in these experiments are either amplitude modulated noise or complex tones. To study temporal modulation independent from spectral resolution, spectral cues are removed by using broadband noise as a carrier. This type of stimulus has waveform envelope variations but no spectral cues. Complex tones are the combination of two or more pure tones. This type of sinusoidal amplitude modulated (SAM) signal has components at  $fc - fm$ ,  $fc$ ,  $fc + fm$  where  $fc$  is the carrier frequency and  $fm$  is the modulation frequency (figure 9b).

The components above and below the centre frequency are called sidebands. If  $fc$  (e.g. 1200, 1400, 1600 Hz) is an integer multiple of  $fm$  (200 Hz), it forms a harmonic structure otherwise, it is called an inharmonic waveform. In signal theory and acoustic literature, amplitude modulated signals are described by the formula:

$$(1 + m \sin 2\pi \cdot fm \cdot t) \sin 2\pi fc t \quad (1)$$

where  $t$  is time and  $m$  is the modulation index ( $m = 1$  means 100% modulation). Figure 9(a) shows an example of an amplitude modulated signal. The pink color waveform shows the depth of modulation. In acoustic hearing,  $m = MD$  (modulation depth) which is defined as:

$$m = (peak - trough) / (peak + trough) \quad (2)$$

where peak and trough refer to the peak and trough (Sound Pressure Level, SPL) levels of the modulation envelope.

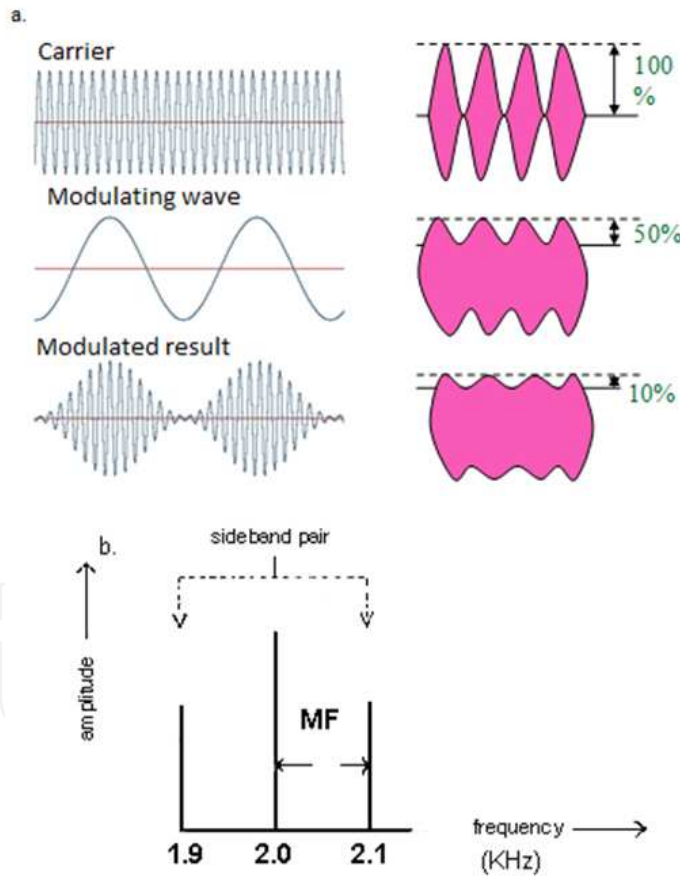
However, in cochlear implant stimulation, the loudness of signals is governed more-so by peaks in the stimulus envelope rather than average stimulus level. Thus, when presenting AM signals to CI users, the peak level of the AM signal is held fixed and only the trough of the modulation envelope is reduced. Eq. 3 describes the SAM formula used in study 2.

$$\sin(2\pi \times fc \times t) \times [1 + 0.5m \times (\sin(2\pi \times fm \times t) - 1)] \quad (3)$$

where  $fc$  is the carrier frequency,  $fm$  is the modulation frequency,  $t$  is time,  $m$  is the modulation index which controls the modulation depth (MD), where

$$m = 1 - 1 / MD \quad (4)$$

and where modulation depth (in cochlear implant stimulation) is defined by the peak over trough (peak/trough) level in the envelope. The psychophysical measure tested in study 2 was modulation detection threshold (MDT) which refers to the depth of modulation necessary to just allow discrimination between a modulated and unmodulated waveform. In this study stimuli were presented through the cochlear implant sound processor. Modulation depth (MD) was converted into modulation index ( $m$ ) using Eq. (3) for all analysis. This was because most of the studies which measured modulation detection in CI recipients have used modulation index ( $m$ ) for analysis purposes.



**Figure 9.** a) An example of an amplitude modulated signal. The pink color waveforms show the depth of modulation. b) The sinusoidally amplitude modulated signal with 2000 Hz carrier frequency ( $fc$ ), 100 Hz modulation frequency (MF).

The limitation with the sinusoidal carriers is that the modulation introduces sidebands which can be heard as separate signals. Also the results of modulation detection may be

influenced by the “off frequency listening”. That is if the carrier and the modulated frequency are separated quite apart the sounds can be heard from the auditory filters centered at the carrier frequency or the sideband frequency depending on the intensity of modulation (Moore and Glasberg, 2001). However, in study 2 a sinusoidal carrier was used instead of noise as the carrier because subjects with cochlear hearing loss have reduced frequency selectivity (Glasberg and Moore, 1986; Moore, 2003) leading to poor spectral resolution of the sidebands. Thus, TMTF in such cases is mainly influenced by temporal resolution over a wide range of modulation frequencies (Moore and Glasberg, 2001). It is also difficult for the CI users to spectrally resolve the components of complex tones (Shannon, 1983). In addition; the noise signal would have its own temporal envelope which can confound the results of modulation detection (Moore and Glasberg, 2001).

### 3. Study 1: Effect of stimulation rate on speech perception

This section provides the details of the study by Arora et al (2009), which examined the effect of low to moderate stimulation rates on speech perception in Nucleus CI users with addition of two more subjects.

#### 3.1. Rationale

If low to moderate stimulation rates do indeed provide equivalent or better speech perception, then recipients may also benefit from reductions in system power consumption and processor/device size and complexity. So far, low to moderate rates have not been explored well in Nucleus™ 24 implants with the ACE strategy, especially in the range of 250-900 pps/ch in spite of the fact that this range of rates is often used clinically<sup>1</sup> with Nucleus devices, (which worldwide are the most used devices so far among CI recipients). The authors thus chose to examine rates of 275, 350, 500, and 900 pps/ch in this study.

This study was specifically designed to determine:

- Whether rates of stimulation (between 275 and 900 pps/ch) have an effect on the speech perception in quiet and noise for the group of adult CI subjects.
- Whether optimal rate varies among various subjects.
- Whether there is a relation between the subjective preference measured with comparative questionnaire and the speech perception scores.

#### 3.2. Method

Ten postlingually deaf adult subjects using the Nucleus™ 24 Contour™ implant and ESPrit™ 3G sound processor participated in the study. Table 1 shows the demographic data for the subjects. Low to moderate stimulation rates of 275, 350, 500 and 900 pulses-per-second/channel (pps/ch) were evaluated.

---

<sup>1</sup> As per the information available from Melbourne Cochlear Implant Clinic (RVEEH, University of Melbourne) and Sydney Cochlear implant Center (The University of Sydney).

Test material comprised CNC open set monosyllabic words (Peterson, & Lehiste, 1962) presented in quiet and Speech Intelligibility test (SIT) open set sentences (Magner, 1972) presented in four talker babble noise. Four lists of CNC words were presented in each session at a level of 60 dB SPL RMS. An adaptive procedure (similar to the procedure used by Henshall and McKay, 2001) was used to measure speech reception threshold (SRT) for the sentence test in noise. Four such SRT estimates were recorded in each session. All four stimulation rate programs were balanced for loudness. A repeated ABCD experimental design was employed.

Take home practice was provided with each stimulation rate. A comparative questionnaire was provided to the CI subjects at the end of the repeated ABCD protocol. Subjects were asked to compare all four rate programs for similar lengths of time over a period of two weeks with a constant sensitivity setting for all stimulation rates.

| Subject | Age | Cause of deafness | Duration of implant use (yr) | Everyday stimulation rate* and strategy. |
|---------|-----|-------------------|------------------------------|--|
| 1       | 58  | Hereditary        | 4                            | 900 pps/ch, ACE                          |
| 2       | 67  | Otosclerosis      | 5                            | 720 pps/ch, ACE                          |
| 3       | 64  | Unknown           | 5                            | 900 pps/ch, ACE                          |
| 4       | 64  | Unknown           | 5                            | 250 pps/ch, SPEAK                        |
| 5       | 74  | Unknown           | 4                            | 1200 pps/ch, ACE                         |
| 6       | 75  | Otosclerosis      | 8                            | 250 pps/ch, SPEAK                        |
| 7       | 62  | Unknown           | 8                            | 250 pps/ch, SPEAK                        |
| 8       | 68  | Unknown           | 6                            | 250 pps/ch, ACE                          |
| 9       | 69  | Unknown           | 4                            | 900 pps/ch, ACE                          |
| 10      | 72  | Unknown           | 5                            | 500 pps/ch, ACE                          |

\*Prior to commencement of the study

**Table 1.** Subject details.

The ACE strategy was used for all stimulation rates. For the 275 pps/ch case, the stimulation rate was jittered in time by approximately 10%, which tends to lower the rate to approximately 250 pps/ch. This was done to minimize the audibility of the constant stimulation rate. It may have been beneficial if all other stimulation rates tested in this study were also jittered (i.e., to avoid a possible confound). The number of maxima was eight for all the conditions. Clinical default settings for pulse width, mode (MP1+2) and frequency to electrode mapping were employed. The pulse width was increased in cases where current level needed to exceed 255 CL units to achieve comfortable levels. The sound processor was set at the client's preferred sensitivity and held constant throughout the study.

Thresholds (T-levels) and Comfortable listening levels (C-levels) were measured for all mapped electrodes and for each rate condition. T-levels were measured using a modified Hughson-Westlake procedure with an ascending step size of 2 current levels (CLs) and a

descending step size of 4 CLs. C-levels were measured with an ascending technique that slowly increases the levels from the baseline T-levels until the client reported that the sound was loud but still comfortable. Loudness balancing was performed at C-levels as well as at 50% level of the dynamic range, using a sweep across four consecutive electrodes at a time. Subjects were asked whether stimulation of all four electrodes sounded equally loud and if not, T- and C-levels were adjusted as necessary. Speech like noise “ICRA” (International Collegium of Rehabilitative Audiology) (Dreschler et al., 2001) was presented at 60 dB SPL RMS for all programs to ensure that each were similar in loudness for conversational speech. The comparison was conducted using a paired-comparison procedure, in which all possible pairings of conditions were compared twice. Adjustments were made to C-levels if necessary to achieve similar loudness across all rate programs.

### 3.3. Results

#### 3.3.1. CNC words

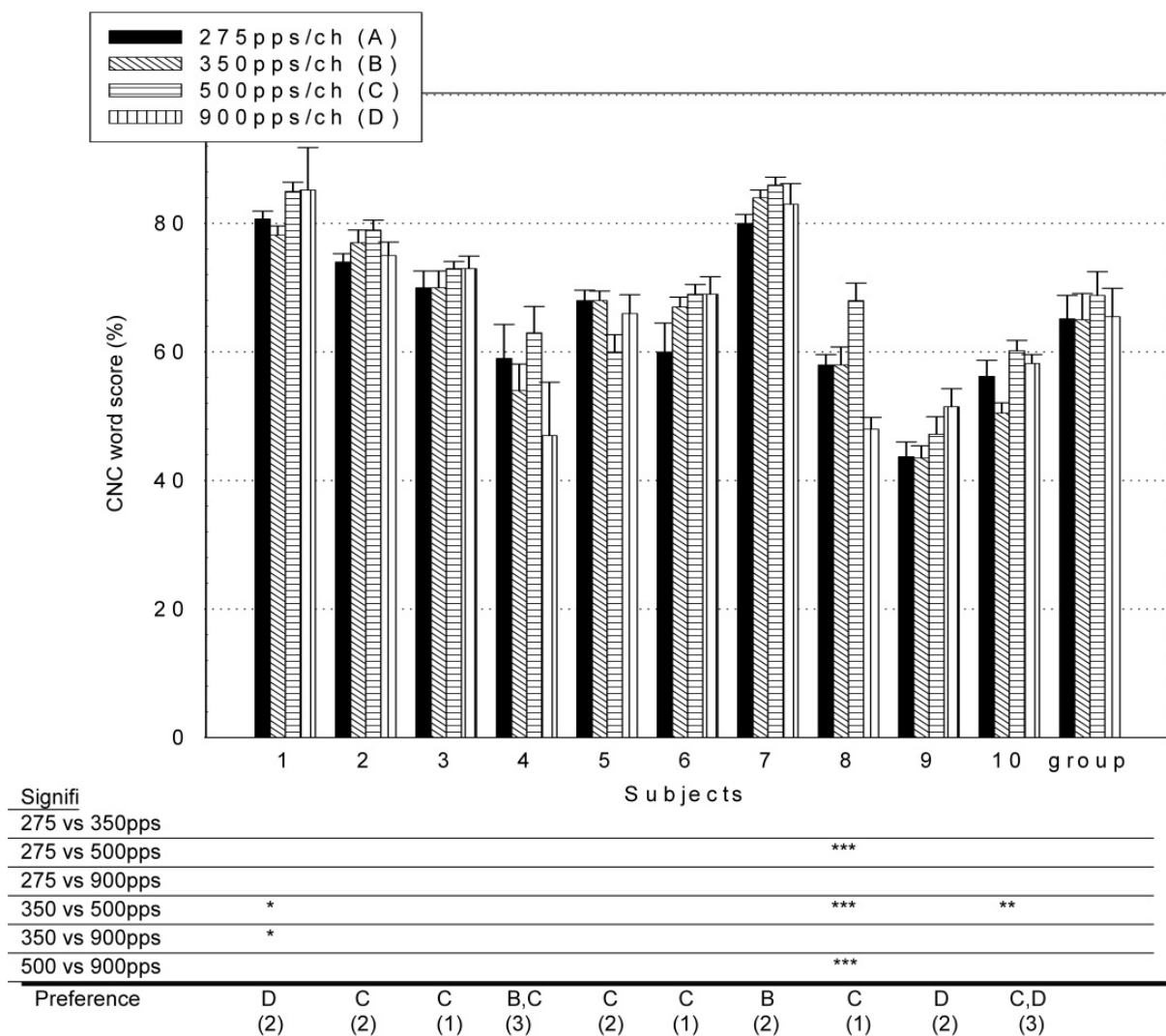
Figure 10 shows percentage correct CNC word scores for the ten subjects for the four stimulation rate programs. The scores were averaged across the two evaluation sessions. Repeated measures two-way analysis of variance (ANOVA) for the group revealed no significant differences across the four rate programs ( $F [3, 27] = 2.14$ ;  $p = 0.118$ ). Furthermore, there was no significant main effect for session ( $F [3, 27] = 2.05$ ;  $p = 0.186$ ). The interaction effect between rate and session was not significant ( $F [3, 27] = 2.30$ ;  $p = 0.099$ ).

In the individual analyses, subject 1 showed significantly better scores for the 500 and 900 pps/ch programs compared to the 350 pps/ch program. There was no significant difference between the 500 and 900 pps/ch programs. Subject 8 showed best CNC scores with the 500 pps/ch program but the 900 pps/ch program showed poorer performance compared to all other programs. Subject 10 showed significantly better scores with 500 pps/ch compared to the 350 pps/ch program.

#### 3.3.2. Sentence test results

Figure 11 shows average SRTs obtained for each subject and the group SRT on the SIT sentences for each stimulation rate program. Lower SRT values indicate better speech perception in noise. Repeated measures two-way analysis of variance revealed a significant main effect for stimulation rate ( $F [3, 27] = 7.79$ ;  $p < 0.001$ ). Group analysis showed significantly better SRT with the 500 pps/ch program compared to 275 pps/ch ( $p = 0.002$ ) and 350 pps/ch ( $p = 0.034$ ). Also 900 pps/ch program showed significantly better SRT compared to 275 pps/ch ( $p = 0.005$ ). Eight out of ten subjects showed improved performance with the 500 or 900 pps/ch rate programs. Small but significant learning effects were observed for sentences in noise scores for all four stimulation rate programs ( $F [3, 27] = 9.39$ ;  $p = 0.013$ ). Mean SRT decreased by 0.6 dB during the second session. There was no significant interaction effect between stimulation rate and the evaluation stage ( $F [3, 27] = 2.04$ ;  $p = 0.13$ ).

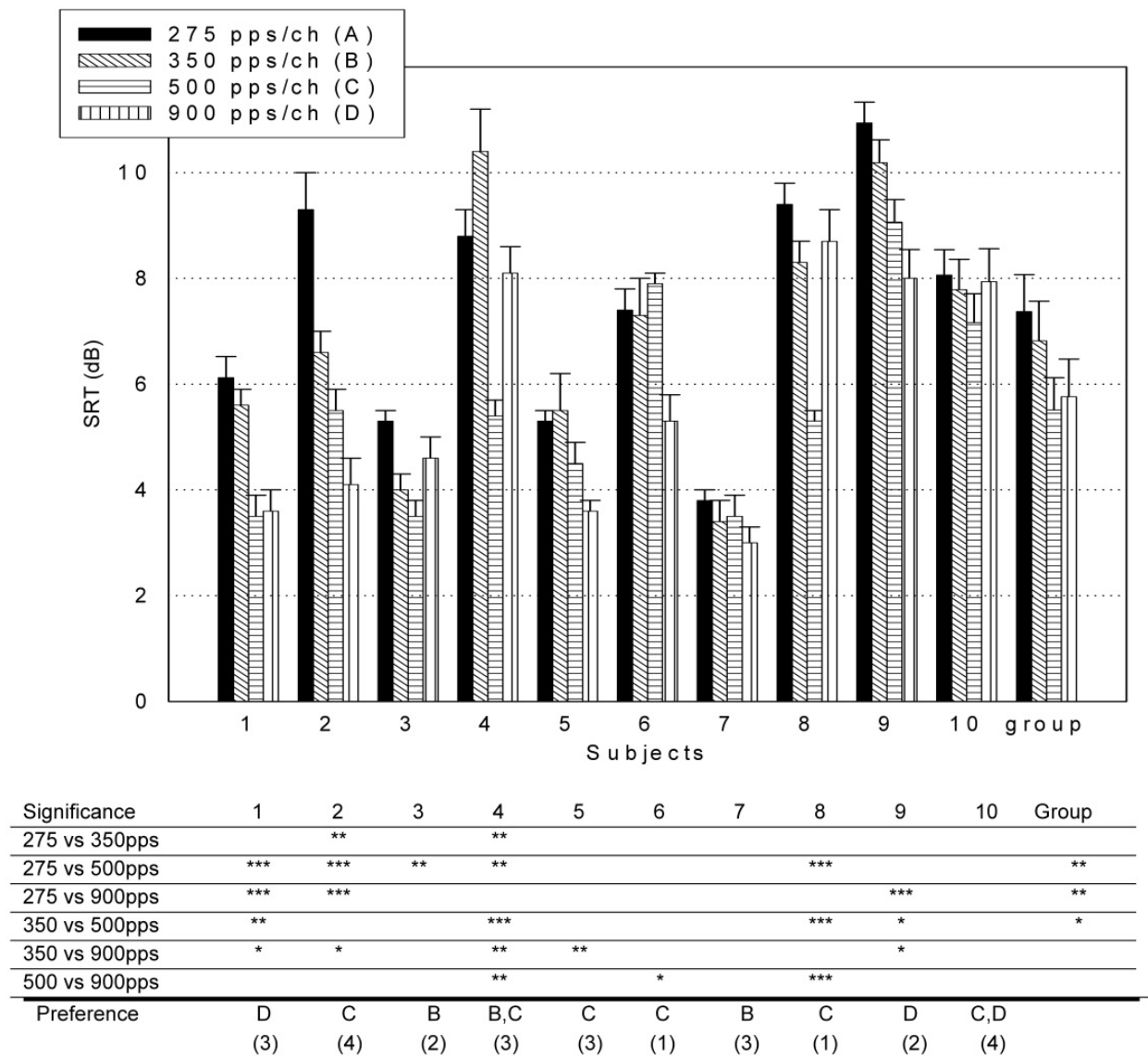




**Figure 10.** Individual patient's percentage correct scores and group mean percentage correct scores for CNC words in quiet. Statistically significant differences (post hoc Tukey test) are shown in the tables presented below each bar graph (\* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ ). Each subject's subjective preference in quiet along with the degree of preference (1 - very similar, 2 - slightly better, 3 - moderately better, 4 - much better) are shown below the chart.

Individual data analysis revealed a significant rate effect for the sentence test ( $p < 0.001$ ) in eight out of ten subjects. All of these eight subjects showed improved performance with the 500 and/or 900 pps/ch rate programs. Subject 1 performed equally well with the 500 and 900 pps/ch stimulation rate programs. The performance was significantly better with both these programs compared to the 275 and 350 pps/ch rate programs ( $p < 0.05$ ). Subject 2 showed improved performance with the 900 pps/ch program. Pair wise multiple comparison with the Tukey test indicated significant differences between the mean SRT obtained with 275 pps/ch program versus all other rate programs ( $p < 0.05$ ), and also between the mean SRT obtained with the 350 and 900 pps/ch programs ( $p = 0.025$ ). No significant differences were observed between the SRTs for the 350 and 500 pps/ch programs and the SRTs for the 500 and 900 pps/ch programs.



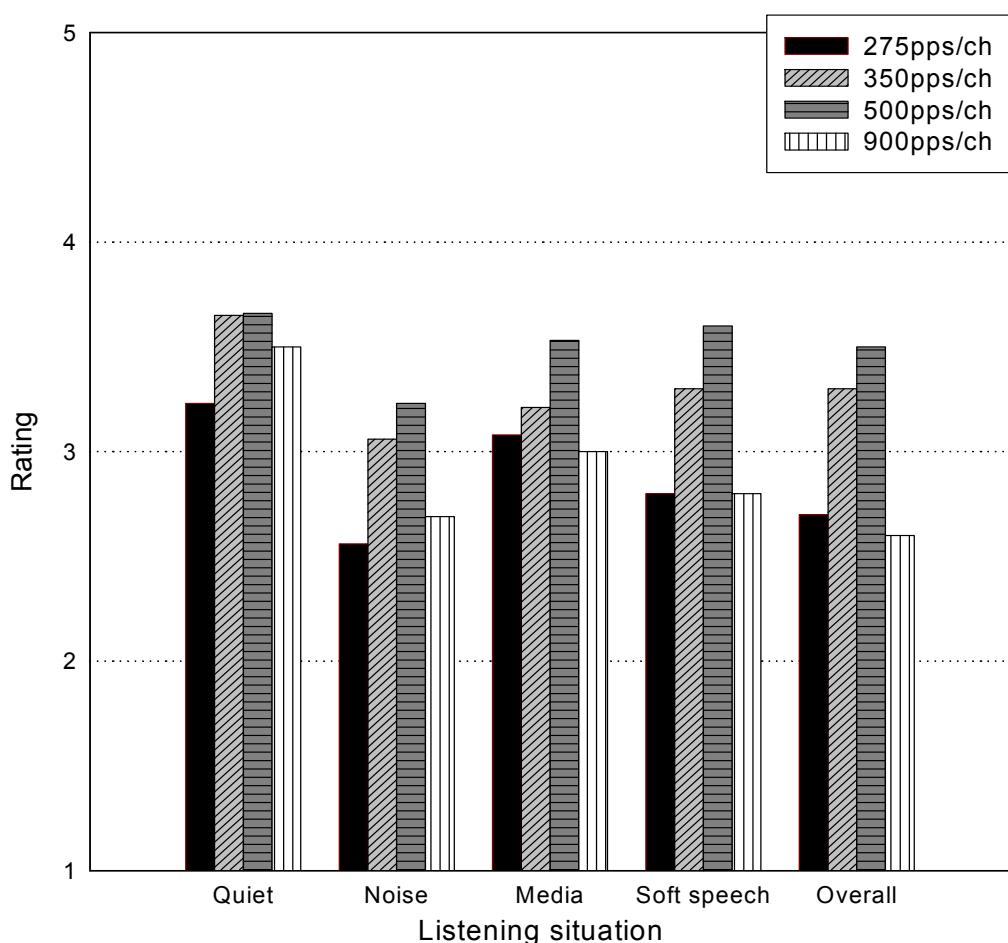


**Figure 11.** Individual patient’s mean speech reception threshold (SRT) and group mean SRT for SIT sentences in competing noise. Statistically significant differences (post hoc Tukey test) are shown in the tables presented below each bar graph (\* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ ). Each subject’s subjective preference in noise along with the degree of preference (1 - very similar, 2 - slightly better, 3 - moderately better, 4 - much better) are shown below the chart.

Subjects 5 and 6 also obtained their best SRTs with the 900 pps/ch compared to 350 and 500 pps/ch stimulation rates ( $p < 0.05$ ). Subject 9 showed improved performance with 900 pps/ch compared to 275 pps/ch ( $p < 0.001$ ) and 350 pps/ch ( $p = 0.01$ ) programs. This subject also showed better SRT for 500 pps/ch compared to 275 pps/ch rate program ( $p = 0.032$ ). Subjects 3, 4 and 8 performed best with the 500 pps/ch stimulation rate. For subject 3, the results for 500 pps/ch condition were significantly better than 275 pps/ch stimulation rate ( $p = 0.001$ ). For subject 4 and subject 8, mean SRTs with 500 pps/ch stimulation rate were significantly better than all other stimulation rates. Subjects 7 and 10 did not show any significant difference in performance when tested for sentences in noise for all four stimulation rates.

### 3.3.3. Comparative performance questionnaire

Figure 12 shows group mean ratings of helpfulness for the stimulation rate programs for the four questionnaire subcategories and averaged across these subcategories. Friedman repeated measures ANOVA on ranks revealed no significant effect of stimulation rate on subjects' average helpfulness ratings across 18 listening situations ( $X^2 [3] = 7.58, p = 0.056$ ). Furthermore, there was no significant effect of rate for the listening categories: listening in quiet ( $X^2 [3] = 1.70, p = 0.63$ ) and listening media devices ( $X^2 [3] = 7.56, p = 0.056$ ). Helpfulness ratings for listening in noise ( $X^2 [3] = 9.16, p = 0.027$ ) and listening to soft speech ( $X^2 [3] = 7.83, p = 0.05$ ) showed a significant effect of stimulation rate, however, pair wise multiple comparisons using Dunn's method revealed no significant differences between any pairs of rate programs for these two categories.



**Figure 12.** Group mean preference ratings of helpfulness for the four rate programs averaged across four categories (listening in quiet, listening in noise, listening media devices & listening to soft speech) and across 18 listening situations (overall). A rating of 1 represented "no help" and a rating of 5 represented "extremely helpful".

After providing helpfulness ratings, subjects were asked to indicate their first preferences in quiet, noise and overall. Table 2 shows the number of subjects reporting their first preferences in quiet, noise and overall for the four rate programs. Chi-square analysis

revealed no significant differences between the distribution of preferences in quiet ( $X^2[5] = 9.24$ ,  $p = 0.099$ ), noise ( $X^2[5] = 5.62$ ,  $p = 0.344$ ) and overall ( $X^2[5] = 5.62$ ,  $p = 0.344$ ). Figures 10 and 11 indicate individual subjects' preferred programs in quiet and in noise respectively.

Subjects were asked to describe if their preferred rate program sounded "very similar", "slightly better", "moderately better" or "much better" than the other programs. As shown in figure 10, five subjects reported their preferred program in quiet to be slightly better than other programs, two reported them moderately better and the remaining three subjects reported them as very similar to other programs. For speech in noise (figure 11), four subjects rated their preferred program in noise as moderately better than other programs; two subjects rated them much better than other programs; two subjects rated them slightly better and the remaining two reported them as very similar to other programs.

### *3.3.4. Relationship between questionnaire results and speech perception outcomes*

The questionnaire results were described in terms of average ratings of helpfulness and the subject's first preferences in quiet, noise and overall. For nine out of the ten subjects there was consistency in the average helpfulness ratings in noise and the subjects' first preferences in noise. There was no close relation between the helpfulness ratings and the first preferences in quiet.

There does not appear to be a close relationship between each subject's subjective preference and the rate program that provided best speech perception. Only two subjects (subject 1 and 8), who scored consistently better on a particular rate program in quiet and noise, chose that program as the most preferred. However, only subject 1 showed consistency between speech test outcomes and helpfulness ratings in quiet and noise. One subject showed consistency between the rate program that provided best speech perception in noise and the most preferred program in noise. Subject 9 scored best with 900 pps/ch rate in noise and preferred this rate in noise. This subject rated 350, 500 and 900 pps/ch equally on helpfulness rating.

For two subjects (subjects 2 and 3) there was a partial agreement between the speech perception scores in noise and the subjective preference. Subject 2 performed best with 900 pps/ch for speech perception in noise, but there was no significant difference in speech performance in noise for 500 and 900 pps/ch. This subject preferred 500 pps/ch stimulation rate in quiet and noise and the average rating of helpfulness in noise was also highest for 500 pps/ch rate. Subject 3 performed best with 500 pps/ch for sentence perception in noise and preferred this program when listening in quiet. This subject preferred 350 pps/ch rate in noise and overall and the average helpfulness rating in noise was highest with this rate program.

Five subjects' (subjects 4, 5, 6, 7 and 10) speech test outcomes did not agree with their subjective preferences. However, the average helpfulness ratings were more or less similar to the first preferences for these five subjects.

At the conclusion of the study, six of the ten subjects (subjects 2, 3, 4, 6, 7 and 10) continued to use a different rate program compared to their everyday rate program (used prior to the commencement of the study). One of these six subjects (subject 6) preferred to continue with the rate program with the best sentence in noise perception score and the remaining five subjects continued with the most preferred program (overall) based on the questionnaire results.

|         | 275 pps/ch | 350 pps/ch | 500 pps/ch | 900 pps/ch | 350 = 500<br>pps/ch | 500 = 900 pps/ch |
|---------|------------|------------|------------|------------|---------------------|------------------|
| Quiet   | 0          | 1          | 5          | 2          | 1                   | 1                |
| Noise   | 0          | 2          | 4          | 2          | 1                   | 1                |
| Overall | 0          | 2          | 4          | 2          | 1                   | 1                |

**Table 2.** The table shows the number of subjects reporting their first preferences in quiet, noise and overall for the four rate programs.

### 3.4. Discussion

#### 3.4.1. Speech perception in quiet and noise

The group averaged scores for monosyllables in quiet showed no significant effect of rate. However, significantly better group results for sentence perception in noise were observed for 500 and 900 pps/ch rates compared to 275 pps/ch stimulation rate and for 500 pps/ch compared to 350 pps/ch rate. Individual data analysis showed improvements with stimulation rates of 500 pps/ch or higher in eight out of ten subjects for sentence perception in noise. Three out of these eight subjects showed benefit with 500 pps/ch and four subjects showed benefit with 900 pps/ch rate. One subject showed improvement with both 500 and 900 pps/ch stimulation rates.

Four out of ten subjects were using 250 pps/ch stimulation rate in their clinical fitted processor before the commencement of the study. Two out of these four subjects showed improved performance with 500 pps/ch, one improved with 900 pps/ch, and the remaining subject showed no effect of rate on speech perception. This suggests that subjects had enough time to become familiar with the higher rate conditions. The remaining six subjects in the study had been using stimulation rates ranging between 500-1200 pps/ch prior to commencement of the study. Four out of these six subjects (including the subject who performed equally with 500 and 900 pps/ch) showed improvement with 900 pps/ch stimulation rate. Better speech perception with 900 pps/ch rate could have been due to the prolonged use of higher stimulation rates prior to commencement of the study.

The CNC test results are somewhat consistent with previous studies that used Nucleus devices with the ACE strategy (Vandali et al., 2000; Holden et al., 2002; Plant et al., 2007 and Weber et al., 2007). In these studies, monosyllabic word or consonant perception was not affected by the increasing stimulation rates. Results in this study are also somewhat consistent with a recent clinical trial by Cochlear Ltd. (2007) (Reference Note 1), which

showed no significant difference in the lower (500-1200 pps/ch) or higher set of rates (1800-3500 pps/ch) for the subjects tested with CNC words.

Most of the previous studies that used Nucleus devices have reported variable rate effects for CI individuals. Some of the subjects in these studies have shown improvement with increased stimulation rates for some of the speech material. Therefore, these studies emphasize the importance of optimizing stimulation rates for individual cochlear implantees. Their results suggest that increasing stimulation rates could provide clinical benefit to some of the cochlear implantees (Vandali et al., 2000; Holden et al., 2002; Plant et al., 2002; Plant et al., 2007). The variable effect of increasing stimulation rates in these studies is consistent with the results of this study, in that not all subjects preferred or showed improved performance with the highest rate (900 pps/ch) tested. For instance, subject 4 and subject 8 performed significantly better with the 500 pps/ch compared to the 900 pps/ch rate. However, in the studies by Vandali et al. (2000) and Holden et al. (2002) the higher stimulation rates probably did not add any extra temporal information due to the limited update rate in SPrint™ processor (760 Hz). In the SPrint™ processor, stimulation rates below 760 Hz provide new information in every cycle because filter analysis rate is set to equal the stimulation rate; however, stimulation rates above 760 Hz are obtained by repeating stimulus frames. Similarly the results of the study by Holden et al. (2002) may have been compromised by the limited analysis rate in the SPrint™ processor. In contrast to these studies, the current study did not use SPrint™ processors. The current study used ESPrit™ 3G processors which have an update rate of 1 kHz for low level sounds and an update rate of 4 kHz for high level sounds.

Analysis of speech perception results across sessions in this study revealed no significant effect of session for the group CNC scores in quiet. However, a significant effect of session was observed for sentence perception in noise scores. Whilst a session effect was observed (which may well have been due to task/rate program learning), scores for all four rate programs showed similar effects of session. Thus given that a balanced design for evaluation of rate was employed in this study, no one rate condition was advantaged by learning within the study.

### 3.4.2. Subjective preferences and speech perception

Some individual variability was also observed in subjective preference results, although the majority of subjects chose 500 or 350 pps/ch rates as their first preferences in quiet, noise and overall.

The individual subjective preference findings in this study initially appear at odds with the results of speech perception outcomes in quiet, in which results for seven of the ten subjects showed no significant effect of rate for monosyllables in quiet. However, six of these seven subjects indicated that there was little difference between their preferred rate program and the other rate programs (see figure 10). On the other hand, three of the five subjects (subjects 1, 2, 4, 8 and 9) who showed some consistencies between the speech perception and the subjective preferences in noise indicated that their preferences were reasonably strong and



that their preferred programs were moderately/much better than the other rate programs (see figure 11). Two of the five subjects, in which inconsistencies between the speech perception in noise and the subjective preference in noise were observed, indicated a weak preference for their preferred rates. The other three, however indicated a strong preference for their preferred rate.

### 3.4.3. *Clinical ramifications*

The present study's findings support previous research, suggesting that optimization of stimulation rate is useful and can lead to better CI recipient outcomes. However, optimization becomes difficult when speech testing outcomes are incompatible with the subject's questionnaire responses. The present study did not reveal a close relation between speech perception outcomes and questionnaire responses. It is possible that the questionnaire data in this study is less likely to be as reliable as the speech perception data. Self reported data in questionnaires is often affected by factors such as how well the subject interprets questions, and compares different rate programs. For example, a recipient may not compare the different rate programs under similar listening conditions. In this study, there was the added difficulty of not being able to toggle between the four rate programs. The recipient instead had to swap processors to compare all four programs. However, all recipients were diligent in taking the two ESPrit™ 3G processors to compare the four rate programs in the 18 listening situations.

Although optimization of stimulation rate appears to be beneficial, time restraints will often prevent clinicians from comparing speech perception outcomes with different stimulation rates. An adaptive procedure called genetic algorithm (Holland, 1975) may offer potential in optimizing stimulation rate along with other parameters. This procedure, based on the genetic "survival of the fittest", guides the recipient through hundreds of processor MAPs towards preferred programs in quiet and in noise. The MAPs vary in terms of speech coding parameters such as, stimulation rate, number of channels, and number of maxima. To date, genetic algorithm (GA) research in experienced CI recipients has not shown better outcomes compared to standard MAPs programmed using default parameters (Wakefield et al., 2005; Lineaweaver et al., 2006). It remains to be seen whether or not the GA algorithm provides significant benefits for newly implanted subjects who are not biased by prolonged use of default parameters such as a particular stimulation rate.

Preference of the majority of subjects in this study for 500 pps/ch rate in quiet, noise and overall is somewhat consistent with the results of Balkany et al. (2007), where 67% of the subjects preferred the slower strategy (ACE) over the faster rate strategy (ACE RE) with the majority of subjects preferring the slowest rate in each strategy (500 pps/ch in ACE and 1800 pps/ch in ACE RE). However, in contrast to the present study, there was no significant effect of rate on the speech perception outcomes in their study.

Subjects who had been using 250 pps/ch stimulation rate for their everyday use prior to commencement of the present study showed improved performance with 500 or 900 pps/ch. In light of this finding, it is recommended that CI recipients who have been using very low



stimulation rates should be mapped with either 500 or 900 pps/ch and given an opportunity to trial the higher rate MAP in different listening environments over a number of weeks.

The findings of previous research (Weber et al., 2007; Cochlear Ltd. 2007, Reference Note 1; Balkany et al., 2007) suggest that for majority of the subjects using Nucleus implants, stimulation rates between 500 pps/ch and 1200 pps/ch should be tried. The present investigation's findings are compatible, suggesting that clinicians should program Nucleus recipients with the rates 500 pps/ch or 900 pps/ch. However, it needs to be remembered that the present study's conclusions are based on a limited number of subjects. Clinicians could consider providing the 500 pps/ch rate as an initial option with the ACE strategy. This rate has the advantage of offering increased battery life compared to the 900 pps/ch rate. Then, if time permits, recipients could compare the 500 pps/ch rate to the 900 pps/ch rate. If for example, the recipient prefers 900 pps/ch and test results in noise show better performance for 900 pps/ch compared to 500 pps/ch, he/she could then be given the opportunity to try 1200 pps/ch.

## **4. Study 2: Effect of stimulation rate on modulation detection**

This section provides the details of the study by Arora et al (2010) which determined whether modulation detection at different stimulation rates predicts speech perception at these rates.

### **4.1. Rationale**

Modulation detection thresholds (MDTs) measured electrically have been found to be closely related to the subjects' speech perception ability with CI and auditory brainstem implant (Cazals et al, 1994, Fu, 2002; Colletti and Shannon, 2005). In addition, studies investigating the effect of relatively high and low stimulation rates on MDTs have shown that MDTs are poorer at high stimulation rates (Galvin and Fu, 2005; Pfingst et al., 2007). The study by Galvin and Fu (2005) showed that rate had a significant effect on the MDTs with lower rates (250 pps/ch) having lower thresholds than the higher rates (2000 pps/ch). Similarly, lower MDTs for 250 pps/ch compared to 4000 pps/ch stimulation rate were observed in the study by Pfingst et al. (2007). These studies suggest that the response properties of auditory neurons to electrical stimulation along with limitation imposed by their refractory behavior must be considered in CI systems (Wilson et al., 1997; Rubinstein et al., 1998). Pfingst et al. (2007) also reported that the average MDTs for 250 pps/ch and 4000 pps/ch were lowest at the apical and the basal end of the electrode array respectively.

Across site variations in modulation detection found by Pfingst et al. (2007, 2008) suggest that testing modulation detection only at one or two sites (as in the studies by Shannon, 1992; Busby et al., 1993; Cazals et al, 1994; Fu, 2002; Galvin and Fu, 2005) may not provide a complete assessment of a CI recipient's modulation sensitivity. In addition, in modern CI sound processors, speech is not coded at one or two specific electrode sites. In current filter bank strategies, many electrode sites are stimulated sequentially based on the amplitude spectrum of the input waveform. In a typical ACE strategy, up to 8 to 10 electrodes are

selected based on the channels with the greatest amplitude. Thus, when measuring MDTs, it may be more realistic to measure them with speech-like signals, which will measure modulation detection across multiple electrodes.

In the current study, modulation sensitivity for low to moderate stimulation rates was measured using acoustic stimuli. Electrode place and intensity coding of the stimuli was representative of vowel-like signals. The vowel-like stimulus stimulated multiple electrodes as in the ACE strategy. It can be argued that MDTs measured across multiple electrodes may be dominated by a few electrodes due to across electrode variations in stimulus levels. However, in the current study all subjects used MP1+2 mode of stimulation and thus the across electrode variations in stimulation levels were small for all subjects. The depth and frequency of sinusoidal amplitude modulation in the stimulus envelope of each channel was controlled in the experiment. Given that it has been found that CI subjects are most sensitive to the modulations between 50-100 Hz (Shannon, 1992; Busby et al., 1993), modulation frequencies of 50 and 100 Hz were examined. Speech recognition has been found to correlate well with MDTs averaged across various stimulation levels of the electrical dynamic range (Fu, 2002). Therefore, this study presented stimuli at an acoustic level that produced electrical levels close to the subjects' most comfortable loudness (MCL) level of stimulation and at a softer acoustic level of 20 dB below this. In previous CI literature, MDTs have been measured using modulated electrical pulse trains; however, in the present study MDTs were measured using acoustic vowel-like stimuli (referred to as Acoustic MDTs in this chapter). Thus for comparison to previous literature, the Acoustic MDTs were transformed to their equivalent current MDTs (referred to as Electric MDTs in this chapter). Acoustic MDTs were of interest because when stimuli are presented acoustically, the differences among different stimulation rate maps are taken into account as in real-life situations for CI users. It is likely that Acoustic MDTs are affected by the subjects' electrical dynamic ranges which can vary with rate of stimulation. This study examined the influence of electrical dynamic range on Acoustic MDTs.

This study was specifically designed to determine:

- Whether rates of stimulation (between 275 and 900 pps/ch) affect modulation detection for vowel-like signals stimulating multiple electrodes.
- Whether modulation detection at different stimulation rates predicts speech perception at these rates.

## 4.2. Method

Modulation detection thresholds were measured for the same 10 subjects who had previously participated in study 1. A repeated ABCD experimental design for the four rate conditions was employed. Evaluation order for rate conditions was balanced across subjects. Four MDT data points were recorded for each rate condition at two modulation frequencies and two stimulation levels (4 data points X 4 rates X 2 modulation frequencies X 2 levels) in each phase of the repeated experimental design.

A sinusoidally amplitude modulated acoustic signal with a carrier frequency of 2 kHz was presented to the audio input of a research processor. The research processor maps were based on the map parameters used in study 1. Most strategy parameters (e.g. number of maxima, pulse width, mode) were kept the same as those used in study 1. However, the maps differed from conventional maps in that only one band-pass filter with a bandwidth of 1.5 to 2.5 kHz (center frequency = 2 kHz) was mapped to all the active electrodes in the map. This was done so that all electrodes received the same temporal information for the test stimuli. The electrical threshold and comfortable levels of stimulation for each electrode were taken from the maps used in study 1.

The signal was used to modulate the envelope of electrical pulse trains interleaved across eight electrode sites. The choice of which 8 electrodes were activated in the maps was based on which electrodes were on average activated in conventional maps for the vowels /a/ and /i/ spoken by a male Australian speaker. This was done by analyzing the spectrograms of each vowel (four different tokens per vowel) and measuring the spectral magnitude at frequencies which coincided with the center frequencies of the bands used in the conventional maps. Two separate vowel maps, one for each vowel, with different sets of fixed electrodes were created. The SAM acoustic stimuli when presented through the experimental map thus provided vowel-like place coding and a SAM temporal envelope code on each channel activated. In addition, the modulation frequency and depth could be controlled systematically via the input SAM signal. For convenience this stimulus will be referred to as a vowel-like SAM stimulus throughout this study.

Modulation frequencies of 50 and 100 Hz were presented at an acoustic level that produced electrical levels close to the subjects' most comfortable level (MCL) of stimulation and at an acoustic level 20 dB below this. Modulation depth was varied in the 3AFC task to obtain a threshold level where the subject could discriminate between the modulated and unmodulated waveform for a particular modulation frequency. A jitter of  $\pm 3$  dB was applied to minimize any loudness effects on measurement of MDTs.

### 4.3. Results

#### 4.3.1. Stimulation rate effect on electrical dynamic range

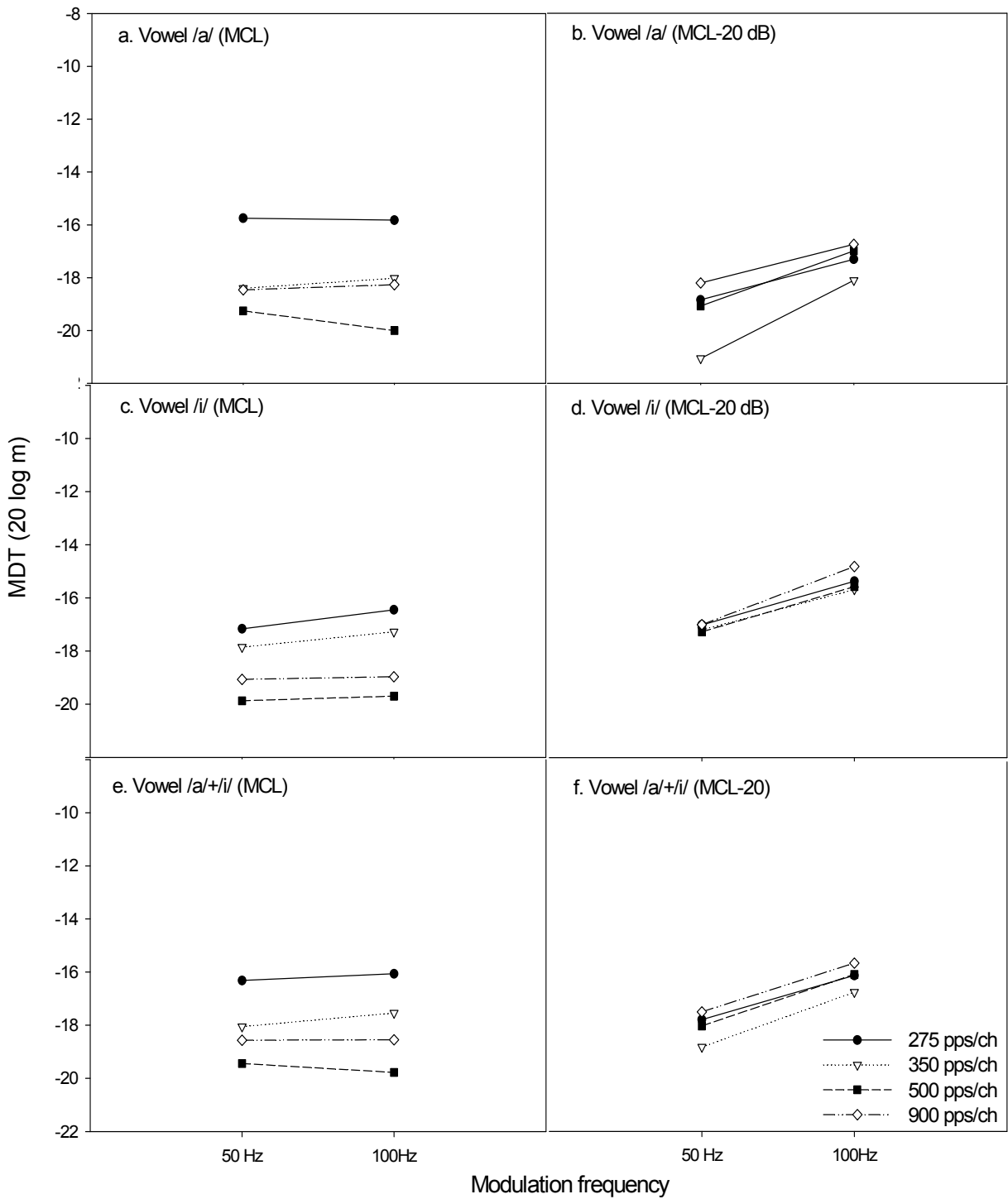
As stimulation rate increased from 250 to 900 pps/ch, mean DR (averaged across the eight most active electrodes that were selected as maxima when coding vowels /a/ and /i/) increased from 40.5 to 51.7 CL (or from  $\sim 6.9$  dB to 8.9 dB in current) for the vowel /a/ map and from 37.6 to 47.8 CL (or from  $\sim 6.5$  dB to 8.2 dB in current) for the vowel /i/ map. These levels were obtained after all four rate programs were balanced for loudness.

#### 4.3.2. Effects of stimulation rate, modulation frequency and presentation level on MDTs

##### 4.3.2.1. Acoustic MDTs

Figure 13 shows Acoustic MDTs for the two modulation frequencies and four stimulation rate conditions measured at MCL in the left most panels (a, c, and e) and at MCL-20 dB in

the right most panels (b, d, and f). MDTs are shown separately for each vowel in the top and middle panels (a and b for vowel /a/, c and d for vowel /i/) and averaged across both vowels in the bottom panels (e and f).



**Figure 13.** Acoustic MDTs, averaged across the subject group, measured at MCL and MCL-20 dB for two modulation frequencies and four stimulation rates.

Repeated measures analysis of variance for Acoustic MDTs averaged across the two vowels revealed a significant effect of rate ( $F [3, 27] = 3.6, p = 0.026$ ). The Mauchley test of Sphericity showed that sphericity was violated for the rate effect. However, the effect remained significant after the G-G (Greenhouse and Geisser, 1959) correction was applied to the degrees of freedom and the  $p$  values.

Post-hoc comparisons for the effect of rate revealed significantly lower MDTs for 500 pps/ch compared to 275 pps/ch rate. There were no significant effects of the other main factors, “modulation frequency” and “level”, on MDTs. The interaction between rate and modulation frequency was not significant, but there was a significant interaction between rate and level. Post-hoc comparisons revealed significantly lower MDTs for 500 and 900 pps/ch rates compared to 275 pps/ch rate at MCL, but no significant effect of rate at MCL-20 dB. MDTs at MCL were significantly lower than those at MCL-20 dB for the rate of 500 pps/ch. The interaction between modulation frequency and level was significant. MDTs for 50 Hz were significantly lower than those for 100 Hz modulation at MCL-20 dB and MDTs at MCL were significantly lower compared to those at MCL-20 dB for the modulation frequency of 100 Hz. A similar pattern of results to those above were observed for separate analyses of each vowel.

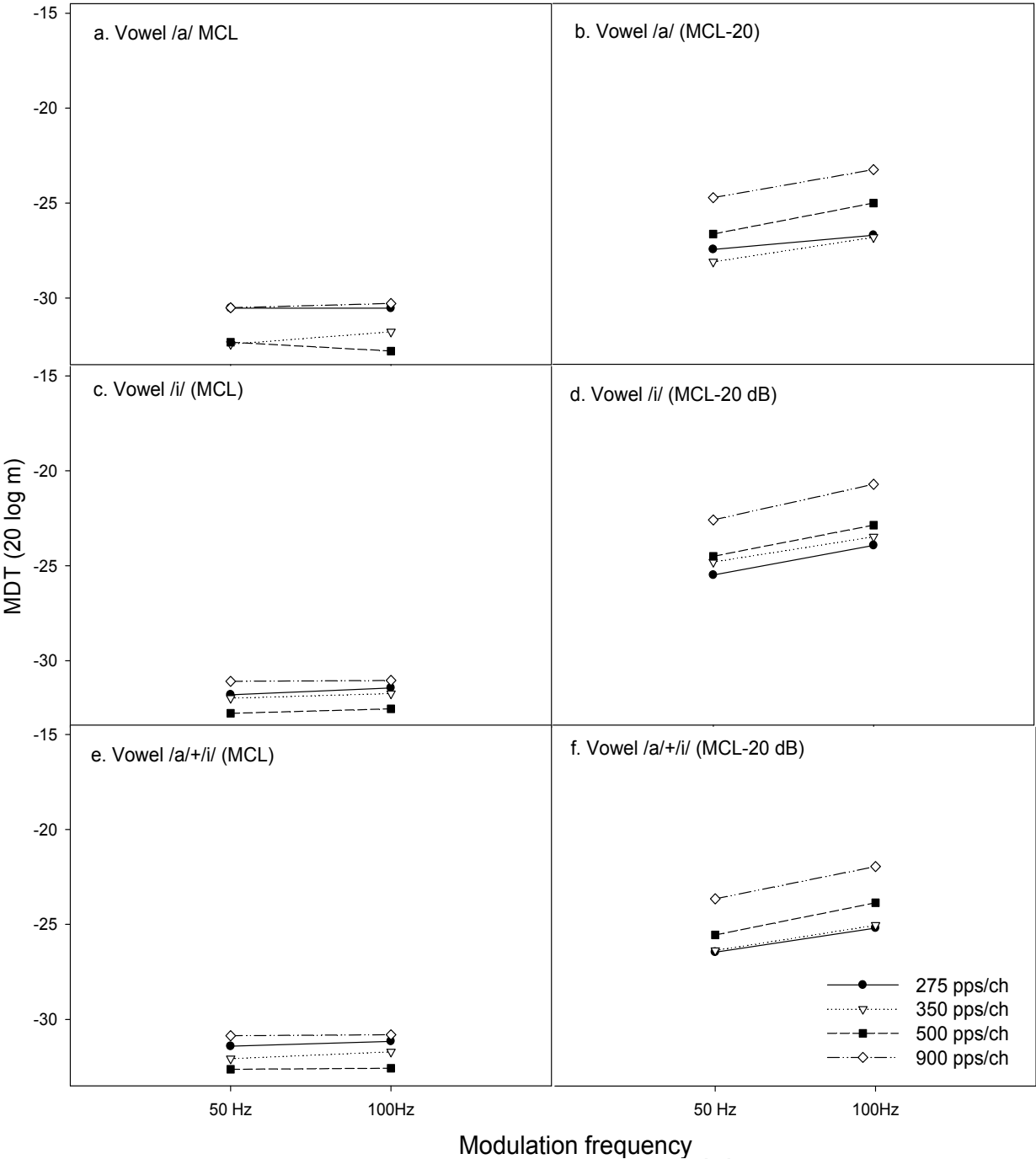
#### 4.3.2.2. Electric MDTs

Figure 14 shows Electric MDTs, averaged across the subject group, for the two modulation frequencies and four stimulation rate programs measured at MCL and MCL-20 dB level. Repeated measures three-way analysis of variance for Electric MDTs averaged across the two vowels revealed a significant effect of rate ( $F [3, 27] = 3.54, p = 0.028$ ), modulation frequency ( $F [1, 27] = 6.66, p = 0.030$ ), and level ( $F [1, 27] = 78.88, p < 0.001$ ). The sphericity assumption was violated for the effect of rate; however, the effect remained significant after the G-G correction was applied.

Post-hoc comparisons for the rate effect revealed no significant comparisons between pairs of stimulation rates. Post-hoc comparisons for the effect of modulation frequency revealed significantly lower MDTs for 50 Hz compared to those for 100 Hz modulation frequency and the post-hoc comparisons for the effect of level revealed significantly lower MDTs at MCL compared to those at MCL-20 dB.

The interaction between rate and modulation frequency was not significant. Interaction between rate and level was significant. At MCL-20 dB, MDTs for 900 pps/ch rate were significantly poorer than all other stimulation rates whereas there was no significant effect of rate at MCL. MDTs at MCL were significantly lower than those at MCL-20 dB for all stimulation rates. The interaction effect between modulation frequency and level was also significant. At MCL, there was no significant effect of modulation frequency on MDTs whereas at MCL-20 dB, MDTs for 50 Hz were significantly lower than those at 100 Hz modulation frequency. MDTs at MCL were significantly lower compared to those at MCL-20 dB for both modulation frequencies (50 and 100 Hz). Again, similar patterns of results to those above at MCL and MCL-20 dB were observed for separate analyses of each vowel.





**Figure 14.** Electric MDTs, averaged across the subject group, measured at MCL and MCL-20 dB for two modulation frequencies and four stimulation rates.

4.3.2.3. Relationship between the speech perception outcomes and MDTs

Analysis of covariance (ANCOVA) to assess the effect of MDTs averaged across both stimulation levels (MCL and MCL-20 dB) on speech perception results (study 1) showed no significant relationships between Acoustic or Electric MDTs and speech perception in quiet and noise. Similar results were obtained on a separate analysis for the Acoustic and Electric MDTs measured at MCL-20 dB.

Results of the ANCOVA showed that MDTs (averaged across 50 and 100 Hz modulation frequencies) at different simulation rates at MCL predicted sentences in noise outcomes (SRTs) at these stimulation rates ( $F [1, 29] = 9.26, p = 0.005$ ). Lower MDTs were associated with lower SRTs. The ANCOVA results also revealed that the estimate for the average slope for subjects was 0.35 ( $se = 0.11$ ). There were no significant effects of Acoustic or Electric MDTs measured at MCL on speech perception in quiet (CNC scores).

ANCOVA results to assess the effect of electrical dynamic range on SRTs showed a significant relationship between electrical dynamic range and SRTs ( $F [1, 29] = 5.52, p = 0.026$ ). The estimate for the average slope for subjects was -0.084 ( $se = 0.035$ ).

## 4.4. Discussion

### 4.4.1. Effects of modulation frequency and presentation level on MDTs

Both Acoustic and Electric MDTs were significantly lower for 50 Hz compared to 100 Hz modulation frequency at the lower presentation level (MCL-20 dB). These results are somewhat consistent with the findings of previous studies in which a progressive increase in modulation thresholds for modulation frequencies above approximately 100 Hz has been reported (Shannon, 1992; and Busby et al., 1993). The Electric MDTs (averaged across 50 and 100 Hz modulation frequencies) presented at MCL were either equivalent to, or better than, those at MCL-20 dB for all rates examined. The Acoustic MDTs at MCL were significantly better than those at MCL-20 dB for 500 pps/ch for MDTs averaged across 50 and 100 Hz. These findings are compatible with those of previous studies (Shannon, 1992; Fu, 2002; Galvin and Fu, 2005; Pfungst et al., 2007).

### 4.4.2. Effects of stimulation rate on MDTs

Significant effects of stimulation rate were found for MDTs; however the effects varied between presentation levels and between how MDTs were defined. For Electric MDTs, which were derived from the modulation depth of current levels (in  $\mu A$ ) averaged across the stimulated electrodes, best MDTs were observed for rates of 500 pps/ch or lower, and poorest MDTs were obtained at 900 pps/ch. This trend was significant at the lower presentation level (MCL-20 dB). In contrast, for Acoustic MDTs derived directly from the modulation depth in the SAM stimulus; the effect of rate was somewhat different and varied across level. At the high presentation level (MCL), best MDTs were observed for rates of 500 pps/ch or higher, and poorest thresholds were observed at the lower stimulation rates. No significant effect of rate was observed at MCL-20 dB.

The differences across rates between the Acoustic and Electric MDTs can be partially attributed to differences in the range of electrical current levels employed across each rate map. Because the electrical dynamic range employed in maps increased with increasing stimulation rate, a smaller depth of modulation in the acoustic input signal is required for higher rate maps in order to produce the same range of modulation in electrical current levels coded for each rate map. Thus, the effects observed for MDTs expressed in acoustic

levels in which modulation sensitivity was better at 900 pps/ch compared to 350 and 275 pps/ch, can be partly accounted for by the increase in electrical dynamic range that was coded in the higher rate maps, at least for MDTs measured at the higher presentation level. At the lower presentation level, it is likely that the increased dynamic range at 900 pps/ch rate could not compensate for the poorer Electric MDTs obtained at that rate.

It can be argued that differences in the absolute current levels at each rate examined (due to different T and C levels for each rate program) might have affected the MDT results. This effect could be more pronounced at MCL-20 dB level, because the effects of stimulation rate on loudness summation are larger at lower stimulation levels (McKay and McDermott, 1998; McKay et al., 2001) which is consistent with the reduction in T levels with increasing rate noted in the current study. However, care was taken to loudness balance all rate programs and thus absolute differences in current levels coded for each rate condition are unlikely to have translated to substantial differences in loudness across rates.

There is some psychophysical evidence (Galvin and Fu, 2005 and Pfingst et al., 2007) to support the findings of the current study, although these studies explored a different range of stimulation rates and modulation frequencies. They did report better MDTs for lower stimulation rates (250 pps/ch) compared to the higher stimulation rates ( $\geq 2000$  pps/ch) which is consistent with the poorer Electric MDTs found at 900 pps/ch compared to the lower rates in the present study.

#### *4.4.3. Relationship between modulation detection and speech perception*

This study did not observe significant relationship between speech perception and Acoustic or Electric MDTs averaged across MCL and MCL-20 dB. This finding is inconsistent with the previous findings which reported a significant correlation between speech perception scores and average modulation detection thresholds at various stimulation levels across the dynamic range (Fu, 2002; Luo et al., 2008). This difference may be in part attributed to the fact that the mean MDTs which were measured through direct electrical stimulation in these previous studies across the various levels of dynamic range are not comparable to the mean MDTs measured at only two stimulation levels in the present study. These studies did not report relationships between speech perception and MDTs at specific stimulation levels. In addition, stimulation rate was not examined in these studies.

For modulation detection measured at MCL, significant effects of Acoustic MDTs and electrical dynamic range (DR) on speech recognition in noise were found in the present study. Acoustic MDTs were of interest, because for both speech and modulation tests, the stimuli were presented through sound processor maps and thus the effects of electrical stimulation level differences between maps with different dynamic ranges were taken into account. Furthermore, a positive correlation between electrical DR and speech test results in noise suggests that the increase in electrical DR with rate contributed to the increase in speech test scores in noise with rate, at least for rates up to 500 pps/ch. These results were somewhat

consistent with the previous findings by Pfingst and Xu (2005) which showed that subjects with larger mean dynamic range had better speech recognition in quiet and noise.

At the highest rate (900 pps/ch), speech perception results were equal to, or worse, than those at 500 pps/ch, particularly for the speech in noise test. In addition, Electric MDTs were poorest at 900 pps/ch compared to 500 pps/ch, which is consistent with findings of other studies where poorer MDTs were observed for higher rates of stimulation (e.g., Galvin and Fu, 2005 and Pfingst et al., 2007). Thus, the benefits to speech perception obtained with an increase in electrical DR may be offset by a reduction in modulation detection sensitivity with increasing rate. The rate for which the effect of increased electrical DR is counteracted by a decrease in modulation detection sensitivity is likely to vary between subjects, speech material, and presentation levels. For the subjects in the current study, although a rate of 500 pps/ch was found to provide the best speech perception results, it is possible that some other rate between 500 and 900 pps/ch, or perhaps even higher, may have provided better speech perception results and/or a better correlation with the electrical DR and modulation detection sensitivity.

At the lower presentation levels (MCL-20 dB), MDTs (Acoustic and Electric) did not correlate with speech perception outcomes. Similar findings were observed by Chatterjee and Peng (2008). In their study no significant correlation was obtained between MDTs measured at soft levels (i.e., at 50% of the dynamic range) and speech intonation recognition presented at comfortable levels.

#### *4.4.4. Clinical ramifications*

An interesting observation of the current study was that the effect of rate on MDTs and speech perception was not monotonic and that lowest (best) MDTs were obtained at a rate of 500 pps/ch. Furthermore, no significant benefit in speech perception was obtained using the higher rate of 900 pps/ch. Thus, clinicians could consider providing the 500 pps/ch rate as an initial option with the ACE strategy, which also has the advantage of offering increased battery life compared to the higher stimulation rates. However, it needs to be remembered that the present study's conclusions are based on a limited number of subjects and that rates between 500 and 900 pps/ch, or even higher, were not examined. Such rates may have provided benefits to speech perception above those obtained at a rate of 500 pps/ch.

## **5. Conclusion**

The above studies investigated the effect of slow and moderate stimulation rates on speech perception and modulation detection in recipients of the Nucleus cochlear implant. Group results for sentence perception in noise showed improved performance for 500 and/or 900 pps/ch stimulation rates but no significant rate effect was observed for monosyllabic perception in quiet. Most subjects preferred the 500 pps/ch stimulation rate in noise. However, a close relationship between each subject's subjective preference and the rate program that provided best speech perception was not observed.

The variable outcomes obtained in the reported studies on stimulation rate could be influenced by factors such as audiological history, length of implant use, and duration of hearing loss, speech processing strategy employed or the implant system used by the implantee. Given the variability of these factors across subjects, it may be important to optimize the stimulation rate for an individual cochlear implantee.

In study 2, cochlear implant subjects' speech perception was compared with their psychoacoustic temporal processing abilities. The aim was to find an objective method for optimizing stimulation rates for cochlear implantees. This study uniquely used multi-channel stimuli to measure modulation detection. Best Acoustic MDTs at MCL were obtained at a rate of 500 pps/ch. MDTs at 900 pps/ch rate were slightly worse and poorest results were observed at the lower rates. No significant effect of rate was observed for Acoustic MDTs at MCL-20 dB. For Electric MDTs at MCL-20 dB, best MDTs were obtained for rates of 500 pps/ch or lower, and poorest MDTs were obtained at 900 pps/ch. Acoustic MDTs are a realistic measure of MDTs since they take into account the map differences in dynamic range across stimulation rates. Acoustic MDTs at MCL at different stimulation rates predicted sentence perception in noise at these rates.

### 5.1. Future research

The ESPrit™ 3G processor offers an option to use a stimulation rate of 720 pps/ch, which was not evaluated in the current study. It would have been interesting to investigate sentence perception in noise with 720 pps/ch for subjects 4 and 8 who showed benefit with 500 pps/ch but showed deterioration with 900 pps/ch. The future work could explore the relationship between speech perception and modulation detection across higher rates of stimulation (>900 pps/ch). In addition, speech perception tests could also be carried out at softer levels to examine possible correlations with MDTs measured at softer levels.

### Author details

Komal Arora and Richard Dowell  
*Department of Otolaryngology, The University of Melbourne, Australia*

Pam Dawson  
*The HEARing CRC, Australia*

### Acknowledgement

The author would like to express appreciation to the research subjects who participated in studies one and two. The studies were supported by University of Melbourne

### Notes

1. Cochlear Ltd. 2007. Selecting stimulation rate with the Nucleus freedom system. White paper prepared by Cochlear Ltd.



## 6. References

- Arora, K., Dawson, P., Dowell, R. C., Vandali A. E. (2009). Electrical stimulation rate effects on speech perception in cochlear implants. *International Journal of Audiology*, Vol 48 (8).
- Arora, K., Vandali A. E., Dawson, P., Dowell, R. C. (2010). Effects of electrical stimulation rate on modulation detection and speech recognition by cochlear implant users. *International Journal of Audiology* Vol 50 (2).
- Assmann, P. F., & Summerfield, Q. (1990). Modeling the perception of concurrent vowels: vowels with different fundamental frequencies. *J Acoust Soc Am*, 88(2), 680-697.
- Bacon, S. P., & Viemeister, N. F. (1985). Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. *Audiology*, 24(2), 117-134.
- Balkany, T., Hodges, A., Menapace, C., Hazard, L., Driscoll, C., Gantz, B., et al. (2007). Nucleus Freedom North American clinical trial. *Otolaryngol Head Neck Surg*, 136(5), 757-762.
- Baskent, D., & Shannon, R. V. (2004). Frequency-place compression and expansion in cochlear implant listeners. *J Acoust Soc Am*, 116(5), 3130-3140.
- Brokx, J. P. L., & Nooteboom, S. G. (1982). Intonation and perceptual separation of simultaneous voices. *Journal of Phonetics*, 10, 23- 36.
- Burian, K. (1979). [Clinical observations in electric stimulation of the ear (author's transl)]. *Arch Otorhinolaryngol*, 223(1), 139-166.
- Burns, E. M., & Viemeister, N. F. (1976). Nonspectral pitch. *J Acoust Soc Am*, 60(4), 863-869.
- Busby, P. A., & Clark, G. M. (2000). Pitch estimation by early-deafened subjects using a multiple-electrode cochlear implant. *J Acoust Soc Am*, 107(1), 547-558.
- Busby, P. A., Tong, Y. C., & Clark, G. M. (1993). The perception of temporal modulations by cochlear implant patients. *J Acoust Soc Am*, 94(1), 124-131.
- Busby, P. A., Whitford, L. A., Blamey, P. J., Richardson, L. M., & Clark, G. M. (1994). Pitch perception for different modes of stimulation using the cochlear multiple-electrode prosthesis. *J Acoust Soc Am*, 95(5 Pt 1), 2658-2669.
- Cazals, Y., Pelizzone, M., Saudan, O., & Boex, C. (1994). Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. *J Acoust Soc Am*, 96(4), 2048-2054.
- Chatterjee, M., & Peng, S. C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hear Res*, 235(1-2), 143-156.
- Clark, G. M., Black, R., Forster, I. C., Patrick, J. F., & Tong, Y. C. (1978). Design criteria of a multiple-electrode cochlear implant hearing prosthesis [43.66.Ts, 43.66.Sr]. *J Acoust Soc Am*, 63(2), 631-633.
- Cohen, L. T., Busby, P. A., Whitford, L. A., & Clark, G. M. (1996). Cochlear implant place psychophysics 1. Pitch estimation with deeply inserted electrodes. *Audiol Neurotol*, 1(5), 265-277.
- Cohen, N. L., and Waltzman, S. B. (1993). Partial insertion of the Nucleus multichannel cochlear implant: Technique and results, *Am. J. Otol*, 14(4), 357-361.
- Colletti, V., & Shannon, R. V. (2005). Open set speech perception with auditory brainstem implant? *Laryngoscope*, 115(11), 1974-1978.
- Cowan, B. (March, 2007). *Historical and BioSafety Overview*. Paper presented at the Cochlear Implant Training Workshop, Bionic Ear Institute, Melbourne.

- Donaldson, G. S., & Nelson, D. A. (2000). Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and SPEAK speech processing strategies. *J Acoust Soc Am*, 107(3), 1645-1658.
- Dorman, M. F., Loizou, P. C., Fitzke, J., & Tu, Z. (1998). The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels. *J Acoust Soc Am*, 104(6), 3583-3585.
- Dynes, S. B., & Delgutte, B. (1992). Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea. *Hear Res*, 58(1), 79-90.
- Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G., & Parkin, J. L. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. *Ann Otol Rhinol Laryngol*, 87(6 Pt 2), 1-39.
- Fearn, R., & Wolfe, J. (2000). Relative importance of rate and place: experiments using pitch scaling techniques with cochlear implants recipients. *Ann Otol Rhinol Laryngol Suppl*, 185, 51-53.
- Fletcher, H. (1940). Auditory Patterns. *Reviews of Modern Physics*, 12, 47-65.
- Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am*, 110(2), 1150-1163.
- Fu, Q. J. (2002). Temporal processing and speech recognition in cochlear implant users. *Neuroreport*, 13(13), 1635-1639.
- Fu, Q. J., Chinchilla, S., & Galvin, J. J. (2004). The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users. *J Assoc Res Otolaryngol*, 5(3), 253-260.
- Fu, Q. J., & Shannon, R. V. (2000). Effect of stimulation rate on phoneme recognition by nucleus-22 cochlear implant listeners. *J Acoust Soc Am*, 107(1), 589-597.
- Fu, Q. J., Shannon, R. V., & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electric hearing. *J Acoust Soc Am*, 104(6), 3586-3596.
- Galvin, J. J., 3rd, & Fu, Q. J. (2005). Effects of stimulation rate, mode and level on modulation detection by cochlear implant users. *J Assoc Res Otolaryngol*, 6(3), 269-279.
- Glasberg, B. R., & Moore, B. C. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J Acoust Soc Am*, 79(4), 1020-1033.
- Holden, L. K., Skinner, M. W., Holden, T. A., & Demorest, M. E. (2002). Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear Hear*, 23(5), 463-476.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*: Arbor, MI: University of Michigan Press.
- Ketten, D. R., Skinner, M. W., Wang, G., Vannier, M. W., Gates, G. A., & Neely, J. G. (1998). In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays. *Ann Otol Rhinol Laryngol Suppl*, 175, 1-16.
- Kiefer, J., Hohl, S., Sturzebecher, E., Pfennigdorff, T., & Gstoettner, W. (2001). Comparison of speech recognition with different speech coding strategies (SPEAK, CIS, and ACE) and their relationship to telemetric measures of compound action potentials in the nucleus CI 24M cochlear implant system. *Audiology*, 40(1), 32-42.
- Kiefer, J., von Ilberg, C., Rupprecht, V., Hubner-Egner, J., & Knecht, R. (2000). Optimized speech understanding with the continuous interleaved sampling speech coding strategy

- in patients with cochlear implants: effect of variations in stimulation rate and number of channels. *Ann Otol Rhinol Laryngol*, 109(11), 1009-1020.
- Loizou, P. C. (1998). Mimicking the human ear. *IEEE Signal Processing Magazine*, 15, 101-130.
- Loizou, P. C., Poroy, O., & Dorman, M. (2000). The effect of parametric variations of cochlear implant processors on speech understanding. *J Acoust Soc Am*, 108(2), 790-802.
- Luo, X., Fu, Q. J., Wei, C. G., & Cao, K. L. (2008). Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users. *Ear Hear*, 29(6), 957-970.
- Magner, M. E. (1972). A speech intelligibility test for Deaf Children (Clark School for the Deaf, Northhampton, MA).
- McDermott, H. J., McKay, C. M., & Vandali, A. E. (1992). A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant. *J Acoust Soc Am*, 91(6), 3367-3371.
- McDermott, H. J. (2004). Music perception with cochlear implants: a review. *Trends Amplif*, 8(2), 49-82.
- McKay, C. M., & McDermott, H. J. (1998). Loudness perception with pulsatile electrical stimulation: the effect of interpulse intervals. *J Acoust Soc Am*, 104(2 Pt 1), 1061-1074.
- McKay, C. M., Remine, M. D., & McDermott, H. J. (2001). Loudness summation for pulsatile electrical stimulation of the cochlea: effects of rate, electrode separation, level, and mode of stimulation. *J Acoust Soc Am*, 110(3 Pt 1), 1514-1524.
- Meddis, R., & Hewitt, M. J. (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *J Acoust Soc Am*, 89(6), 2866-2882.
- Middlebrooks, J. C. (2008). Cochlear-implant high pulse rate and narrow electrode configuration impair transmission of temporal information to the auditory cortex. *J Neurophysiol*, 100(1), 92-107.
- Moore, B. C., & Glasberg, B. R. (2001). Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. *J Acoust Soc Am*, 110(2), 1067-1073.
- Moore, B. C. J., & Carlyon, R. P. (2005). Perception of pitch by people with cochlear hearing loss and by cochlear implant users. In C. J. Plack, A. J. Oxenham, R. R. Fay & A. N. Popper (Eds.), *Pitch: Neural Coding and Perception* (Vol. 24, pp. 234-277): New York: Springer-Verlag.
- Moore, B. (2008). Basic auditory processes involved in the analysis of speech sounds. *Philos Trans Roy Soc B-Bio Sc*, 363(1493), 947-963.
- Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., & Levine, S. (1995). Electrode ranking of "place pitch" and speech recognition in electrical hearing. *J Acoust Soc Am*, 98(4), 1987-1999.
- Nie, K., Barco, A., & Zeng, F. G. (2006). Spectral and temporal cues in cochlear implant speech perception. *Ear Hear*, 27(2), 208-217.
- Parkins, C. W. (1989). Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys. *Hear Res*, 41(2-3), 137-168.
- Pasanisi, E., Bacciu, A., Vincenti, V., Guida, M., Berghenti, M. T., Barbot, A., et al. (2002). Comparison of speech perception benefits with SPEAK and ACE coding strategies in pediatric Nucleus CI24M cochlear implant recipients. *Int J Pediatr Otorhinolaryngol*, 64(2), 159-163.

- Pfingst, B. E., Zwolan, T. A., & Holloway, L. A. (1997). Effects of stimulus configuration on psychophysical operating levels and on speech recognition with cochlear implants. *Hear Res*, 112(1-2), 247-260.
- Pfingst, B. E., Xu, L., & Thompson, C. S. (2007). Effects of carrier pulse rate and stimulation site on modulation detection by subjects with cochlear implants. *J Acoust Soc Am*, 121(4), 2236-2246.
- Pialoux, P. (1976). [Cochlear implants]. *Acta Otorhinolaryngol Belg*, 30(6), 567-568.
- Plant, K., Holden, L., Skinner, M., Arcaroli, J., Whitford, L., Law, M. A., et al. (2007). Clinical evaluation of higher stimulation rates in the nucleus research platform 8 system. *Ear Hear*, 28(3), 381-393.
- Psarros, C. E., Plant, K. L., Lee, K., Decker, J. A., Whitford, L. A., & Cowan, R. S. (2002). Conversion from the SPEAK to the ACE strategy in children using the nucleus 24 cochlear implant system: speech perception and speech production outcomes. *Ear Hear*, 23(1 Suppl), 18S-27S.
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci*, 336(1278), 367-373.
- Rubinstein, J. T., Abbas, P. J., & Miller, C. A. (1998). *The neurophysiological effects of simulated auditory prostheses stimulation*. Paper presented at the Eighth Quarterly Progress Report N01-DC-6 2111.
- Rubinstein, J. T., Wilson, B. S., Finley, C. C., & Abbas, P. J. (1999). Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. *Hear Res*, 127(1-2), 108-118.
- Seligman, P., & McDermott, H. (1995). Architecture of the Spectra 22 speech processor. *Ann Otol Rhinol Laryngol Suppl*, 166, 139-141.
- Seligman, P. (March, 2007). Behind-The-Ear Speech Processors. Paper presented at Cochlear Implant Training Workshop, Bionic Ear Institute, Melbourne.
- Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hear Res*, 11(2), 157-189.
- Shannon, R. V. (1992). Temporal modulation transfer functions in patients with cochlear implants. *J Acoust Soc Am*, 91(4 Pt 1), 2156-2164.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303-304.
- Simmons, F. B. (1966). Electrical stimulation of the auditory nerve in man. *Arch Otolaryngol*, 84(1), 2-54.
- Skinner, M. W., Arndt, P. L., & Staller, S. J. (2002a). Nucleus 24 advanced encoder conversion study: performance versus preference. *Ear Hear*, 23(1 Suppl), 2S-17S.
- Skinner, M. W., Holden, L. K., Whitford, L. A., Plant, K. L., Psarros, C., & Holden, T. A. (2002b). Speech recognition with the nucleus 24 SPEAK, ACE, and CIS speech coding strategies in newly implanted adults. *Ear Hear*, 23(3), 207-223.
- Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., Neely, J. G., Kletzker, G. R., Brunsden, B., and Blocker, B. (2002c). CT-Derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus 22 recipients. *J Assoc. Res. Otolaryngol*, 3(3), 332-350.
- Stark, H., & Tuteur, F. B. (1979). *Modern Electrical Communications*: Englewood Cliffs, NJ: Prentice-Hall.



- Tong, Y. C., Black, R. C., Clark, G. M., Forster, I. C., Millar, J. B., O'Loughlin, B. J., et al. (1979). A preliminary report on a multiple-channel cochlear implant operation. *J Laryngol Otol*, 93(7), 679-695.
- Tong, Y. C., Clark, G. M., Blamey, P. J., Busby, P. A., & Dowell, R. C. (1982). Psychophysical studies for two multiple-channel cochlear implant patients. *J Acoust Soc Am*, 71(1), 153-160.
- Townshend, B., Cotter, N., Van Campenolle, D., & White, R. L. (1987). Pitch perception by cochlear implant subjects. *J Acoust Soc Am*, 82(1), 106-115.
- Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. D., & McDermott, H. J. (2005). Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies. *J Acoust Soc Am*, 117(5), 3126.
- 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 Hear*, 21(6), 608-624.
- Verschuur, C. A. (2005). Effect of stimulation rate on speech perception in adult users of the Med-El CIS speech processing strategy. *Int J Audiol*, 44(1), 58-63.
- Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am*, 66(5), 1364-1380.
- Weber, B. P., Lai, W. K., Dillier, N., von Wallenberg, E. L., Killian, M. J., Pesch, J., et al. (2007). Performance and preference for ACE stimulation rates obtained with nucleus RP 8 and freedom system. *Ear Hear*, 28(2 Suppl), 46S-48S.
- Wilson, B. S. (1991). Better speech recognition with cochlear implants. *Nature*, 352(6332), 236-238.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., & Zerbi, M. (1993). Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. *J Rehabil Res Dev*, 30(1), 110-116.
- Wilson, B. S., Finley, C. C., Lawson, D. T., & Zerbi, M. (1997). Temporal representations with cochlear implants. *Am J Otol*, 18(6 Suppl), S30-34.
- Wilson, B. S., Rebscher, S., Zeng, F. G., Shannon, R. V., Loeb, G. E., Lawson, D. T., et al. (1998). Design for an inexpensive but effective cochlear implant. *Otolaryngol Head Neck Surg*, 118(2), 235-241.
- Xu, L., Thompson, C. S., & Pfungst, B. E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am*, 117(5), 3255-3267.
- Xu, L., Tsai, Y., & Pfungst, B. E. (2002). Features of stimulation affecting tonal-speech perception: implications for cochlear prostheses. *J Acoust Soc Am*, 112(1), 247-258.