

RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 25 (2001-2002)
Indiana University

Cochlear Implants: Signal Processing and Speech Perception¹

Philipos C. Loizou² and Mario A. Svirsky³

*Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405*

¹ This work was supported in part by Grants No. R01 DC03421, R01 DC03739, R01 DC 00421 and R01 DC00064 from the National Institute of Deafness and other Communication Disorders, NIH.

² The University of Texas-Dallas.

³ DeVault Otologic Research Laboratory, Indiana University School of Medicine, Indianapolis, IN.

Cochlear Implants: Signal Processing and Speech Perception

Abstract. Several million Americans today have profound hearing loss, and for years they had to rely on conventional hearing aids. Although hearing aids have been found to benefit hearing impaired individuals who suffer mild to moderate deafness, they do not seem to provide much benefit to individuals with profound (sensorineural) deafness. Today, a prosthetic device, a *cochlear implant*, can be implanted in the inner ear to restore partial hearing to profoundly deaf people. Cochlear implants are now established as a new option for individuals with profound hearing impairments. Most individuals who are implanted with cochlear prosthesis are able to understand speech without lip-reading and can communicate over the phone. This article presents an overview of cochlear implants, describing the signal processing and stimulation strategies they employ and summarizing key research findings.

Background

Hearing and Deafness.

In a healthy ear, sound undergoes a series of transformations as it travels through the outer ear, middle ear, inner ear, auditory nerve and into the brain. The outer ear picks up acoustic pressure waves, which are converted to mechanical vibrations by a series of small bones in the middle ear. In the inner ear, the cochlea, a snail-shaped cavity filled with fluid, transforms the mechanical vibrations to vibrations in fluid. Pressure variations within the fluid of the cochlea lead to displacements of a flexible membrane, called the basilar membrane. These displacements contain information about the frequency of the acoustic signal. Attached to the basilar membrane are hair cells, which are bent according to the displacements of the basilar membrane. The bending of the hairs releases a substance that causes neurons to fire, signaling the presence of excitation at a particular site in the inner ear. These neurons communicate with the central nervous system and transmit information about the acoustic signal to the brain.

The hair cells in conjunction with the basilar membrane are responsible for translating mechanical information into neural information. If the hair cells are damaged, the auditory system has no way of transforming acoustic pressure waves (sound) to neural impulses, resulting in hearing impairment. The hair cells can be damaged by certain diseases (e.g., meningitis, Meniere's disease), congenital disorders, drug treatments, or by many other causes. Damaged hair cells can subsequently lead to degeneration of adjacent auditory neurons. If a large number of hair cells or auditory neurons are damaged, then the condition is called profound deafness. Research (e.g., Hinojosa & Marion, 1983) has shown that the most common cause of deafness is the loss of hair cells rather than the loss of auditory neurons. This was very encouraging for cochlear implantation because remaining neurons could be excited directly through electrical stimulation. A cochlear prosthesis therefore bypasses the normal hearing mechanism (outer, middle, and part of the inner ear, including the hair cells) and electrically stimulates the remaining auditory neurons directly. The challenge we face is finding the optimal signal processing to stimulate (electrically) auditory neurons so that meaningful information about speech is conveyed to the brain. For example, information about the amplitude and the frequency of the acoustic signal should be conveyed.

Encoding Frequency.

The question then arises: "How does the auditory system encode frequencies?" The pioneering work of Georg von Békésy in the 1950's showed that the basilar membrane in the inner ear is responsible for analyzing the input signal into different frequencies. Different frequencies cause maximum vibration

amplitude at different points along the basilar membrane (see Fig. 1). Low frequency sounds create traveling waves in the fluids of the cochlea, which cause the basilar membrane to vibrate, with largest amplitude of displacement at the apex (see Fig. 1) of the basilar membrane. On the other hand, high frequency sounds create traveling waves with largest displacement at the base of the basilar membrane (near the stapes). If the signal is composed of multiple frequencies, then the resulting traveling wave will create maximum displacement at different points along the basilar membrane. The cochlea thereby acts as a spectrum analyzer, decomposing complex sounds into their frequency components.

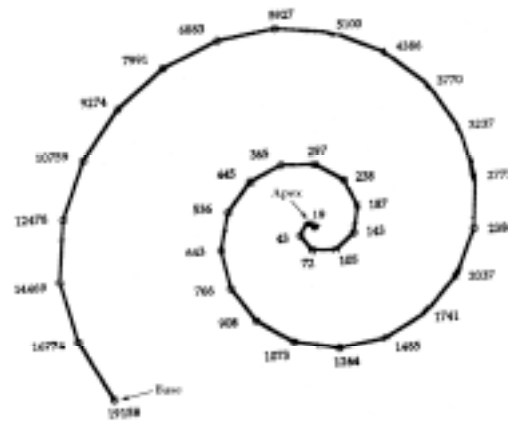


Figure 1. Diagram of the basilar membrane showing the base and the apex. The position of maximum displacement in response to sinusoids of different frequency (in Hz) is indicated.

The corresponding hair cells bent by the displacement in the membrane stimulate adjacent nerve fibers, which are organized according to the frequency at which they are most sensitive. Each place or location in the cochlea is therefore responding "best" to a particular frequency. This mechanism for determining frequency is referred to as *place coding*. The place mechanism for coding frequencies has motivated multi-channel cochlear implants.

Cochlear Implants.

Cochlear implants are based on the premise that there are sufficient auditory nerve fibers remaining for stimulation in the vicinity of the electrodes. Once the nerve fibers are stimulated, they fire and propagate neural impulses to the brain. The brain interprets these impulses as sounds. The perceived loudness of the sound may depend on the number of nerve fibers activated and their rates of firing. If a large number of nerve fibers is activated, then the sound is perceived as loud and vice versa. The number of fibers activated is a function of the amplitude of the stimulus current. The loudness of the sound can therefore be controlled by varying the amplitude of the stimulus current. The pitch, on the other hand, is related to the place in the cochlea that is being stimulated. Low pitch sensations are elicited when electrodes near the apex are stimulated, while high pitch sensations are elicited by stimulation of electrodes near the base. Thus, the implant can effectively transmit information to the brain about the loudness of the sound that is a function of the amplitude of the stimulus current and the sound pitch that is a function of the place in the cochlea being stimulated. Additionally, the rate at which electrical pulses are delivered also affects the perceived pitch.

Several cochlear implant devices have been developed over the years (Wilson, 1993; Loizou, 1998). All of the implant devices have the following features in common: a microphone that picks up the sound, a signal processor that converts the sound into electrical signals, a transmission system that

transmits the electrical signals to the implanted electrodes, and an electrode or an electrode array (consisting of multiple electrodes) that is implanted into the cochlea. In single-channel implants (which are now considered obsolete, due to the inferior perceptual results they provide) only one electrode is used. In multi-channel cochlear implants, an electrode array is inserted in the cochlea so that different auditory nerve fibers can be stimulated at different places, thereby exploiting the place mechanism for coding frequencies. Different electrodes are stimulated, depending on the frequency of the signal. Electrodes near the base of the cochlea are stimulated with high frequency signals, while electrodes near the apex are stimulated with low frequency signals. The signal processor is responsible for breaking the input signal into different frequency bands or channels, and delivering the filtered signals to the appropriate electrodes. The main function of the signal processor is to decompose the input signal into its frequency components, much like a healthy cochlea. The designers of cochlear prosthesis are faced with the challenge of developing signal processing techniques that mimic the function of a healthy cochlea.

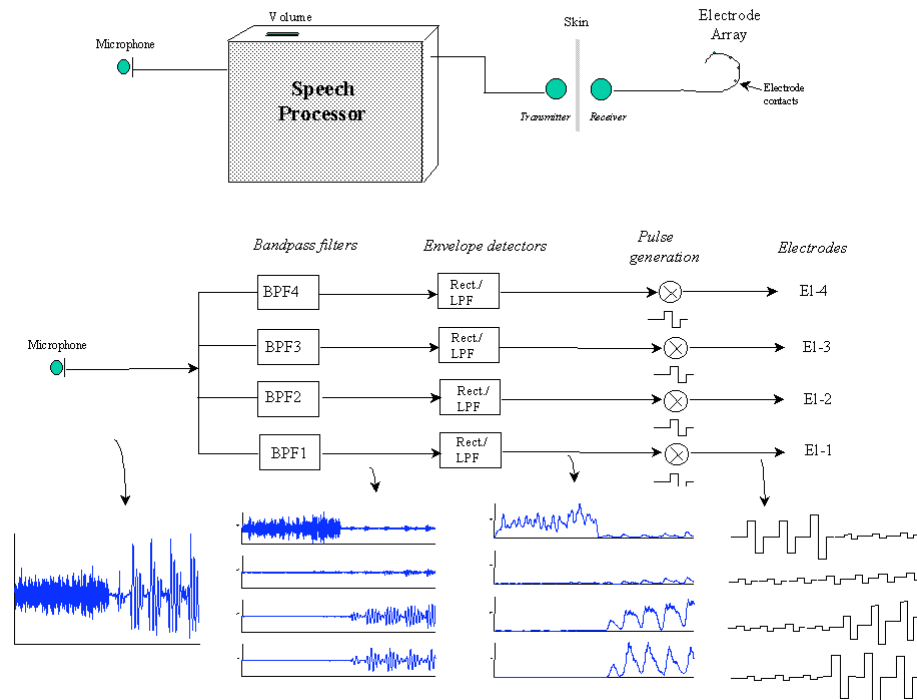


Figure 2. Diagram showing the operation of a four-channel cochlear implant. Sound is picked up by a microphone and sent to a speech processor box worn by the patient. The sound is then processed, and electrical stimuli are delivered to the electrodes through a radio-frequency link. Bottom figure shows a simplified implementation of the CIS signal processing strategy using the syllable "sa." The signal first goes through a set of four bandpass filters that divide the acoustic waveform into four channels. The envelopes of the bandpassed waveforms are then detected by rectification and low-pass filtering. Current pulses are generated with amplitudes proportional to the envelopes of each channel and transmitted to the four electrodes through a radio-frequency link. Note that in the actual implementation, the envelopes are compressed to fit the patient's electrical dynamic range.

Figure 2 shows, as an example, the operation of a four-channel implant. Although the example uses four channels for explanatory purposes, cochlear implants currently in use typically employ between eight and twenty channels. Sound is picked up by a microphone and sent to a speech processor box worn by the patient. Speech processors can be body-worn or behind-the-ear (BTE). In the case of some devices, the BTE version is less flexible than the body-worn processor. After the sound is picked up by the microphone, it is processed through a set of four bandpass filters, which divide the acoustic signal into four channels. Current pulses are generated with amplitudes proportional to the energy in each channel, and transmitted to the four electrodes through a radio-frequency link. The relative amplitudes of the current pulses delivered to the electrodes reflect the spectral content of the input signal (Fig. 2). For

instance, if the speech signal contains mostly high frequency information (e.g., /s/), then the pulse amplitude of channel 4 will be large relative to the pulse amplitudes of channels 1-3. Similarly, if the speech signal contains mostly low frequency information (e.g., vowel /a/) then the pulse amplitude of channels 1 and 2 will be large relative to the amplitudes of channels 3 and 4 (Fig. 2).

Who can be implanted?

Not all people with hearing impairment are candidates for cochlear implantation. Certain audiological criteria need to be met. First, the hearing loss has to be severe or profound and it must be bilateral. Hearing loss is typically measured as the average of pure tone hearing thresholds at 500, 1000 and 2000 Hz, expressed in dB with reference to normal thresholds. Profound deafness is defined as a hearing loss of 90 dB or more, and severe deafness is a hearing loss of 70 dB or more. Second, the candidate should not derive substantial benefit from hearing aids. For adults, this means that they must have acoustically aided scores of 50% or less with the ear to be implanted, and 60% or less with the unimplanted ear (or bilaterally), in a sentence recognition test such as the Hearing-In-Noise-Test (HINT) (Nilsson, Soli & Sullivan, 1994). Typically, the test is administered with no visual cues and in quiet conditions. Children age two years or older with bilateral, severe to profound sensorineural loss are also candidates for cochlear implantation if they receive little or no useful benefit from hearing aids and show a lack of progress in the development of auditory skills. Children between 12 and 24 months of age may also be implanted if they have profound bilateral deafness and do not show progress in the development of auditory skills. In special cases, implantation may be indicated even before 12 months of age. One case is deafness due to meningitis, which may cause ossification of the cochlea, making a delayed implantation more difficult and potentially less successful.

Implant Characteristics.

Commercially available implant devices differ in the following characteristics: Electrode design (e.g., number of electrodes, electrode configuration); type of stimulation (e.g., analog or pulsatile); and signal processing (e.g., waveform representation, envelope representation or spectral features). A brief description of each of the above characteristics is given below.

Electrode Design.

Electrodes are commonly placed in the scala tympani because it brings them in close proximity with auditory neurons that lie along the length of the cochlea. This electrode placement is preferred because it preserves the "place" mechanism of the normal cochlea for coding frequencies. That is, auditory neurons that are "tuned" for high frequencies are stimulated whenever the electrodes near the base are stimulated, whereas auditory neurons that are "tuned" for low frequencies are stimulated whenever the electrodes near the apex are stimulated. In most cases, the electrode arrays can be inserted in the scala tympani to depths of 22-30 mm within the cochlea.

The number of electrodes, as well as the spacing between them, affects the place resolution for coding frequencies. In principle, the larger the number of electrodes, the finer the place resolution for coding. Frequency coding is constrained, however, by two inherent factors: (1) number of surviving auditory neurons that can be stimulated at a particular site in the cochlea, and (2) spread of excitation associated with electrical stimulation. Unfortunately, not much can be done about the first problem, because it depends on the etiology of deafness. Ideally, we would like to have surviving auditory neurons lying along the length of the cochlea. Such a neuron survival pattern would support good frequency representation through the use of multiple electrodes, each stimulating a different site in the cochlea. At the other extreme, consider the situation where the number of surviving auditory neurons is restricted to a small area in the cochlea. In that situation, a few electrodes implanted near that area would be as good as 100 electrodes distributed along the cochlea. So, using a large number of electrodes will not necessarily

result in better performance, because frequency coding is constrained by the number of surviving auditory neurons that can be stimulated.

In addition, frequency coding is constrained by the spread of excitation caused by electrical stimulation. Electric current injected into the cochlea tends to spread out symmetrically from the source. As a result, the current does not stimulate just a single (isolated) site of auditory neurons but several. Such a spread in excitation is most prominent in the monopolar electrode configuration. In this configuration, the active electrode is located far from the reference electrode, which acts as a ground for all electrodes. The spread of excitation can be constrained, to a degree, by using a bipolar electrode configuration. In this configuration, the active and the reference (ground) electrodes are placed close to each other. Bipolar electrodes have been shown to produce a more localized stimulation than monopolar (van den Honert & Stypulkowski, 1987; Merzenich & White, 1977). Although the patterns of electrical stimulation produced by these two configurations are different, it is still not clear which will result in better performance for a particular patient.

Currently, some implant devices employ monopolar electrodes, other devices employ bipolar electrodes, and yet other devices support both types. Table 1 lists some current implant devices and their characteristics. The Nucleus device uses 22 electrodes spaced 0.75 mm apart. Electrodes that are 1.5 mm apart are used as bipolar pairs. The Clarion device provides both monopolar and bipolar configurations, with 8 electrodes spaced 2 mm apart. The Med-El device uses 12 electrodes in monopolar configuration.

| Device | Number of electrodes | Electrode configuration | Type of stimulation |
|---------|----------------------|-------------------------|---------------------|
| Nucleus | 22 | Bipolar | pulsatile |
| Clarion | 8 | Monopolar/Bipolar | analog/pulsatile |
| Med-El | 12 | Monopolar | pulsatile |

Table 1: Characteristics of commercially available cochlear implant devices.

Some companies have developed “modiolar-hugging” electrodes as an upgrade or an alternative to their standard electrodes. Modiolar-hugging electrodes are located closer to the inner wall of the cochlea and therefore are closer to the neural elements that receive electrical stimulation (Cohen et al., in press). It has been demonstrated that this electrode design results in lower perceptual thresholds and comfortable levels, which in turn decreases power consumption and therefore increases battery life (Saunders et al., in press; Parkinson et al., in press). Another potential benefit of modiolar-hugging electrodes is that their stimulation may be more focused on smaller groups of neurons. This may result in better discrimination of place of stimulation in the cochlea (or, equivalently, frequency discrimination) and in improved speech perception. However, this potential benefit has not yet been demonstrated in a clear way.

Type of Stimulation.

Information is presented either in analog or pulsatile form. In analog stimulation, an electrical analog of the acoustic waveform is presented to the electrode. In multi-channel implants, the acoustic waveform is bandpass filtered, and the filtered waveforms are presented to all electrodes simultaneously. The rationale behind this type of stimulation is that the nervous system will sort out and/or make use of all the information contained in the raw acoustic waveforms. One disadvantage of analog stimulation is that its simultaneous action may cause channel interactions. The SAS strategy implemented in the Clarion device is the only strategy that uses analog stimulation.

In pulsatile stimulation, the information is delivered to the electrodes using a set of narrow pulses. In some devices, the amplitudes of these pulses are extracted from the envelopes of the filtered waveforms (Fig. 2). The advantage of this type of stimulation is that the pulses can be delivered in a non-overlapping (i.e., non-simultaneous) fashion, thereby minimizing channel interactions. The majority of the commercial implant devices utilize pulsatile stimulation.

Information is transmitted to the electrodes using a transcutaneous connection, where an external transmitter encodes the stimulus information for radio-frequency transmission from an external coil to an implanted coil. The internal receiver decodes the signal and delivers the stimuli to the electrodes. All cochlear implant devices (e.g., Nucleus, Clarion, Med-El) today use transcutaneous connections. Implanted receivers use a small magnet to facilitate alignment of the external coil to the implanted coil. Although the use of this magnet is well justified, it does present compatibility problems with MRI scanners.

Signal Processing.

The last, and perhaps most important, difference among implant devices is in the signal processing strategy used for transforming the speech signal to electrical stimuli. Some of these strategies are aimed at preserving waveform information, others are aimed at preserving envelope information, and yet others aimed at preserving spectral features (e.g., formants). A more detailed discussion on these signal-processing techniques can be found in Loizou (1998).

Speech Processing Strategies.

Several speech-processing strategies have been developed over the years for multi-channel implants (Loizou, 1998). The effectiveness and performance of these speech processing strategies improved significantly over the years. The Nucleus multi-channel implant, for instance, underwent several changes before developing the latest speech processing strategy, the Advanced Combined Encoder (ACE) strategy. Below, we give the description of the three most common and most successful speech processing strategies used commercially today: the Continuous Interleaved Sampling (CIS), the Simultaneous Analog Stimulation (SAS) and the ACE/SPEAK/n-of-m strategies. The CIS strategy is available in all three implant devices, the SAS is available only in the Clarion device and the ACE/SPEAK/n-of-m strategy is available in the Nucleus-24 and Med-El devices.

Continuous Interleaved Sampling (CIS) Strategy.

The CIS strategy was developed by researchers at the Research Triangle Institute (Wilson et al., 1991). The block diagram of the CIS strategy is shown in Figure 3 (which is a more complete version of the block diagram shown in Figure 2). The signal is first pre-emphasized and then passed through a bank of n bandpass filters, where n typically corresponds to the number of electrodes. The envelopes of the n filtered waveforms are extracted by full-wave rectification and low-pass filtering. The envelope outputs are finally compressed and then used to modulate biphasic pulses. A logarithmic compression function is used to ensure that the envelope outputs fit the patient's dynamic range of electrically evoked hearing. Trains of balanced biphasic pulses, with amplitudes proportional to the envelopes, are delivered to the six electrodes at a constant rate in a non-overlapping fashion i.e., such that only one electrode is stimulated at a time. The rate at which the pulses are delivered to the electrodes has been found to influence speech recognition. High pulse-rate stimulation (i.e., 800 pulses per second, per channel, or higher) typically yields higher performance than low stimulation rate (e.g., Loizou, Poroy & Dorman, 2000). There are also other CIS parameters that may affect performance, and these are discussed in a later section. The CIS strategy is currently being used in the Med-El, the Clarion and the Nucleus CI24M devices. It should be pointed out that the CIS strategy is implemented differently in the various devices, and this can account for some of the differences in performance obtained with the CIS strategy in the commercial devices.

Figure 4 shows, as an example, the pulsatile waveforms produced for the syllable "sa" using a simplified, four-channel implementation of the CIS strategy.

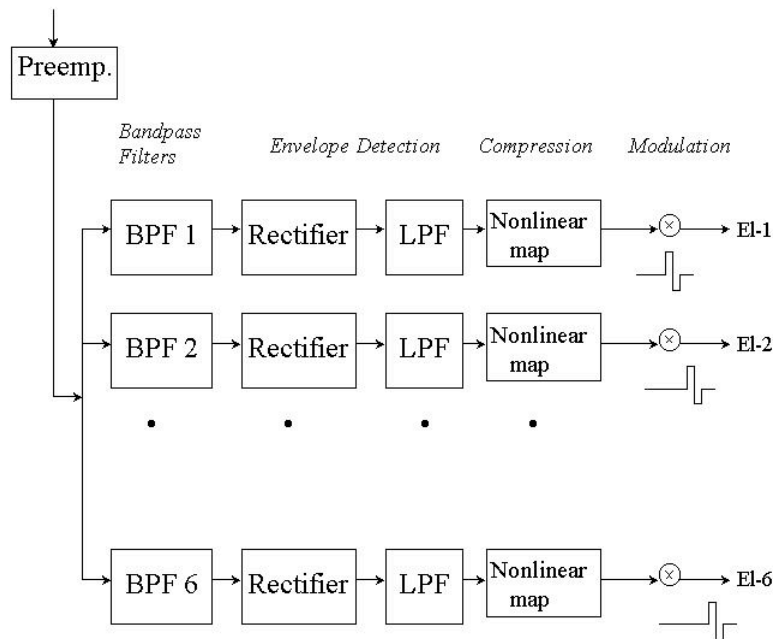


Figure 3. Block diagram of the CIS strategy. The signal is first pre-emphasized and filtered into six frequency bands. The envelopes of the filtered waveforms are then extracted by full-wave rectification and low-pass filtering. The envelope outputs are compressed to fit the patient's dynamic range and then modulated with biphasic pulses. The biphasic pulses are transmitted to the electrodes in an interleaved fashion

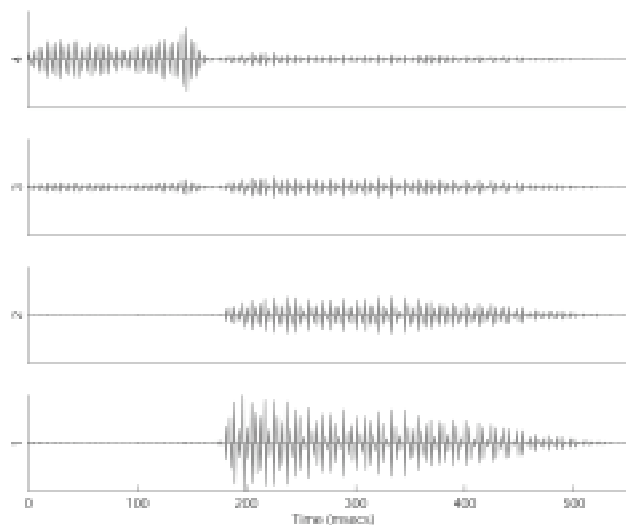


Figure 4. Pulsatile waveforms of the syllable "sa" produced by a simplified implementation of the CIS strategy using a 4-channel implant. The pulse amplitudes reflect the envelopes of the bandpass outputs for each channel. The pulsatile waveforms are shown prior to compression.

In the CIS strategy, the electrodes are stimulated in a non-overlapping manner, i.e., not simultaneously. This is done to address a major concern associated with simultaneous stimulation, which is channel interaction. This interaction is caused by the summation of electrical fields from individual electrodes, particularly when the electrodes are configured in monopolar mode. Neural responses to stimuli from one electrode may be significantly distorted by stimuli from other electrodes. These interactions may distort speech spectrum information and therefore degrade speech understanding.

ACE/SPEAK/N-of-M Strategy.

Unlike previous strategies developed for the Nucleus implant, the SPEAK strategy does not extract any features (e.g., F1, F2) from the speech waveform. Instead it analyzes the speech signal using a bank of 20 bandpass filters and a spectral maxima detector. The signal from the microphone is first pre-amplified and then sent through a bank of 20 bandpass filters with center frequencies ranging from 250 Hz to 10 kHz. The output of each filter is rectified and low-pass filtered with a cutoff frequency of 200 Hz. The SPEAK processor continuously estimates the outputs of the 20 filters and selects the ones with the largest amplitude. The number of maxima selected varies from 5 to 10, depending on the spectral composition of the input signal, with an average number of 6 maxima. The selected electrodes are stimulated at a rate that varies between 180 and 300 Hz depending on: (1) the number of maxima selected and (2) on the patient's individual parameters. The selected amplitudes of the spectral maxima are then logarithmically compressed, to fit the patient's electrical dynamic range, and transmitted to the selected electrodes through a radio-frequency link.

One electrode is allocated for each of the 20 filter outputs, according to the tonotopic order of the cochlea. That is, the most apical electrode is allocated to the filter with the lowest center frequency, while the most basal electrode is allocated to the filter with the highest center frequency. Figure 5 illustrates the pattern of electrical stimulation for the word "choice." As can be seen, the electrodes selected for stimulation in each cycle vary depending upon the spectral content of the signal.

The SPEAK strategy is available in the Nucleus CI24M device. A strategy called n-of-m that is similar to the SPEAK strategy is available in the Med-El device. In the n-of-m strategy, the number of maxima selected in each cycle is fixed. Therefore, "n" corresponds to the number of maxima selected, and "m" corresponds to the total number of channels available. The SPEAK strategy is therefore similar (but not exactly identical) to a 6-of-20 strategy, with an average stimulation rate of 250 cycles per second.

The ACE strategy, described in detail by Vandali et al. (2000) is the latest strategy available for the Nucleus device. It is essentially similar to the n-of-m strategy although some implementation details do differ. For example, the frequency analysis is done using a 128 Hanning window and FFT rather than using an actual software filter bank. Another implementation detail of the ACE strategy is that the FFT's are calculated at a rate of 760 times per second or less. When the stimulation rate is less than 760 cycles per second, the FFT analysis rate is set to equal the stimulation rate, as is normally done in all other DSP-based strategies used by cochlear implant speech processors. However, when the stimulation rate exceeds 760 cycles per second, the analysis rate is limited to 760 times per second and the higher stimulation rate is achieved by repeating stimulation frames as necessary. This compromise is not likely to affect the speech perception of ACE users because the envelopes of the signals coming out of the filterbank change at rates that are typically much lower than 760 Hz (and indeed, as discussed above, these envelopes are usually lowpass filtered using cutoff frequencies of a few hundred Hz at most).

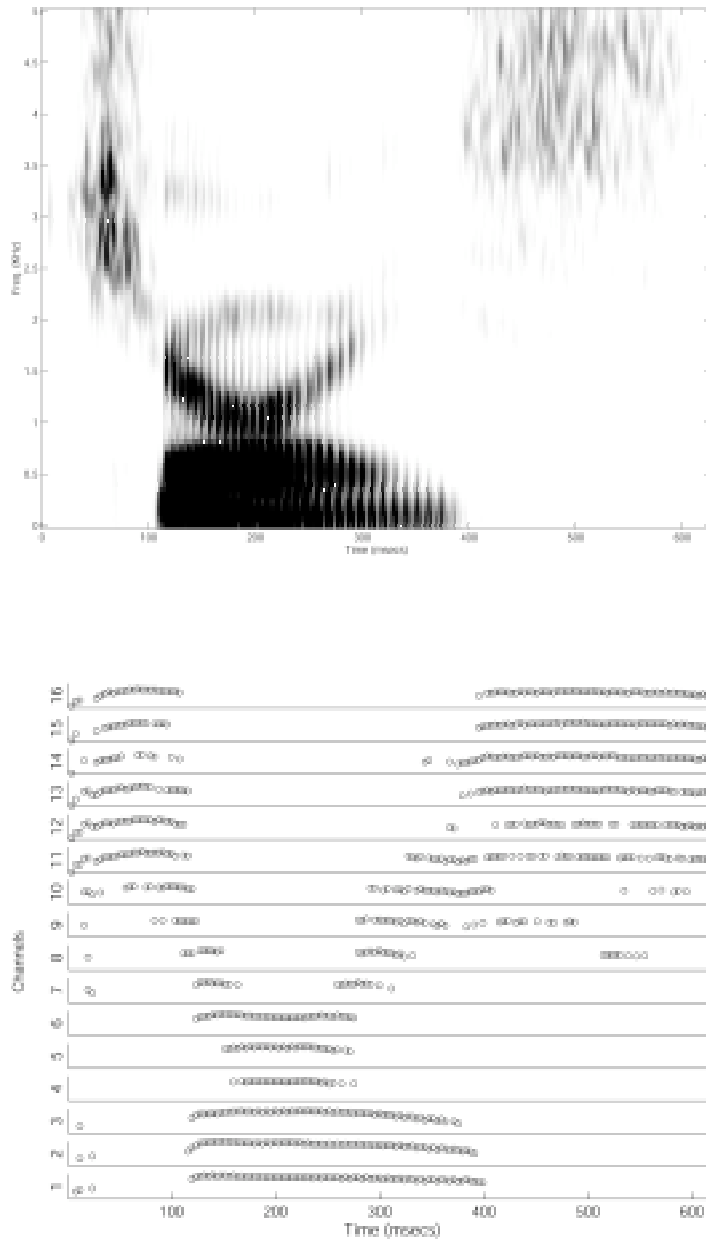


Figure 5. Example of the SPEAK strategy using the word "choice." The top panel shows the spectrogram of the word "choice," and the bottom panel shows the filter outputs selected at each cycle. The channels selected for stimulation depend upon the spectral content of the signal. As shown in the bottom panel, during the "s" portion of the word, high frequency channels (10-16) are selected and during the "o" portion of the word, low frequency channels (1-6) are selected.

Simultaneous Analog Stimulation (SAS) Strategy.

In the SAS strategy (Kessler, 1999), which is available only in the Clarion device, the signal is first pre-emphasized and then passed through a bank of seven bandpass filters. The seven filtered

waveforms are compressed to fit the patient's dynamic range and then used to stimulate seven electrodes simultaneously. Pseudo-analog waveforms are delivered to each electrode at a rate of 13,000 samples/sec per channel. To minimize possible channel interaction, the electrodes are configured in bipolar mode that provides a more selective electrical stimulation pattern than monopolar coupling.

Importance of Fitting: Optimizing Patient Performance.

Cochlear implant manufacturers now offer a multitude of speech processing strategies in their speech processors. It is generally not known, however, which strategy will work the best for a particular patient. As a result, clinicians now have the option to program the patient with multiple strategies and have the patient select the strategy they prefer. In addition, clinicians have the capability, thanks to the flexible fitting software, to change certain speech processing parameters to optimize performance.

For patients fitted with the CIS strategy, clinicians can vary a number of parameters to optimize speech recognition performance for each patient. These parameters include pulse rate, pulse duration and stimulation order.

Pulse Rate and Pulse Duration.

The pulse rate defines the number of pulses per sec (pps) delivered to each electrode. Pulse rates as low as 100 pulses/sec and as high as 2400 pulses/sec have been used commercially. The "optimal" pulse rate for speech recognition varies from patient to patient. Wilson et al. (1995) reported that some patients obtain a maximum performance on the 16-consonant recognition task with 833 pulses/sec and pulse duration of 33 msec/phase. Other patients obtained maximum performance at different combinations of pulse rate and pulse duration (Wilson, Lawson & Zerbi, 1993). Loizou et al. (2000) showed that the performance of some patients on word recognition increased monotonically from 400 pps to 2100 pps. The performance obtained at 2100 pps and with a 40-ms pulse duration, was found to be significantly higher than the performance obtained at 800 pps. However, note that studies conducted with users of the ACE strategy did not find a similar, consistent advantage for higher stimulation rates. Vandali et al. (2000) found that the use of rates higher than 250 cycles per second (up to 1615 cps) did not provide significant improvement in speech comprehension. Holden, Skinner and Holden (in press) found that higher rates resulted in significantly better performance for some subjects and significantly worse performance for other individuals.

Stimulation Order.

The stimulation order refers to the order that electrodes are stimulated. The stimulation order can be varied to minimize possible interaction between channels. One possibility is to stimulate the electrodes in an apex-to-base order, i.e., first stimulate electrode 1, then 2, etc., and lastly, 6. This way, signals in the low frequencies (apex) are stimulated first, and signals in the high frequencies (base) are stimulated last. This apex-to-base order, however, does not minimize the spatial separation between sequentially stimulated electrodes. Alternatively, the electrodes can be stimulated in a so-called "staggered" order, i.e., 6-3-5-2-4-1, which maximizes the spatial separation between stimulated electrodes. As with the pulse rate, preference for stimulation order varies from patient to patient. Some patients prefer the apex-to-base stimulation because they say speech sounds more natural and intelligible while other patients prefer the staggered order stimulation.

For patients fitted with the Nucleus' CI24M device, clinicians have the option to select a subset of electrodes for stimulation. Research has shown that stimulating only a subset of electrodes, rather than all 22 electrodes, can produce significant benefits for some patients fitted with the CIS strategy. Plant et al. (1999) for instance, showed that some subjects preferred and performed better on word recognition with

the CIS strategy when it was programmed with 8 channels than when it was programmed with 16 channels.

For patients fitted with the Clarion device, clinicians have the option to fit the patient with the CIS, the SAS or the Paired Pulsatile Stimulation (PPS) strategies. The PPS strategy is similar to CIS with the exception that two simultaneous pulses are delivered at a time instead of just one. Research has shown that some patients prefer the SAS strategy to the CIS strategy (Osberger & Fisher, 1999; Battmer, Zilberman, Haake & Lenarz, 1999), while other patients prefer the PPS strategy to the CIS strategy (Armstrong-Bendall et al., 1999). In most cases, Clarion patients seem to perform better on speech recognition tasks with the strategy they prefer.

For patients fitted with the Med-El device (COMBI 40+), clinicians now have the option to fit patients with the n-of-m strategy in addition to the CIS strategy. Research has shown that Med-El patients performed better with the 7-of-12 strategy (operating at a higher rate) than the CIS strategy (12 channels) on monosyllabic word recognition (Ziese et al., 2000). This is consistent with research reported by Brill et al. (1997) showing that by trading channels with higher stimulation rates (i.e., by using fewer number of channels to obtain higher rates of stimulation) performance can be improved.

Factors Affecting the Performance of Cochlear Implant Patients.

There is a great variability in the speech recognition performance of cochlear implant patients. For a given type of implant, auditory performance may vary from zero to nearly 100% correct. Auditory performance is defined here as the ability to discriminate, detect, identify, or recognize speech (a typical measure of auditory performance is the percent correct score on open-set speech recognition tests). The factors responsible for such variability in auditory performance have been the focus of research for many years (Blamey et al., 1996). Some of the factors that have been found to affect auditory performance are listed below.

Duration of Deafness.

The duration of deafness prior to implantation has been found to have a negative effect on auditory performance. Individuals with shorter durations of auditory deprivation tend to achieve better auditory performance than individuals with longer durations of auditory deprivation.

Age of Onset of Deafness.

The age of onset of deafness has a major impact on the success of cochlear implants, depending on whether the deafness was acquired before (prelingual) or after (postlingual) learning speech and language. It is now well established that children and adults with postlingual deafness perform better than those with prelingual or congenital deafness. However, this factor interacts with age at implantation, because congenitally deaf children can achieve excellent communication skills if they are implanted early enough. This includes speech perception (Meyer, Svirsky, Kirk & Miyamoto, 1998; Svirsky & Meyer, 1999; Meyer & Svirsky, 2000), speech production (Svirsky, Sloan, Caldwell & Miyamoto, 2000b) and language development (Svirsky, Robbins, Kirk, Pisoni & Miyamoto, 2000a).

Age at Implantation.

Prelingually deafened persons who were implanted in adolescence have been found to obtain higher levels of auditory performance than those implanted in adulthood. People implanted at an early age seem to perform better than people implanted in adulthood. In addition, recent results suggest that congenitally deaf children who are implanted earlier in life have a better communicative prognosis (Svirsky et al., 2000a).

Duration of Cochlear Implant Use.

Duration of experience with the implant has been found to have a strong positive effect on auditory performance for both adults and children.

Other Factors.

Other factors that may affect auditory performance include: (1) number of surviving spiral ganglion cells, (2) electrode placement and insertion depth, (3) electrical dynamic range, and (4) signal processing strategy. There are also factors, such as patient's level of intelligence and communicativeness, which are unrelated to deafness that may also affect auditory performance. It is important to note, however, that much of the substantial intersubject variability that we observe in cochlear implant users remains unexplained.

Commercial Implant Processors.

There are currently three cochlear implant processors in the United States approved by the Food and Drug Administration (FDA), the Nucleus 24, the Clarion processor and the Med-El processor.

Nucleus 24 (CI24M).

The Nucleus 24 is the latest speech processor manufactured by Cochlear Pty. Limited, Australia. The processor can be programmed with either the SPEAK, the CIS or the Advanced Combined Encoder (ACE) strategies. In the SPEAK strategy, the signal is sent through a bank of 20 bandpass filters and out of the 20 filter outputs extracted through envelope detection, the 5-10 filter outputs with the largest amplitude are selected for stimulation. The selected electrodes (six on the average) are stimulated at a rate that varies from 100 to 300 pps. The CIS strategy does not use all 20 frequency bands available with the SPEAK strategy. A single frequency band is dedicated to each of the n electrodes, where n ranges from 4 to 12. The electrodes can be stimulated up to a maximum rate of 2400 pulses/sec depending on the value of n . This is a considerably higher rate than the rate (180-300 pulses/sec) used in the SPEAK strategy. The ACE strategy provides a faster (in terms of stimulation rate) implementation of the SPEAK strategy as the selected channels can be stimulated up to a maximum rate of 2400 pulses/sec.

The Nucleus 24 (CI24M) is available in two sizes, the regular size (SPrint) worn on the waist and the ear-level size worn behind the ear. The ear-level version (ESPrint) is the size of a behind-the-ear hearing aid and can be only programmed with the SPEAK strategy.

Clarion.

The Clarion cochlear implant system is the result of cooperative efforts among the University of California at San Francisco (UCSF), Research Triangle Institute (RTI) and the device manufacturer, Advanced Bionics Corporation (evolved from MiniMed Technologies). The Clarion implant supports a variety of speech processing options and stimulation patterns (Kessler, 1999). The stimulating waveform can be either analog or pulsatile. The stimulation can be simultaneous, sequential or both and the stimulation mode can be either monopolar or bipolar. The processor can be programmed with either the SAS strategy, the CIS strategy or the Paired Pulsatile Sampler (PPS) strategy. In the CIS strategy, pulses are delivered to 8 electrodes at a maximum rate of 833 pulses/sec per channel in an interleaved fashion. The PPS strategy is very similar to the CIS strategy, except that a pair of electrodes is stimulated simultaneously, thereby increasing the maximum rate per channel to 1666 pulses/sec. In the SAS strategy, the acoustic signal is processed through seven filters, compressed and then delivered simultaneously to seven electrode pairs. Pseudo-analog waveforms are delivered to each electrode at a rate of 13,000 samples/sec per channel. The Clarion processor also has the capability of supporting other speech

processing strategies including a hybrid (CIS in some channels and SAS in others) strategy, however these strategies are still under investigation.

Med-El.

The Med-El cochlear implant processor is manufactured by Med-El Corporation in Austria. The Med-El processor (COMBI 40+) uses a 12-electrode array configured in monopolar mode and can be programmed with either the CIS or the n-of-m strategy. In the CIS strategy, the envelope amplitudes extracted from 12 bandpass filters are used to modulate biphasic pulses that are then delivered in an interleaved fashion to 12 monopolar electrodes at a maximum rate of 1,515 pulses/sec per channel.

The Med-El processor can also be programmed with a high-rate n-of-m strategy that is similar to the ACE strategy used in the CI24M device. In the n-of-m strategy, from a maximum of m channels, the n ($n < m$) channels with the highest energy are selected for stimulation in each cycle. Reducing the number of channels allows the stimulation rate per channel to increase, possibly leading to better temporal resolution of the sound signal without compromising spectral resolution. Due to limitations in RF transmission rates, it is sometimes necessary to reduce the number of electrodes to be stimulated in order to obtain higher rates of stimulation.

Clinical Results.

Adults. After the preceding description of cochlear implant systems and the signal processing and stimulation strategies they use, it may be of interest to review the kinds of perceptual benefit that patients typically achieve. Although many adult patients experience immediate results, scores in speech perception tests improve gradually over the first few weeks or months after implantation and then stay more or less constant. Figure 6 shows CNC word identification scores obtained by 29 postlingually deaf cochlear implant users grouped by the device they use. These results were obtained in the laboratory of the second author. All subjects had at least one year of experience with the device. Sets of 50 words were prerecorded and presented in open-set fashion at 70 dB SPL.

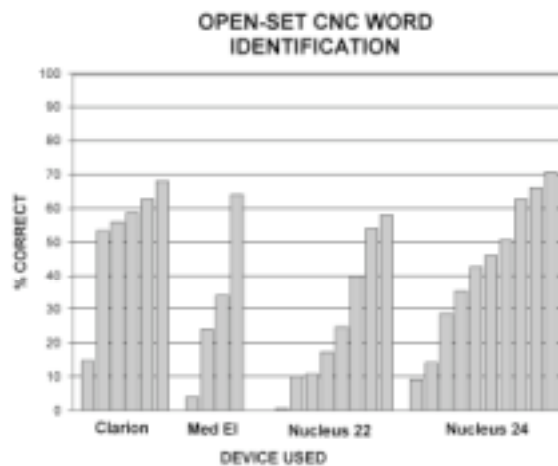


Figure 6. Perception of monosyllabic words without lipreading by postlingually, profoundly deaf adults with cochlear implants. These open-set scores were obtained in quiet, at least one year after initial activation of the patients' speech processors.

Subjects were not selected. Instead, they volunteered for studies that included several other tests. Due to the nature of the study, patients using different devices were not matched along variables that may influence speech perception, such as duration of deafness, or amount of residual hearing. Therefore, it would be inappropriate to use these numbers to compare speech perception scores obtained with different devices. Instead, we can draw two conclusions that are valid for all four devices. First, almost all patients obtain some level of open set speech perception (only one out of 29 subjects was unable to identify any of the words). Second, there are large individual differences among patients. In this group of subjects, there were one Clarion user, one Med-El user, four Nucleus-22 users and two Nucleus-24 users who scored 15% words correct or less. On the other hand, there were several users of all devices who scored over 50% correct. A listener who can identify more than half of the words in a relatively difficult test like this that does not provide any linguistic context or allow for lipreading, probably can communicate in relatively fluent fashion in a face-to-face situation. Furthermore, the second author has observed many of these patients communicating fluently over the telephone, even when they didn't know who was calling or the nature of the call.

It is important to stress that the preceding results are representative of those obtained by postlingually deaf listeners. In contrast, speech perception scores are much worse for prelingually deaf adult cochlear implant users (people who become profoundly deaf before the age of three years, before their language skills are ingrained). Only rarely can a prelingually deafened adult cochlear implant recipient achieve any open-set speech perception. For example, Zwolan et al. (1994) report that none of the 17 subjects they tested were able to demonstrate any open-set speech recognition. Zwolan et al. (1996) also assessed changes in *closed-set* speech perception performance in 11 prelingually deaf adults before they received cochlear implants and one year after activation of the speech processor.

Seven tasks were used and tested for statistically significant differences at a 5% level using the binomial distribution. If changes pre- to post-implantation were completely random, it would be expected that about four comparisons (77 tests times the 5% significance level equals 3.85, which is rounded to 4) would yield “significant” increases and another four comparisons would yield “significant” decreases. Instead, Zwolan et al. found 11 significant decreases and four significant increases. Two of the increases were from 10% to 35% in a four-choice spondee test and from 12% to 30% in a four-choice final consonant test. In other words, the two post-increase scores may not be statistically different from the 25% that would be expected from chance alone. The Zwolan et al. article, which may be the largest most comprehensive and careful study of this population, found that most prelingually deafened patients who receive cochlear implants as adults used their devices regularly, were satisfied with them, and reported communicative benefit from its use. Nevertheless, it is clear that long-term sound deprivation that includes the first few years of life has a large, deleterious effect on speech perception after cochlear implantation. But what if a patient is congenitally or prelingually deaf and receives a cochlear implant within a few years after hearing loss, so that the period of sound deprivation is minimized?

Children. As we have indicated above, it has been well documented that implantation of congenitally or prelingually deaf children results in substantial benefit to speech perception (Meyer et al., 1998; Svirsky & Meyer, 1999; Meyer & Svirsky, 2000), and speech production (Svirsky et al., 2000b). Waltzman et al. (1997), in a study of 38 consecutively implanted, prelingually deafened and orally educated children showed that all subjects achieved significant open-set speech recognition. However, in the case of children who are born profoundly deaf, the auditory signal provided by a cochlear implant may also provide enough information for the development of oral language skills. In other words, one of the main potential outcomes of pediatric cochlear implantation (in addition to improvement of speech perception and speech production skills) is the development of skills in an oral language such as English. Numerous studies have found that prelingually deaf children lag in their English language abilities with respect to normal-hearing children.

A recent study (Svirsky et al., 2000a) assessed the expressive language of 23 prelingually deaf children before receiving cochlear implants as well as 6, 12 and 18 months after initial stimulation with a

Nucleus-22 device. Figure 7 shows the most important results of that study. The vertical axis indicates the children's expressive language skills, expressed in terms of age equivalent scores, and the horizontal axis shows their chronological age. Therefore, the linguistic development of an average child with normal hearing should follow the diagonal line. Black circles indicate the mean language age of the 23 patients at all four testing intervals and the white circles indicate the language age that would be expected at each interval if the children had not received cochlear implants. It is observed that at the preimplant interval, the children who would later receive cochlear implants were delayed with respect to children with normal hearing (because the first black circle is well below the diagonal). It may also be observed that these implanted children developed linguistic skills faster than would have been predicted for deaf children without implants (because the black circles fall above the white circles). Analysis of these and other data showed that language development over 2.5 years after implantation kept pace with the development that would have been expected from normal-hearing children at the same starting point of linguistic skills. In other words, the language gap that was present at implantation remained the same size (when measured in terms of language age) instead of increasing month after month as would be expected from deaf children without cochlear implants.

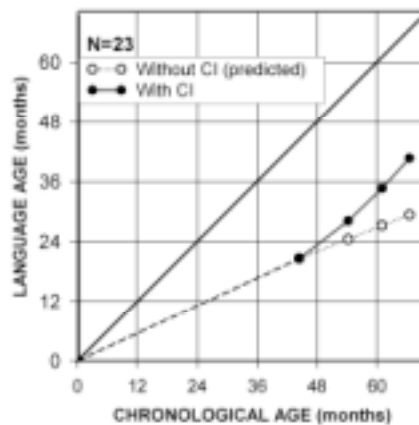


Figure 7. Expressive language skills (expressed in terms of language age) as a function of chronological age for 23 prelingually deaf cochlear implant users before implantation and at three intervals after implantation (black circles). White circles represent the expressive language growth expected from these children if they had not received cochlear implants. The solid diagonal line represents the scores expected from children with normal hearing.

Summary

Cochlear implants have been very successful in restoring partial hearing to profoundly deaf people. Most individuals with implants are now able to communicate and understand speech without lip-reading and many are able to talk over the telephone. The greatest benefits with cochlear implantation have occurred in patients who (1) acquired speech and language before their hearing loss and (2) have a shorter duration of deafness. Implant patients in United States can choose from three implant processors that support a variety of speech processing strategies. Clinicians now have the option to program the patient with multiple strategies and have the patient select the strategy they prefer. In addition, clinicians have the capability, thanks to the flexible fitting software, to change certain speech processing parameters to optimize performance.

References

- Armstrong-Bendall G, Goodrum-Clarke K, Stollwerck L, Nunn T, Wei J, & Boyle P. (1999). Clarion paired pulsatile sampler (PPS): user preference and performance. *1999 Conference on Implantable Auditory Prostheses*, Pacific Grove, California.
- Battmer R., Zilberman Y, Haake P, & Lenarz T. (1999). SAS-CIS pilot comparison study in Europe. *Annals of Otolaryngology, Rhinology & Laryngology* 108:69-73.
- Brill S, Gstottner W, Helms J, Ilberg CV, Baumgartner W, Muller J, & Kiefer J. (1997). Optimization of channel number and stimulation rate for the fast CIS strategy in the COMBI 40+. *American Journal of Otolaryngology* 18: S104-6.
- Blamey P, Arndt P, Bergeron F, Bredberg G, Brimacombe J, Facer G, Larky J, Lindstrom B, Nedzelski J, Peterson A, Shipp D, Staller S, & Whitford L. (1996). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiology and Neuro-Otology*. 1:293-306.
- Cohen NL, Roland JT, & Fishman A. (in press). Surgical technique for the Nucleus Contour cochlear implant. *Ear & Hearing*.
- Hinojosa R, & Marion M. (1983). Histopathology of profound sensorineural deafness. *Annals of New York Academy of Sciences*. 405: 459-84.
- Holden LK, Skinner MW, & Holden TA. (in press). Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear and Hearing*.
- Kessler D. (1999). The Clarion multi-strategy cochlear implant. *Annals of Otolaryngology, Rhinology and Laryngology*. Suppl. 177. 108: 8-16.
- Loizou P. (1998). Mimicking the human ear: an overview of signal processing techniques used for cochlear implants. *IEEE Signal Processing Magazine*. 15: 101-30.
- Loizou P, Poroy O, & Dorman M.. (2000). The effect of parametric variations of cochlear implant processors on speech understanding. *Journal of the Acoustical Society of America*. 108: 790-802.
- Merzenich M & White M. (1977). Cochlear implant - the interface problem. In: Hambrecht F and Reswick J, (eds.) *Functional electrical stimulation: applications in neural prostheses*. Pp. 321-340. Marcel Dekker.
- Meyer TA, Svirsky MA, Kirk KI, & Miyamoto RT. (1998). Improvements in speech perception by children with profound prelingual hearing loss: effects of device, communication mode, and chronological age. *Journal of Speech, Language and Hearing Research*. 41: 846-58.
- Meyer TA & Svirsky MA. (2000). Speech perception by children with the Clarion (CIs) or Nucleus 22 (SPEAK) cochlear implant or hearing aids. *Annals of Otolaryngology, Rhinology, & Laryngology*. Suppl.185: 49-51.
- Nilsson M, Soli SD, & Sullivan J. (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*. 95:1085-99.
- Osberger M & Fisher L. (1999). SAS-CIS preference study in postlingually deafen adults implanted with the Clarion cochlear implant. *Annals of Otolaryngology, Rhinology & Laryngology*. Suppl. 177. 108: 74-79.
- Parkinson AJ, Arcaroli J, Staller SJ, Arndt PL, Cosgriff A, & Ebinger K. (in press). The Nucleus 24 Contour cochlear implant system: adult clinical trial results. *Ear & Hearing*.
- Plant K, Whitford L, Psarros C, Vandali A, & Clark G. (1999). Parameter selection and programming recommendations for the ACE and CIS speech processing strategies. *1999 Conference on Implantable Auditory Prosthesis*, Pacific Grove, CA.
- Saunders E, Cohen L, Aschendorff A, Shapiro W, Knight M, Stecker M, Righter B, Waltzman S, Tykocinski M, Roland T, Laszig R, & Cowan R. (in press). Threshold, comfortable level and impedance changes as a function of electrode-modiolar distance. *Ear & Hearing*.
- Svirsky MA, & Meyer A. (1999). A comparison of speech perception for pediatric Clarion and hearing aid users. *Annals of Otolaryngology, Rhinology, and Laryngology*. Suppl. 177. 108: 104-10.
- Svirsky MA, Robbins AM, Kirk KI, Pisoni DB, & Miyamoto RT. (2000a). Language development in profoundly deaf children with cochlear implants. *Psychological Science*. 11: 153-58.

- Svirsky MA, Sloan RB, Caldwell M, & Miyamoto RT. (2000b). Speech intelligibility of prelingually deaf children with multichannel cochlear implants. *Annals of Otolaryngology, Rhinology, & Laryngology*. Suppl. 185: 123-5.
- van den Honert, & Stypulkowski P. (1987). Single fiber mapping of spatial excitation patterns in the electrically stimulated nerve. *Hearing Research*. 29: 195-206.
- Vandali AE, Whitford LA, Plant KL, & Clark GM. (2000). Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear & Hearing*. 21:608-24.
- Waltzman SB, Cohen NL, Gomolin RH, Green JE, Shapiro WH, Hoffman RA, & Roland JT, Jr. (1997). Open-set speech perception in congenitally deaf children using cochlear implants. *American Journal of Otolaryngology*. 18:342-9.
- Wilson B. (1993). Signal processing. In: Tyler R, (ed.) *Cochlear Implants: Audiological Foundations*. Pp. 35-86. Singular Publishing Group, Inc.
- Wilson B, Finley C, Lawson D, Wolford R, Eddington D, & Rabinowitz W. (1991). Better speech recognition with cochlear implants. *Nature*. 352: 236-38.
- Wilson B, Lawson D, & Zerbi M. (1993). Speech Processors for auditory prostheses. NIH Project N01-DC-2-2401, *Fifth Quarterly Progress Report*.
- Wilson B, Lawson D, & Zerbi M. (1995). Advances in coding strategies for cochlear implants. *Advances in Otolaryngology - Head and Neck Surgery*, 9: 105-29.
- Ziese M, Stutzel A, von Specht H, Begall K, Freigang B, Sroka S, & Nopp P. (2000). Speech understanding with the CIS and N-of-M strategy in the MED-EL COMBI 40+ system. *ORL*. 62: 321-29.
- Zwolan TA, Loughheed LA, & Kileny PR (1994). Clinical results of prelingually deafened adult cochlear implant patients. *Abstracts of the Seventeenth Midwinter Research Meeting, Association for Research in Otolaryngology*, February 6-10, 1994, St. Petersburg Beach, FL.