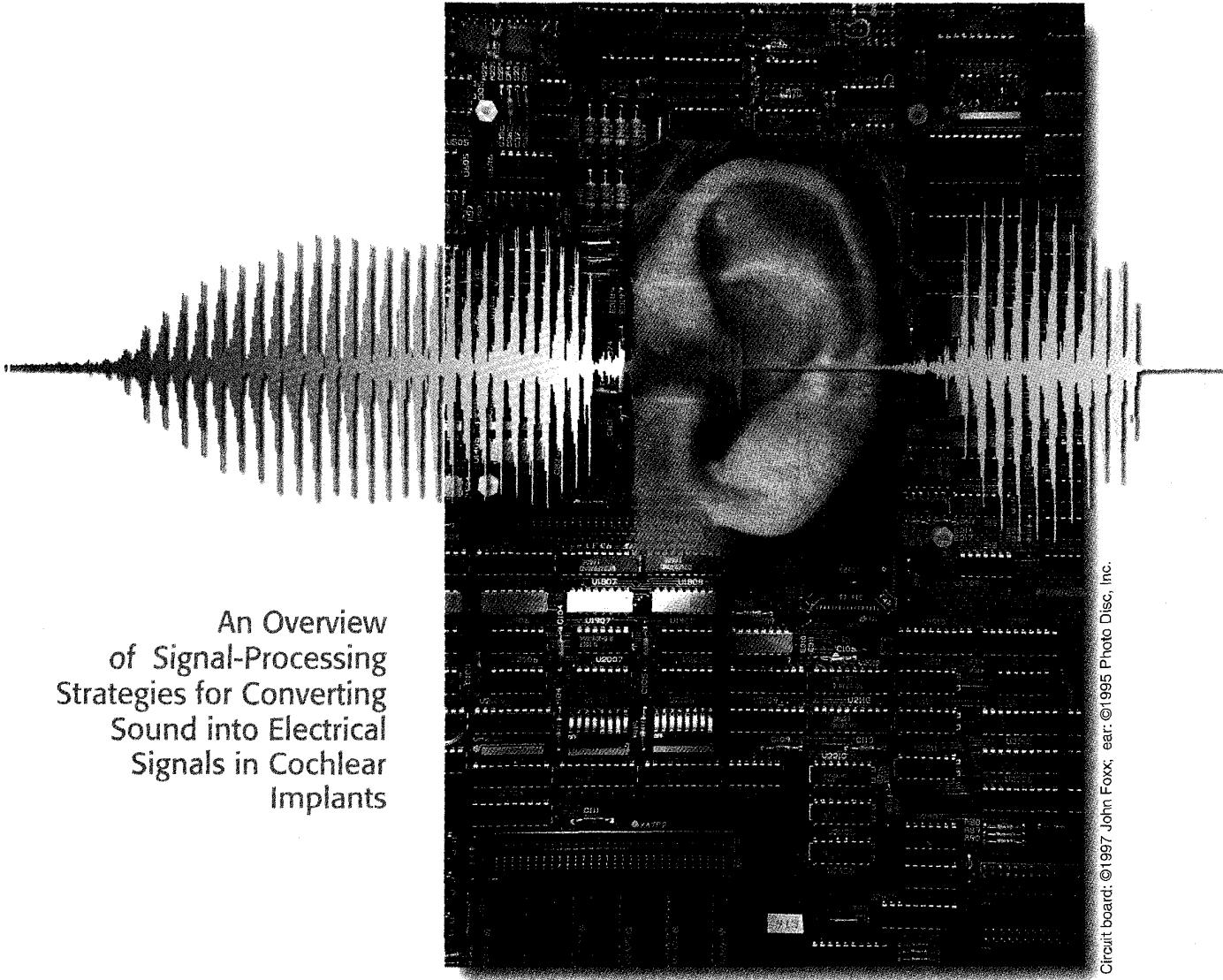


# Mimicking the **HUMAN EAR**



Circuit board: © 1997 John Foxx; ear: © 1995 Photo Disc, Inc.

An Overview  
of Signal-Processing  
Strategies for Converting  
Sound into Electrical  
Signals in Cochlear  
Implants

**F**or centuries, people believed that only a miracle could restore hearing to the deaf. It was not until 40 years ago that scientists first attempted to restore normal hearing to the deaf by electrical stimulation of the auditory nerve. The first experiments were discouraging as the patients reported that speech was unintelligible. However, as researchers kept investigating different techniques for delivering electrical stimuli to the auditory nerve, the auditory sensations elicited by electrical stimulation gradually came closer to sounding more like normal speech. Today, a prosthetic device, called a cochlear implant, can be implanted in the inner ear and can restore partial hearing to profoundly deaf people. Some individuals with implants can now communicate without lip-reading or signing, and some can communicate over the telephone.

The success of cochlear implants can be attributed to the combined efforts of scientists from various disciplines including bioengineering, physiology, otolaryngology, speech science, and signal processing. Each of these disciplines contributed to various aspects of the design of cochlear prostheses. Signal processing, in particular, played an important role in the development of different techniques for deriving electrical stimuli from the speech signal. Designers of cochlear prosthesis were faced with the challenge of developing signal-processing techniques that would mimic the function of a normal cochlea.

The purpose of this article is to present an overview of various signal-processing techniques that have been used for cochlear prosthesis over the past 25 years. The signal-processing strategies described here will be only a subset of the many that have been developed for cochlear prosthesis. Information on other signal processing strategies may be found in the excellent review by Wilson [1] as well as in [2], [3], and [4].

## Background

An understanding of how our auditory system works is central to the development of a successful cochlear implant. After all, it is important to know how a normal auditory system functions before we can fix a system that is damaged. Here we review background material necessary for understanding cochlear prosthesis including a brief overview of the type of information contained in the acoustic signal and its relation to speech perception and a description of how the auditory system works, how it analyzes the speech signal, and what causes hearing loss.

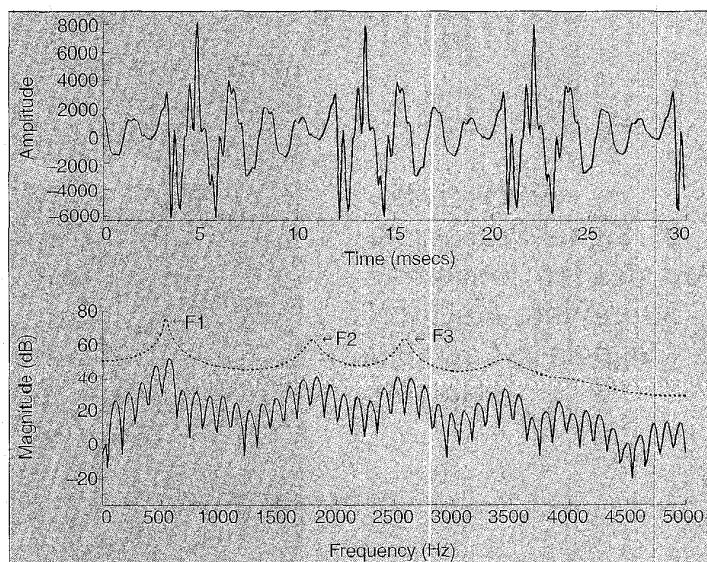
## The Speech Signal

The designers of cochlear prosthesis need to know what information in the speech signal is perceptually important. This information needs to be preserved in order for the patient to be able to hear

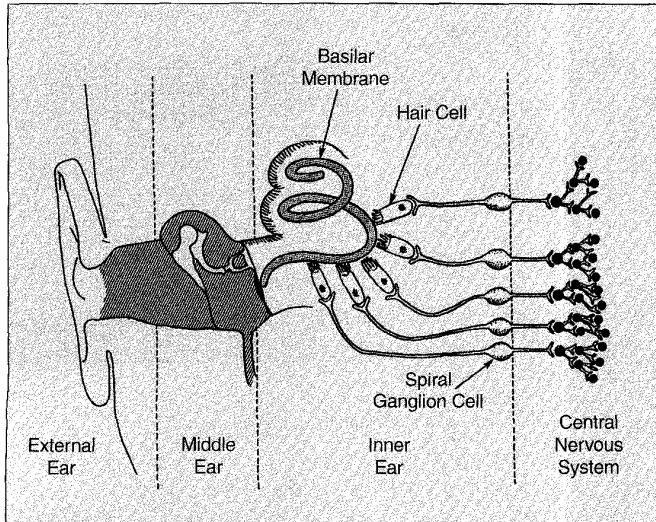
speech that is intelligible. This section reviews some fundamental elements of the speech signal.

According to the source-filter model of speech production [5, 6], the speech signal can be considered to be the output of a linear system. Depending on the type of input excitation (source), two classes of speech sounds are produced: voiced and unvoiced. If the input excitation is noise, then unvoiced sounds such as /s/, /t/, etc., are produced, and if the input excitation is periodic then voiced sounds such as /a/, /i/, etc., are produced. In the unvoiced case, noise is generated either by forcing air through a narrow constriction (e.g., production of /f/) or by building air pressure behind an obstruction and then suddenly releasing that pressure (e.g., production of /t/). In contrast, the excitation used to produce voiced sounds is periodic and is generated by the vibrating vocal cords. The frequency of the voiced excitation is commonly referred to as the fundamental frequency ( $F_0$ ).

The vocal tract shape, defined in terms of tongue, velum, lip and jaw position, acts like a "filter" that filters the excitation to produce the speech signal. The frequency response of the filter has different spectral characteristics depending on the shape of the vocal tract. The broad spectral peaks in the spectrum are the resonances of the vocal tract and are commonly referred to as formants. Figure 1 shows, for example, the formants of the vowel /eh/ (as in "head"). The frequencies of the first three formants (denoted as  $F_1$ ,  $F_2$ , and  $F_3$ ) contain sufficient information for the recognition of vowels as well as other voiced sounds. Formant movements have also been found to be extremely important for the perception of unvoiced sounds (i.e., consonants) (e.g., [6, 7]). In summary, the formants carry some information about the speech signal; because of this, some of the early cochlear implant devices conveyed formant information to the electrodes.



▲ 1. The top panel shows the time waveform of a 30-msec segment of the vowel /eh/, as in "head." The bottom panel shows the spectrum of the vowel /eh/ obtained using the short-time Fourier transform (solid lines) and linear prediction (LPC) analysis (dashed lines). The peaks in the LPC spectrum correspond to the formants  $F_1$ ,  $F_2$ , and  $F_3$ .



▲ 2. A diagram (not to scale) of the human ear (reprinted with permission from [85]).

### Normal Hearing

Figure 2 shows a simplified diagram of the human ear consisting of the outer, middle, and inner ear [8]. 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 that 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 fluids of the cochlea lead to displacements of a flexible membrane, called the basilar membrane, in the cochlea. These displacements contain information about the frequency of the acoustic signal. Attached to the basilar membrane are hair cells that are bent according to the displacements of the basilar membrane. The bending of the hairs releases an electrochemical 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.

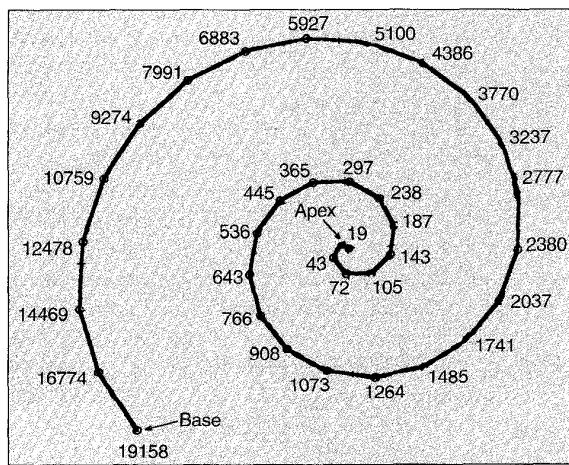
### Deafness

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, and that in turn leads to hearing impairment. Put simply, the sound travels through the outer ear, the middle ear, and the inner ear but never makes it all the way to the brain because of the broken link—the damaged hair cells. The hair cells can be damaged by certain diseases (e.g., meningitis, Meniere's disease), congenital disorders, by certain drug

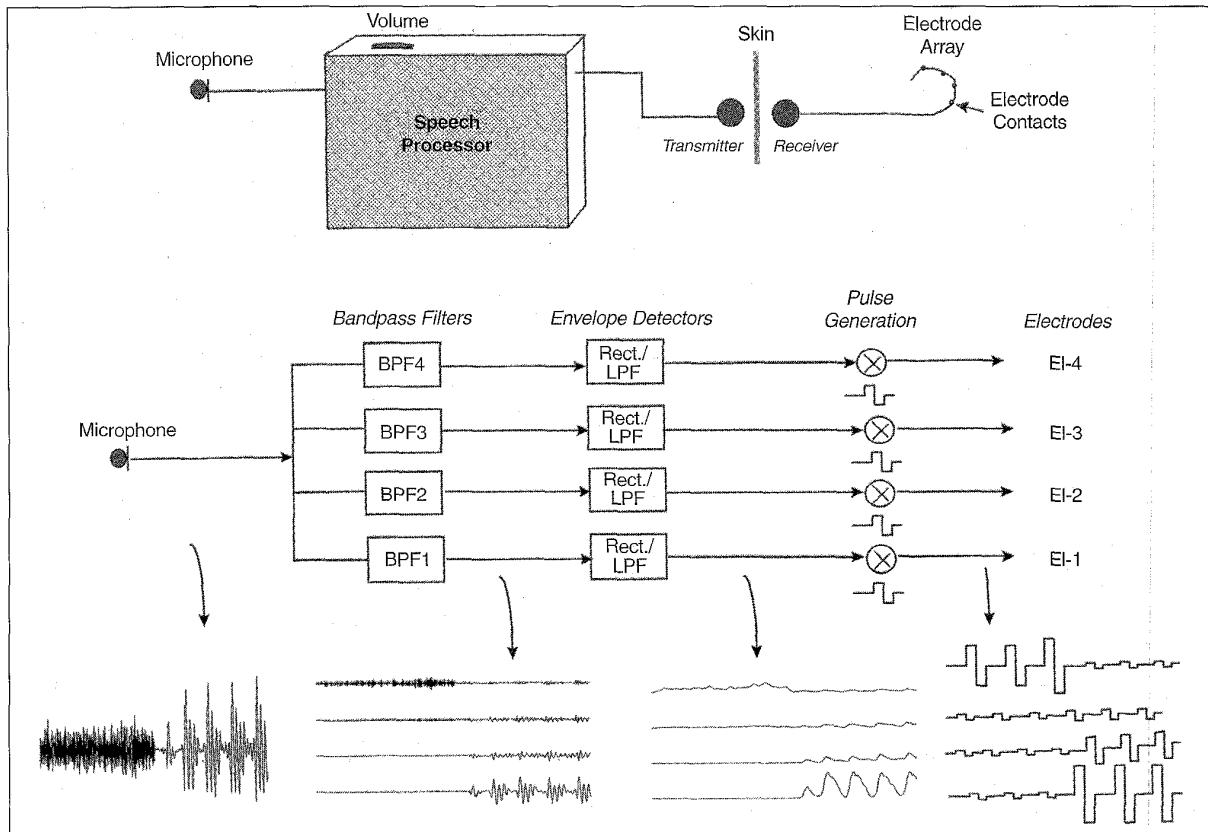
treatments, or by many other causes. Damaged hair cells can subsequently lead to degeneration of adjacent auditory neurons, and if a large number of hair cells or auditory neurons throughout the cochlea are damaged, then the person with such a loss is diagnosed as profoundly deaf. Research [9] has shown 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 implants because the remaining neurons could be excited directly through electrical stimulation. A cochlear prosthesis is therefore based on the idea of bypassing the normal hearing mechanism (outer, middle, and part of the inner ear including the hair cells) and electrically stimulating the remaining auditory neurons directly. The challenge we face is finding out how to stimulate (electrically) auditory neurons so that meaningful information about speech is conveyed to the brain. The electrical stimulation should, for example, convey information about the amplitude and the frequency of the acoustic signal.

### Encoding Frequency

This leads us to the question "How does the auditory system encode frequencies?" The pioneering work of Georg von Bekesy in the 1950s 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. 3). Low-frequency sounds create traveling waves in the fluids of the cochlea that cause the basilar membrane to vibrate with largest amplitude of displacement at the apex (see Fig. 3) of the basilar membrane. On the other hand, high-frequency sounds create traveling waves with largest amplitude of displacement at the base (near the stapes) of



▲ 3. 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.



▲ 4. 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" as an input signal. 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.

the basilar membrane. 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 therefore acts like a spectrum analyzer. It decomposes complex sounds into their frequency components.

The cochlea is one of the mechanisms used by our auditory system for encoding frequencies. The traveling wave of the basilar membrane in the cochlea vibrates with maximum amplitude at a place along the cochlea that is dependent on the frequency of stimulation. 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 theory*. The place mechanism for coding frequencies has motivated multichannel cochlear implants. Another theory, called *volley theory*, suggests that frequency is determined by the rate at which the neurons are fired. According to the volley theory, the audi-

tory nerve fibers fire at rates proportional to the period of the input signal up to frequencies of 5,000 Hz. At low frequencies, individual nerve fibers fire at each cycle of the stimulus; i.e., they are "phase locked" with the stimulus. At high frequencies, frequency is indicated by the organized firing of groups of nerve fibers.

## Cochlear Implants

Several cochlear implant devices have been developed over the years [1]. All 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 inserted into the cochlea by a surgeon (Fig. 4). In single-channel implants only one electrode is used. In multichannel cochlear implants, an electrode array is inserted in the cochlea so that different auditory nerve fibers can be stimulated at different places in the cochlea,

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 analyzes the input signal into its frequency components. The designers of cochlear prosthesis are faced with the challenge of developing signal-processing techniques that mimic the function of a healthy cochlea.

The cochlear implant is based on the idea that there are enough auditory nerve fibers left 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 them 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. Likewise, if a small number of nerve fibers is activated, then the sound is perceived as soft. 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 when electrodes near the base are stimulated. In summary, the implant can effectively transmit information to the brain about the loudness of the sound, which is a function of the amplitude of the stimulus current, and the pitch, which is a function of the place in the cochlea being stimulated.

As an example, Fig. 4 shows the operation of a four-channel implant. Sound is picked up by a microphone and sent to a speech processor box (the size of a pager) worn by the patient. The sound is then processed through a set of four bandpass filters that divide the acoustic waveform into four channels. Current pulses are generated with amplitudes proportional to the energy of 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. 4). For instance, if the speech signal contains mostly high frequency information (e.g., /s/), then the pulse amplitude of the fourth channel 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 the first and second channels will be large relative to the amplitudes of channels 3 and 4 (Fig. 4). The electrodes are therefore stimulated according to the energy level of each frequency channel.

## The cochlea is one of the mechanisms used by our auditory system for encoding frequencies.

### **Implant Characteristics**

Figure 4 showed one type of cochlear implant that is being used. Several other types of implant devices have been developed over the years [1]. These devices differ in the following characteristics:

- ▲ Electrode design (e.g., number of electrodes, electrode configuration),
- ▲ Type of stimulation—analog or pulsatile,
- ▲ Transmission link—transcutaneous or percutaneous,
- ▲ Signal processing—waveform representation or feature extraction.

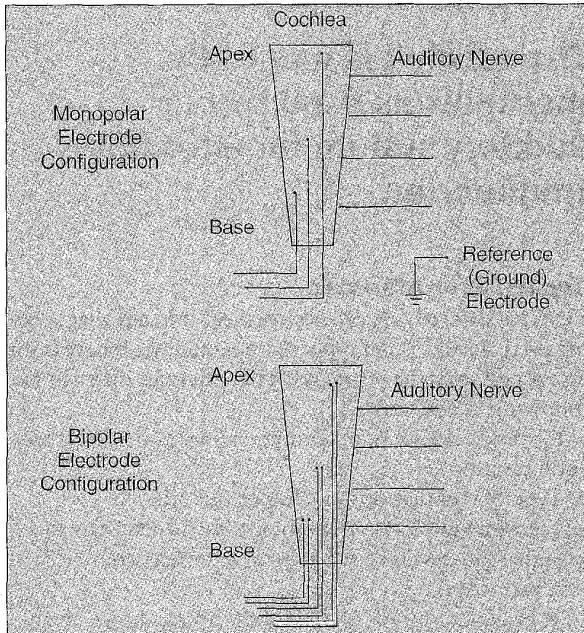
A brief description of each of the above device characteristics is given below.

### **Electrode Design**

The design of electrodes for cochlear prosthesis has been the focus of research for over two decades [10, 11]. Some of the issues associated with electrode design are: (1) electrode placement, (2) number of electrodes and spacing of contacts, (3) orientation of electrodes with respect to the excitable tissue, and (4) electrode configuration.

Electrodes may be placed near the round window of the cochlea (extracochlear), or in the scala tympani (intracochlear) or on the surface of the cochlear nucleus. Most commonly, the electrodes are placed in the scala tympani because it brings the electrodes 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 used in a 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 the electrodes affects the place resolution for coding frequencies. In principle, the larger the number of electrodes, the finer the place resolution for coding frequencies. Frequency coding is constrained, however, by two factors that are inherent in the design of cochlear prosthesis: (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, there is not much that can be done about the first problem, because it depends on the etiology of deafness. Ideally, we would like to have sur-



▲ 5. Diagram showing two electrode configurations: monopolar and bipolar. In the monopolar configuration the active electrodes are located far from the reference electrode (ground), while in the bipolar configuration the active and reference electrodes are placed close to each other.

viving auditory neurons lying along the length of the cochlea. Such a neuron survival pattern would support a 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 all 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. When electric current is injected to the cochlea, it tends to spread out symmetrically from the source. As a result, the current stimulus does not stimulate just a single (isolated)

site of auditory neurons, but several. Such a spread in excitation is most prominent in monopolar electrode configuration. In this configuration, the active electrode is located far from the reference electrode, which acts as a ground for all electrodes (see Fig. 5). The spread of excitation due to electrical stimulation can be constrained to a degree by using a bipolar electrode configuration. In the bipolar configuration, the active and the reference (ground) electrodes are placed close to each other (Fig. 5). Bipolar electrodes have been shown to produce a more localized stimulation than monopolar electrodes [12, 13]. Although the patterns of electrical stimulation produced by monopolar and bipolar configurations are different, it is still not clear which of the two electrode configurations will result in better performance for a particular patient.

Currently, some implant devices employ monopolar electrodes, other devices employ bipolar electrodes, and other devices provide both types of electrodes. Table 1 shows a list of current implant devices and their characteristics. The Ineraid (also called Symbion) device uses 6 electrodes spaced 4 mm apart. Only the four most apical electrodes are used in monopolar configuration. 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. Eight electrodes are used, which are spaced 2 mm apart. The Med-El device uses 12 electrodes (spaced 2.4 mm apart) in monopolar configuration.

#### Type of Stimulation

There are generally two types of stimulation depending on how information is presented to the electrodes. If the information is presented in analog form, then the stimulation is referred to as analog stimulation, and if the information is presented in pulses, then the stimulation is referred to as pulsatile stimulation.

In analog stimulation, an electrical analog of the acoustic waveform itself 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 in analog form. 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

Table 1: Characteristics of commercially available cochlear implant devices.

Device	Electrodes			Type of stimulation	Transmission link
	Number	Spacing	Configuration		
Ineraid	6	4 mm	Monopolar	Analog	Percutaneous
Nucleus	22	0.75 mm	Bipolar	Pulsatile	Transcutaneous
Clarion 1.0	8	2 mm	Monopolar/Bipolar	Analog/Pulsatile	Transcutaneous
Med-El	12	2.4 mm	Monopolar	Pulsatile	Transcutaneous

of analog stimulation is that the simultaneous stimulation may cause channel interactions.

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. 4). The advantage of this type of stimulation is that the pulses can be delivered in a nonoverlapping (i.e., nonsimultaneous) fashion, thereby minimizing channel interactions. The rate at which these pulses are delivered to the electrodes, i.e., the pulse rate, has been found to affect speech-recognition performance [14]. High pulse rates tend to yield better performance than low pulse rates.

#### *Transmission Link*

Once the electrodes are in place, how do signals get transmitted from the external processor to the implanted electrodes? There are currently two ways of transmitting the signals: (1) through a transcutaneous connection and (2) through a percutaneous connection (see Fig. 6).

The transcutaneous system transmits the stimuli through a radio-frequency link. In this system, an external transmitter is used to encode 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 (Fig. 6). The transmitter and the implanted receiver are held in place on the scalp by a magnet. The advantage of this system is that the skin in the scalp is closed after the operation, thus avoiding possible infection. The disadvantage of this system is that the implanted electronics (i.e., the receiver circuitry) may fail, and it would require surgery to replace them. Another disadvantage of this system is that the transcutaneous connector contains magnetic materials that are incompatible with MRI scanners. Most cochlear implant devices (e.g., Nucleus, Clarion, Med-El) today use transcutaneous connections.

The percutaneous system transmits the stimuli to the electrodes directly through plug connections (Fig. 6). In this system, there are no implanted electronics, other than the electrodes. The major advantage of the percutaneous

## **The cochlear implant is based on the idea that there are enough auditory nerve fibers left for stimulation in the vicinity of the electrodes.**

system is flexibility and signal transparency. The signal transmission is in no way constrained by the implanted receiver circuitry. It is therefore ideal for research purposes for investigating new signal-processing techniques. The Ineraid device is the only device that uses percutaneous connectors.

#### *Signal Processing*

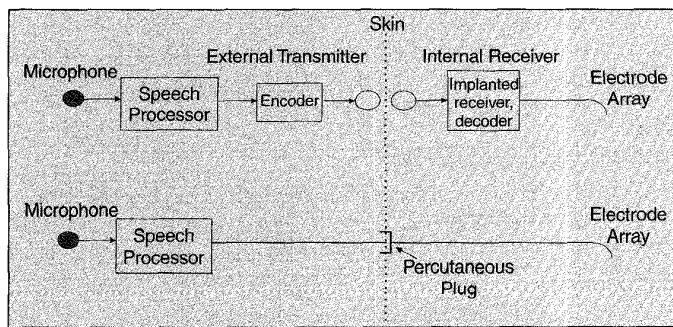
The last, and perhaps most important, difference among implant devices is in the signal-processing strategy used for transforming the speech signal to electrical stimuli. Several signal-processing techniques have been developed over the past 25 years. Some of these techniques are aimed at preserving waveform information, others are aimed at preserving envelope information, and others are aimed at preserving spectral features (e.g., formants). A more detailed description of each of these signal-processing techniques will be presented in the following sections. Representative results for each signal-processing strategy will be presented.

#### **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 has to be bilateral (i.e., in both ears). Profound deafness [15] is defined as a hearing loss of 90 dB or more. Hearing loss is typically measured as the average of pure-tone hearing thresholds at 500, 1,000, and 2,000 Hz expressed in dB with reference to normal thresholds. Second, the candidate has to obtain sentence recognition scores of 30% correct or less under best-aided conditions. Children age two years or older with profound ( $> 90$  dB HL) sensorineural loss in both ears are also candidates for cochlear implantation.

#### **Evaluating Performance**

Once a patient has been fit with a cochlear implant, how do we evaluate his/her ability to identify or recognize speech? Patient's speech perception abilities are typically evaluated using sentence, monosyllabic word, vowel, and consonant tests. Implant patients tend to achieve higher scores on sentence tests than on any other test. This is because they can use higher-level knowledge such as grammar, context, semantics, etc., when recognizing words in sentences. For



▲ 6. Diagram showing two different ways of transmitting electrical stimuli to the electrode array. The top panel shows a transcutaneous (radio-frequency link) connection and the bottom panel shows a percutaneous (direct) connection.

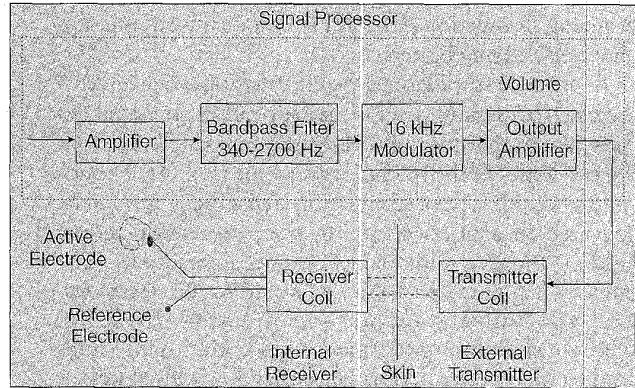
example, a patient might only hear the first two words and the final word in a sentence but can use context to "fill in" the blanks. Sentence tests are considered to be open sets because the patient does not know the list of all possible word choices. Tests of vowel and consonant recognition, on the other hand, are considered closed-set tests. In these tests the patient knows all of the possible choices, but the tests themselves are not necessarily easier because all the items in the list are phonetically similar. In a vowel test, for example, the patient may listen to words like "heed, had, hod, head, hud, hid, hood, who'd" which only differ in the middle segment (i.e., the vowel) of the word. Vowel and consonant tests are aimed at assessing a patient's ability to resolve spectral and temporal information. The most difficult test, by far, is the recognition of mono-syllabic words. One such test, the NU-6 word lists, was developed by Northwestern University and consists of lists of 50 mono-syllable words [16]. Other standardized tests include the recognition of 100 keywords from the Central Institute for the Deaf (CID) sentences of everyday speech, recognition of 25 two-syllable words (spondees), and the Iowa test [17] which consists of sentences, vowels, and consonants recorded on a laserdisc in audio, visual, and audio-visual format.

Different tests are used to evaluate the speech perception abilities of children. These tests are specially designed to reflect the language and vocabulary level of the child. It makes no sense, for example, to include the word or picture of a "turtle" in the test, if the child does not know what a turtle is. A good review on various tests developed to evaluate the speech perception abilities of children can be found in [18].

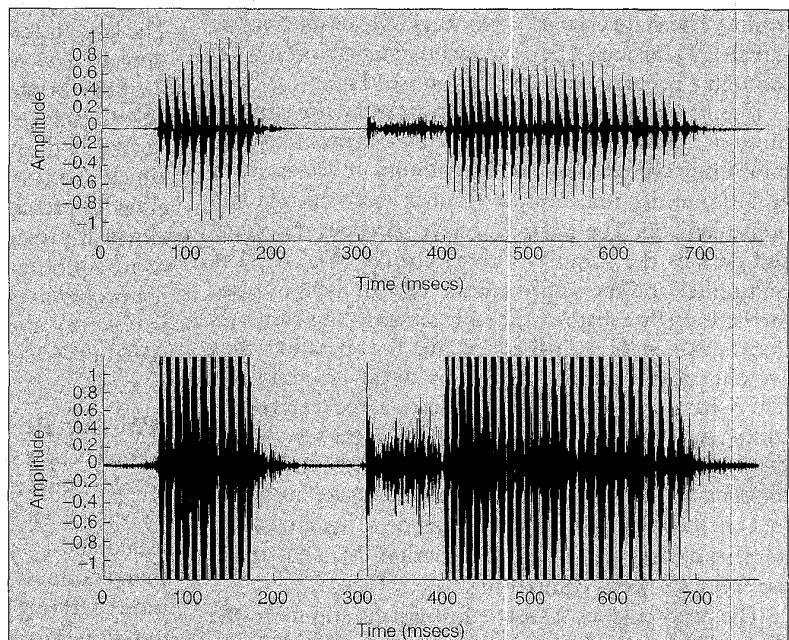
## Single-Channel Implants

Single-channel implants provide electrical stimulation at a single site in the cochlea using a single electrode. These implants are of interest because of their simplicity in design and their low cost compared to multichannel implants. They are also appealing because they do not require much hardware and conceivably all the electronics could be packaged into a behind-the-ear device.

Single-channel implants were first implanted in human subjects in the early 1970s. At the time, there was a lot of skepticism about whether single-channel stimulation could really work [19]. Doctors and scientists argued that electrical stimulation of the auditory nerve could produce nothing but noise. Despite the controversy, researchers in United States and in Europe kept working on the develop-



▲ 7. Block diagram of the House/3M single-channel implant. The signal is processed through a 340-2,700 Hz filter, modulated with a 16 kHz carrier signal, and then transmitted (without any demodulation) to a single electrode implanted in the scala tympani.

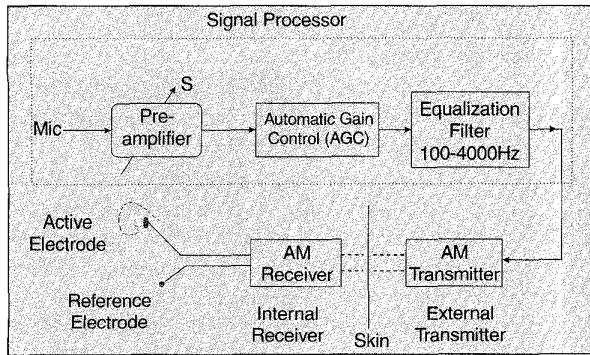


▲ 8. The time waveform (top) of the word "aka," and the amplitude modulated waveform (bottom) processed through the House/3M implant for input signal levels exceeding 70 dB SPL.

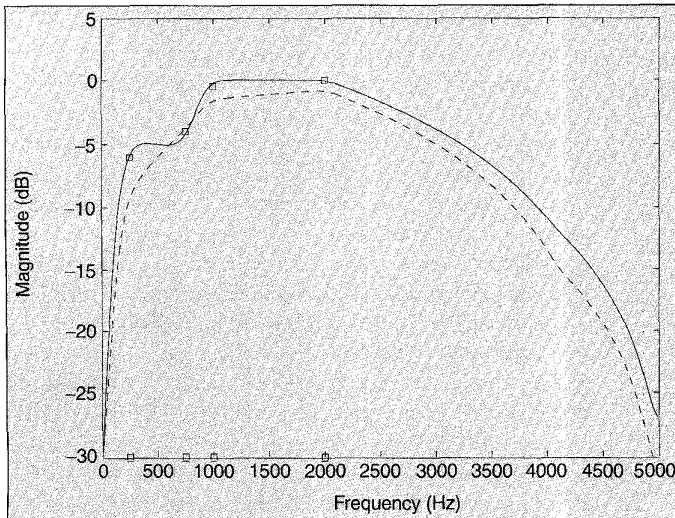
ment of single-channel prosthesis. These early efforts led to, among other devices, the House/3M single-channel implant and the Vienna/3M single-channel implant.

### House/3M Device

The House single-channel implant was originally developed by William House and his associates in the early 1970s [20, 21]. Improvements to the implant were later undertaken jointly with the 3M company, and the device was henceforth referred to as House/3M. Figure 7 shows the block diagram of the House/3M device. It consists of three main components: (1) the signal processor, (2) the signal transmitter/receiver, and (3) the implanted electrodes [22]. The acoustic signal is picked up by a micro-



▲ 9. Block diagram of the Vienna/3M single-channel implant. The signal is first processed through a gain-controlled amplifier that compresses the signal to the patient's electrical dynamic range. The compressed signal is then fed through an equalization filter (100-4,000 Hz), and is amplitude modulated for transcutaneous transmission. The implanted receiver demodulates the radio-frequency signal and delivers it to the implanted electrode.



▲ 10. The equalization filter used in the Vienna/3M single-channel implant. The solid plot shows the ideal frequency response and the dashed plot shows the actual frequency response. The squares indicate the corner frequencies, which are adjusted for each patient for best equalization.

phone, amplified, and then processed through a 340-2,700 Hz bandpass filter. The bandpassed signal is then used to modulate a 16 kHz carrier signal. The modulated signal goes through an output amplifier and is applied to an external induction coil. The output amplifier allows the patient to control the intensity of the stimulation. The output of the implanted coil is finally sent (without any demodulation) to the implanted active electrode in the scala tympani.

In the House/3M device, it is the modulated speech signal that is being transmitted to the electrodes rather than the speech signal itself. Information about gross temporal fluctuations of the speech signal are contained in the envelope of the modulated signal. However, the shape of the modulated envelope signal is affected by the

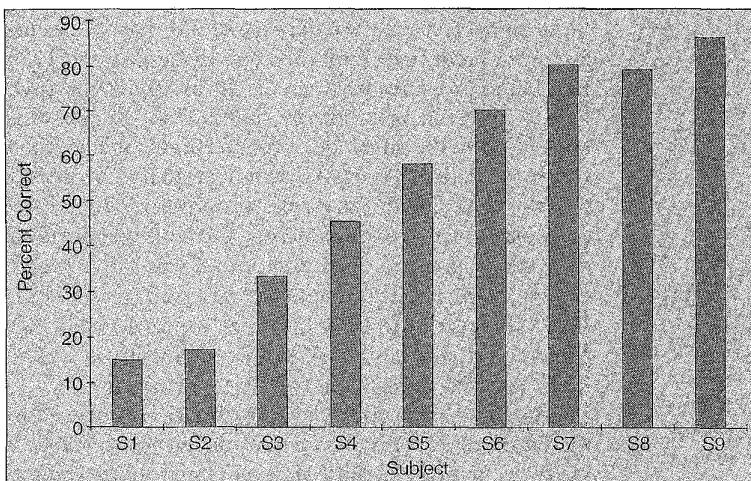
input signal level because the House/3M device does not attempt to reduce or limit the input dynamic range [22]. For sound pressures between 55 dB to 70 dB, the envelope output changes linearly, but for sound pressures above 70 dB, the envelope output saturates at a level just below the patient's level of discomfort; that is, for speech signals above 70 dB the envelope output is clipped (see example in Fig. 8). Consequently, the temporal details in the speech signal may be distorted or discarded. The periodicity, however, of the signal is preserved. As shown in Fig. 8, bursts of the 16 kHz carrier appear to be in synchrony with the period of voiced segments as well as other low-energy segments of the input signal.

Given the limited temporal information conveyed by the House/3M device, it was not surprising that the majority of the patients did not obtain open-set speech recognition with hearing alone (e.g., [23]). Rosen et al. [24] found that the average percent correct score on consonant identification was 37% for four patients. Only exceptional patients were able to obtain scores above zero on monosyllabic word (NU-6) identification. In a study by Danhauer et al. [25], only four patients (out of 18) achieved a 2% correct score and only one patient achieved a 4% correct score on monosyllabic word identification.

### Vienna/3M Device

The Vienna single-channel implant was developed at the Technical University of Vienna, Austria, in the early 1980s [26]. The block diagram of the Vienna/3M implant is shown in Fig. 9. The signal is first preamplified and then compressed using a gain-controlled amplifier with a short attack time (0.5 msec). The amount of compression is adjusted according to the patient's dynamic range. The compressed signal is then fed through a frequency-equalization filter (Fig. 10) that attenuates frequencies outside the range 100-4,000 Hz. The filtered signal is amplitude modulated for transcutaneous transmission. The implanted receiver demodulates the radio-frequency signal and sends the demodulated stimuli to the implanted electrode.

The Vienna/3M device was designed so that: (1) the temporal details in the analog waveform would be preserved and (2) frequencies in the range 100-4,000 Hz would be audible to the patients. The automatic gain control ensures that the temporal details in the analog waveform are preserved regardless of the input signal level. It therefore prevents high-level input signals from being clipped. The frequency-equalization filter ensures that all frequencies in the range of 100 Hz to 4 kHz, which are very important for understanding speech, are audible to the patients. Without the equalization filter, only low-frequency signals would be audible. This is because the electrical threshold level (i.e., the electrical stimulus level that is barely audible to the patient) is typically lower



▲ 11. Percentage of words identified correctly on sentence tests by nine "better-performing" patients wearing the Vienna/3M device (Tyler et al. [29]).

at low frequencies and higher at high frequencies ( $> 300$  Hz) [27][28]. The frequency response of the equalization filter (Fig. 10) is adjusted for each patient so that sinusoids with frequencies in the range of 100 Hz to 4 kHz are equally loud.

Unlike the House/3M device, the Vienna/3M device managed to preserve fine temporal variations in the speech signal. Some Vienna/3M patients were able to recognize speech. In the study by Tyler [29], some of the exceptional patients were able to identify words in sentences with 86% accuracy. Word identification scores ranged from 15% to 86% correct across nine patients (Fig. 11). Hochmair-Desoyer et al. [30] also report, for a group of 22 patients, a mean score of 30% correct for monosyllabic word identification and a mean score of 45% correct for words in sentences. Unfortunately, not all patients did as well. Other researchers (e.g., Gantz et al. [23]), found that patients using the Vienna/3M device were not able to obtain scores above zero on open-set speech recognition.

### Speech Perception Using Single-Channel Implants

It was not surprising that relatively few patients could obtain open-set speech understanding with single-channel implants given the limited spectral information. Single-channel stimulation does not exploit the place code mechanism used by a normal cochlea for encoding frequencies, since only a single site in the cochlea is being stimulated. Temporal encoding of frequency by single nerve fibers is restricted (due to the neural refractory period) to 1 kHz [31]. It is also conceivable that patients could extract frequency information from the periodicity of the input stimulus. This is possible, but only for stimulus frequencies up to 300-500 Hz. Experiments [27] showed that implant patients (as well as normal-hearing listeners [32]) cannot discriminate differences in pitch for stimulus frequencies above 300 Hz. (It should be noted

that pitch is related to frequency. As the frequency increases, we hear a rise in pitch and as the frequency decreases we hear a lowering in pitch. The relationship, however, is not linear [35].)

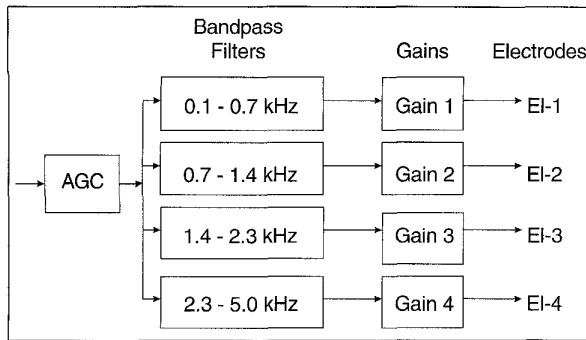
Single-channel stimulation restricts the amount of spectral information that an implant patient can receive to frequencies below 1 kHz. However, this is not sufficient for speech perception because there is important information in the speech signal up to 4,000 Hz, and beyond. But what kind of information is available in the speech signal below 1 kHz? The speech signal contains information about the fundamental frequency, the first formant (F1), and sometimes (depending on the vowel and the speaker) the second formant (F2). The presence of fundamental frequency indicates the presence of voiced sounds (e.g., vowels), and, therefore, the patient could discriminate between voiced (vowels) and unvoiced sounds (majority of consonants). Changes in fundamental frequency also give information about sentence prosody; i.e., the patients should be able to tell whether a sentence is a statement or a question. Patients could also discriminate between certain vowels that differ in F1 frequency, i.e., vowels /i, u/ and /a, ae/. Finally, assuming that the temporal details in the waveform are preserved (as in the Vienna/3M device), the patients should be able to discriminate among the consonant sets /s sh th f/, /b d g p t k/ and /w r l y/, which have different waveform characteristics [34].

In summary, single-channel implants are capable of conveying time/envelope information as well as some frequency information. The transmitted frequency information however is limited and insufficient for speech recognition. Yet, some of the exceptional patients achieved high scores on open-set speech recognition tests. It remains a puzzle how some single-channel patients can perform so well given the limited spectral information they receive.

### Multichannel Implants

Unlike single-channel implants, multichannel implants provide electrical stimulation at multiple sites in the cochlea using an array of electrodes. An electrode array is used so that different auditory nerve fibers can be stimulated at different places in the cochlea, 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.

When multichannel implants were introduced in the 1980s, several questions were raised regarding multichannel stimulation:



▲ 12. Block diagram of the compressed analog approach used in the Ineraid device. The signal is first compressed using an automatic gain control. The compressed signal is then filtered into four frequency bands (with the indicated frequencies), amplified using adjustable gain controls, and then sent directly to four intracochlear electrodes.

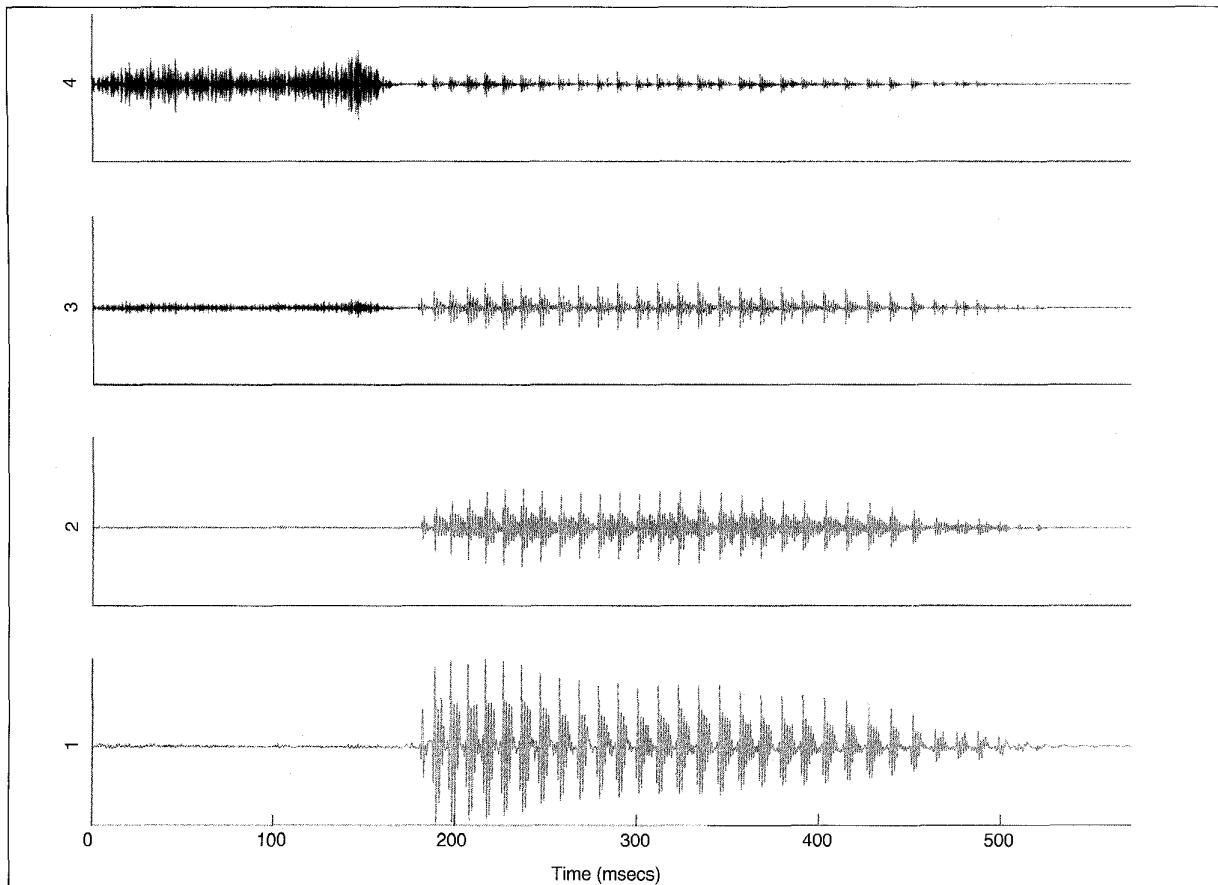
▲ How many electrodes should be used? If one channel of stimulation is not sufficient for speech perception, then how many channels are needed to obtain high levels of speech understanding?

▲ Since more than one electrode will be stimulated, what kind of information should be transmitted to each elec-

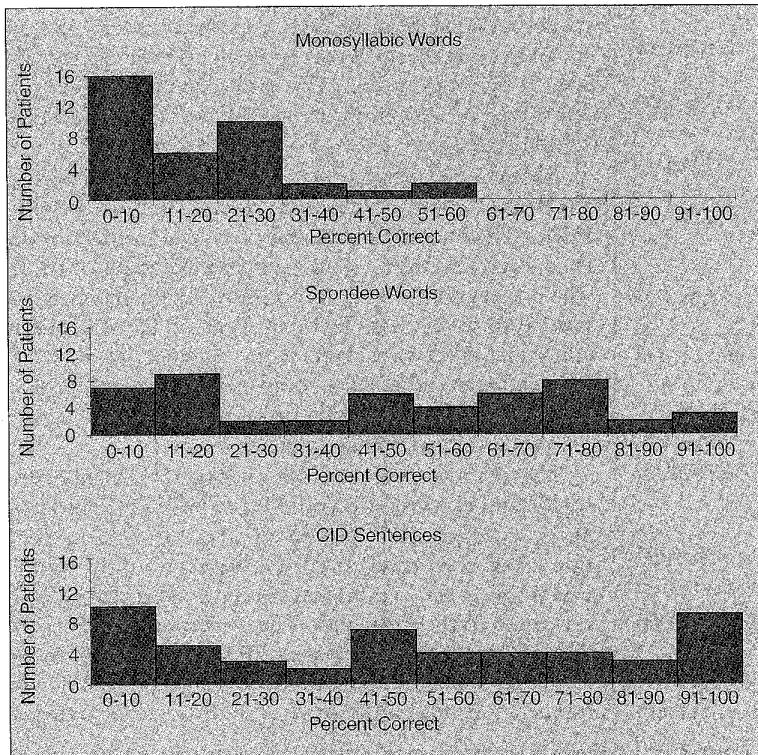
trode? Should it be some type of spectral feature or attribute of the speech signal that is known to be important for speech perception (e.g., first and second formants), or some type of waveform derived by filtering the original speech signal into several frequency bands?

Researchers experimented with different numbers of electrodes. Some devices used a large number of electrodes (22) but only stimulated a few, while other devices used a few electrodes (4-8) and stimulated all of them. The answer to the question on how many channels are needed to obtain high levels of speech understanding is still the subject of debate (e.g., Shannon et al. [35], Dorman et al. [36]). Depending on how researchers tried to address the second question, different types of signal-processing techniques were developed.

The various signal-processing strategies developed for multichannel cochlear prosthesis can be divided into two main categories: waveform strategies and feature-extraction strategies. These strategies differ in the way information is extracted from the speech signal and presented to the electrodes. The waveform strategies try to present some type of waveform (in analog or pulsatile form) derived by filtering the speech signal into different frequency bands, while the feature-extraction strategies



▲ 13. Bandpassed waveforms of the syllable "sa" produced by a simplified implementation of the compressed analog approach. The waveforms are numbered by channel, with channel 4 being the high-frequency channel (2.3-5 kHz), and channel 1 being the low frequency channel (0.1-0.7 kHz).



▲ 14. The distribution of scores for 50 Ineraid patients tested on monosyllabic word recognition, spondee word recognition, and sentence recognition (Dorman et al. [39]).

try to present some type of spectral features, such as formants, derived using feature extraction algorithms. A review of these signal-processing strategies is given next, starting with waveform strategies and continuing with feature-extraction strategies.

### Compressed-Analog (CA) Approach

The compressed-analog (CA) approach was originally used in the Ineraid device manufactured by Symbion, Inc., Utah [37]. The CA approach was also used in the UCSF/Storz device [38], which is now discontinued. The block diagram of the CA approach is shown in Fig. 12. The signal is first compressed using an automatic gain control and then filtered into four contiguous frequency bands, with center frequencies at 0.5, 1, 2, and 3.4 kHz. The filtered waveforms go through adjustable gain controls and then are sent directly through a percutaneous connection to four intracochlear electrodes. The filtered waveforms are delivered simultaneously to four electrodes in analog form. The electrodes, spaced 4 mm apart, operate in monopolar configuration. Figure 13 shows, as an example, the four bandpass filtered waveforms produced for the syllable "sa" using a simplified implementation of the CA approach.

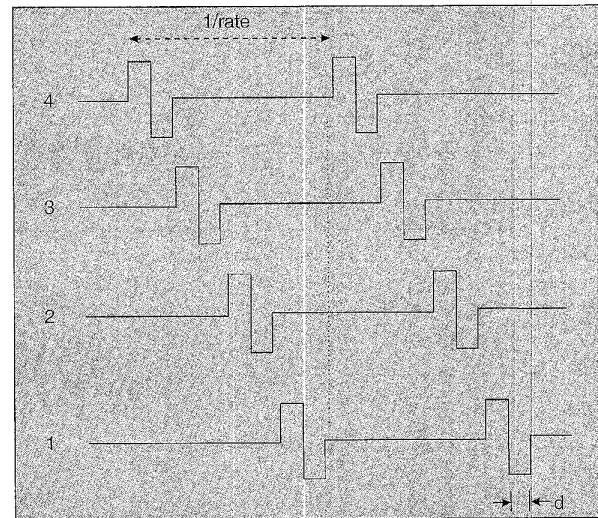
The CA approach, used in the Ineraid device, was very successful as it enabled many patients to obtain open-set speech understanding. Dorman et al. [39] reported, for a sample of 50 Ineraid patients (Figure 14), a

median score of 45% correct for word identification in CID sentences, a median score of 14% correct for monosyllabic word identification, and a median score of 14% correct for spondee (two-syllable) words. The CA multichannel approach clearly yielded superior speech-recognition performance over the single-channel approach [37]. This was not surprising given the increased frequency resolution provided by multiple-channel stimulation.

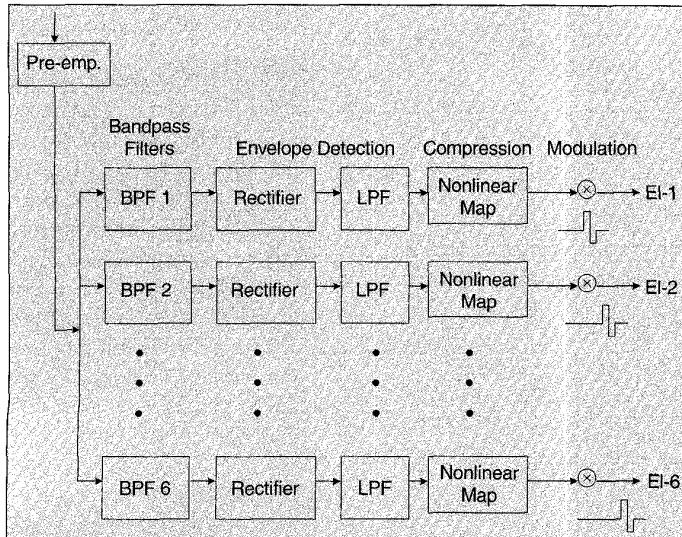
### Continuous Interleaved Sampling (CIS) Approach

The CA approach uses analog stimulation that delivers four continuous analog waveforms to four electrodes simultaneously. A major concern associated with simultaneous stimulation is the interaction between channels caused by the summation of electrical fields from individual electrodes [40]. 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.

Researchers at the Research Triangle Institute (RTI) developed the Continuous Interleaved Sampling (CIS) approach [41] which addressed the channel interaction issue by using nonsimultaneous, interleaved pulses. Trains of biphasic pulses are delivered to the electrodes in a nonoverlapping (nonsimultaneous) fashion; that is, in a way such that only one electrode is stimulated at a time (Fig. 15). The amplitudes of the pulses are derived by ex-

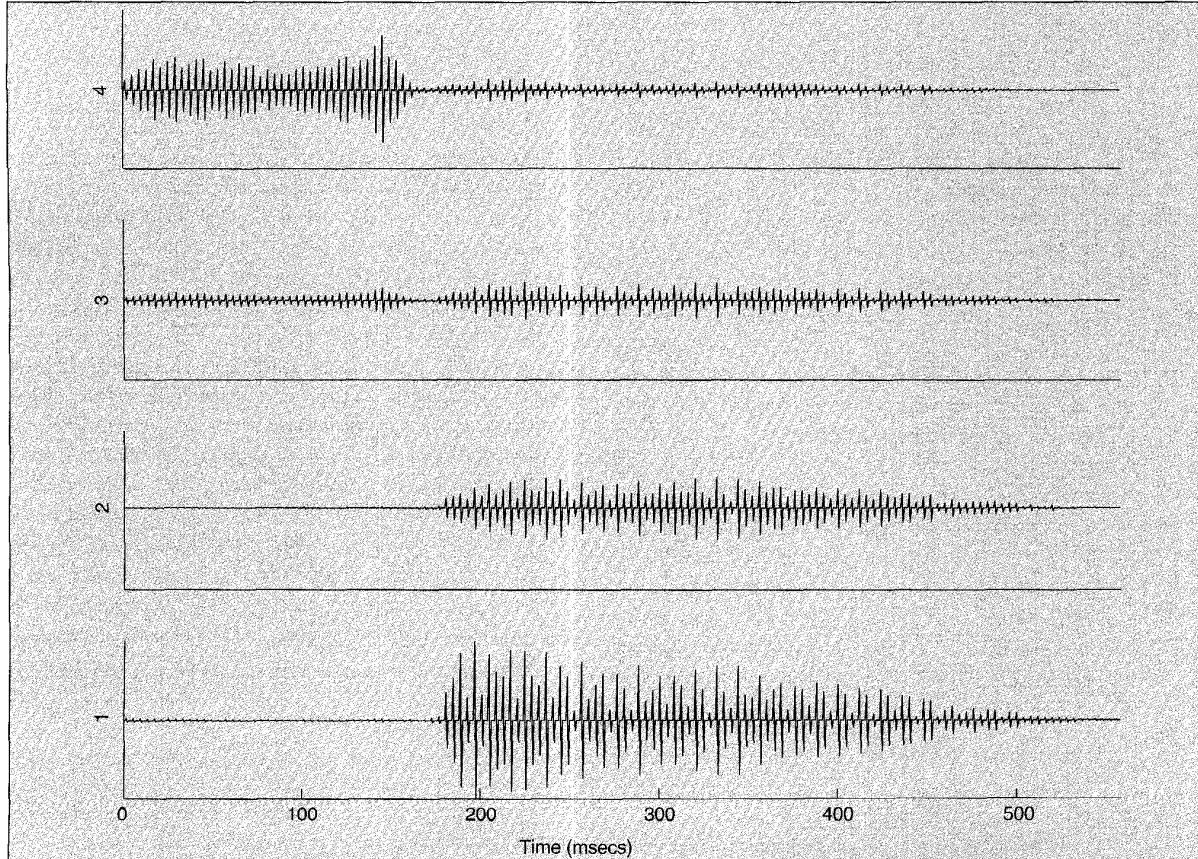


▲ 15. Interleaved pulses used in the CIS strategy. The period between pulses on each channel ( $1/\text{rate}$ ) and the pulse duration ( $d$ ) per phase are indicated.

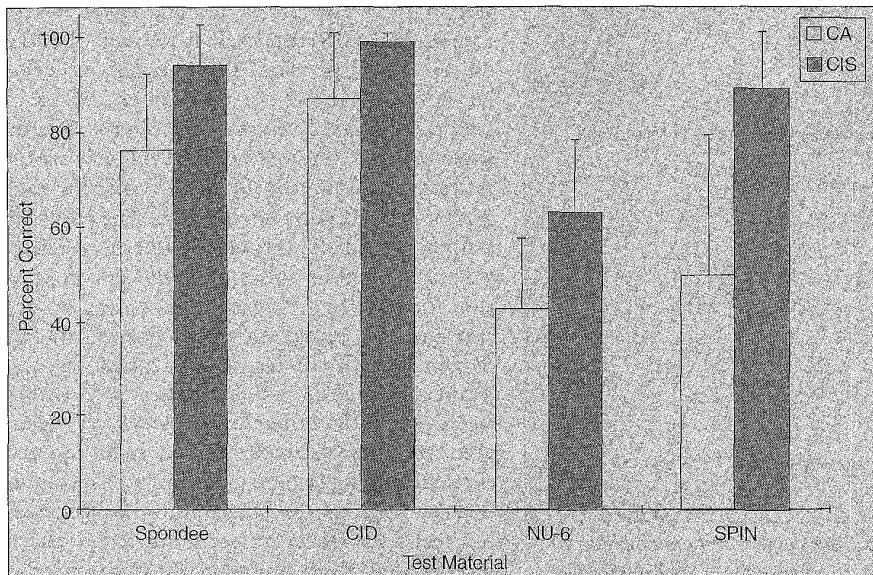


▲ 16. Block diagram of the CIS strategy. The signal is first pre-emphasized and filtered into six frequency bands. The envelopes of the filtered waveforms are then extracted by full-wave rectification and low-pass filtering. The envelope outputs are compressed to fit the patient's dynamic range and then modulated with biphasic pulses. The biphasic pulses are transmitted to the electrodes in an interleaved fashion (see Fig. 15).

tracting the envelopes of bandpassed waveforms. The CIS approach is shown in more detail in Fig. 16. The signal is first pre-emphasized and passed through a bank of bandpass filters. The envelopes of the filtered waveforms are then extracted by full-wave rectification and low-pass filtering (typically with 200 or 400 Hz cutoff frequency). The envelope outputs are finally compressed and then used to modulate biphasic pulses. A nonlinear compression function (e.g., logarithmic) 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 nonoverlapping fashion (see Fig. 15). The rate at which the pulses are delivered to the electrodes has been found to have a major impact on speech recognition. High pulse-rate stimulation typically yields better performance than low pulse-rate stimulation. Figure 17 shows, as an example, the pulsatile waveforms produced for the syllable "sa" using a simplified implementa-



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



▲ 18. Comparison between the CA and the CIS approach [41]. Mean percent correct scores for monosyllabic word (NU-6), keyword (CID sentences), spondee (two syllable words), and final word (SPIN sentences) recognition. Error bars indicate standard deviations.

tion of the CIS strategy. The pulse amplitudes were estimated by extracting the envelopes of the filtered waveforms (Fig. 13).

Several studies (e.g., [14, 41, 42]) were conducted by RTI and other institutions comparing the differences in performance between the CA and CIS strategies. The results [41] for seven patients tested on open-set recognition of 50 monosyllabic words (NU-6) and 100 keywords from the CID test are shown in Fig. 18. These results were obtained at RTI using patients fitted with the Ineraid device, which employs the CA approach. The same patients were also tested using a lab implementation of the CIS approach. As shown in Fig. 18, the mean scores obtained with the CIS processor were significantly higher than the corresponding scores obtained with the CA approach. Several other investigators replicated RTI's findings (e.g., Dorman and Loizou [43, 44], Boex et al. [42]). Several factors could be responsible for the success of the CIS approach over the CA approach: (1) use of nonsimultaneous stimulation that minimizes channel interaction, (2) use of six channels rather than four, and (3) representation of rapid envelope variations with the use of high pulse-rate stimulation. The CIS strategy is currently being used in three commercially available implant devices, namely the Clarion device, the Med-El device, and the new Nucleus CI24M device.

#### CIS parameters

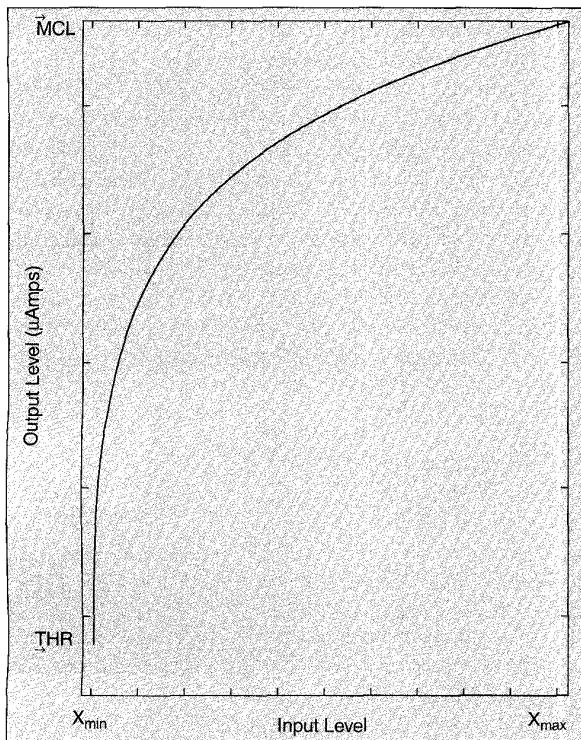
There are a number of parameters associated with the CIS approach that could be varied to optimize speech-recognition performance for each patient [14, 43]. These parameters include:

▲ **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 2500 pulses/sec have been used. The "optimal" pulse rate, as far as speech-recognition performance is concerned, varies from patient to patient. Wilson et al. [14] reported that some patients obtain a maximum performance on the 16-consonant recognition task with a pulse rate of 833 pulses/sec and a pulse duration of 33  $\mu$ sec/phase. Other patients obtain small but significant increases in performance as the pulse rate increases from 833 pps to 1365 pps, and from 1365 pps to 2525 pps, using 33

$\mu$ sec/phase pulses. One would expect that better recognition performance would be obtained with very high pulse rates, since high pulse-rate stimulation can better represent fine temporal variations. However, this was not found to be true for all patients, at least over this tested range of pulse rates.

▲ **Stimulation order.** The stimulation order can be varied to minimize possible interaction between channels. The stimulation order refers to the order with which the electrodes are stimulated. One possibility is to stimulate the electrodes in an apex-to-base order, i.e., first stimulate electrode 1, then electrode 2, etc., and lastly, stimulate electrode 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., in the order of 6-3-5-2-4-1, which maximizes the spatial separation between stimulated electrodes. Like the pulse rate, preference for stimulation order varies from patient to patient. Some patients prefer the apex-to-base stimulation because, as they say, speech sounds more natural and more intelligible, while other patients prefer the staggered order stimulation.

▲ **Compression function.** The compression (of envelope outputs) is an essential component of the CIS processor because it transforms acoustical amplitudes into electrical amplitudes. This transformation is necessary because the range in acoustic amplitudes in conversational speech is considerably larger than the implant patient's dynamic range. Dynamic range is



▲ 19. Example of a logarithmic compression map commonly used in the CIS strategy. The compression function maps the input acoustic range  $[x_{\min}, x_{\max}]$  to the electrical range [THR, MCL].  $x_{\min}$  and  $x_{\max}$  are the minimum and maximum input levels respectively, THR is the threshold level, and MCL is the most comfortable level.

defined here as the range in electrical amplitudes between threshold (barely audible level) and loudness uncomfortable level (extremely loud). In conversational speech, the acoustic amplitudes may vary over a range of 30 dB. Implant listeners, however, may have a dynamic range as small as 5 dB. For that reason, the CIS processor compresses, using a nonlinear compression function, the acoustic amplitudes to fit the patient's electrical dynamic range. The logarithmic function is commonly used for compression because it matches the loudness between acoustic and electrical amplitudes [45, 46]. It has been shown that the loudness of an electrical stimulus in microamps is analogous to the loudness of an acoustic stimulus in dB. Logarithmic compression functions of the form  $Y = A \log(x) + B$  are typically used, where  $x$  is the acoustic amplitude (output of envelope detector),  $A$  and  $B$  are constants, and  $Y$  is the (compressed) electrical amplitude. Other types of compression functions used are power-law functions of the form  $Y = Ax^p + B$ , ( $p < 1$ ). The advantage of using power-law functions is that the shape, and particularly the steepness of the compression function, can be easily controlled by simply varying the value of the exponent  $p$ . The constants  $A$  and  $B$  are chosen such that the input acoustic range  $[x_{\min}, x_{\max}]$  is mapped to the electrical dynamic range [THR, MCL], where THR is the threshold level and MCL is the most

comfortable level measured in  $\mu$ amps (see Fig. 19). For the power-law compression function, the constants  $A$  and  $B$  can be computed as follows:

$$A = \frac{MCL - THR}{x_{\max}^p - x_{\min}^p} \quad (1)$$

$$B = THR - Ax_{\min}^p \quad (2)$$

The values THR and MCL may vary from electrode to electrode.

### Nucleus Multi-Electrode Implant

The Nucleus multi-electrode implant, manufactured by Nucleus Limited, was developed at the University of Melbourne, Australia, by Clark and his colleagues [47]. The Nucleus device has gone through a number of improvements since it was first introduced in the early 1980s. Initially, the Nucleus device employed a feature-extraction approach that was based on a fundamentally different principle than the CA or CIS approach. Rather than presenting waveform information, obtained by filtering the speech signal into a few frequency bands, the Nucleus processor presented spectral features, such as formants, obtained by formant-extraction algorithms. Several refinements were made over the years to the feature-extraction strategy that improved the speech-recognition performance significantly. In the early 1990s, the Nucleus device adopted a new signal-processing strategy that was based on a filterbank approach that did not require the extraction of any features other than the extraction of spectral maxima. This strategy is currently being employed in the Nucleus Spectra 22 processor, a commercially available implant. The following sections give a more detailed description of the evolution of the Nucleus multichannel implant.

### F0/F2

The F0/F2 strategy was the first strategy developed for the Nucleus device in the early 1980s [47, 48]. In this strategy, the fundamental frequency (F0) and the second formant (F2) are extracted from the speech signal using zero crossing detectors. One zero-crossing detector is used to estimate F0 from the output of a 270 Hz low-pass filter, and another zero-crossing detector is used to estimate F2 from the output of a 1,000-4,000 Hz bandpass filter. The amplitude of F2 is estimated with an envelope detector by rectifying and low-pass filtering (at 35 Hz) the bandpassed signal. The F0/F2 processor conveys F2 frequency information by stimulating the appropriate electrode in the 22-electrode array. Voicing information is conveyed with F0 by stimulating the selected electrode at a rate of F0 pulses/sec. The amplitude of the pulses is set in proportion to the amplitude of F2. During unvoiced segments (e.g., consonant /s/) the selected electrode is stimulated at quasi-random intervals with an

## It is not easy to assess the significance of individual factors on speech perception due to the interaction among factors.

average rate of 100 pulses/sec. Initial results [49] with the F0/F2 strategy were encouraging as it enabled some patients to obtain open-set speech understanding.

### F0/F1/F2

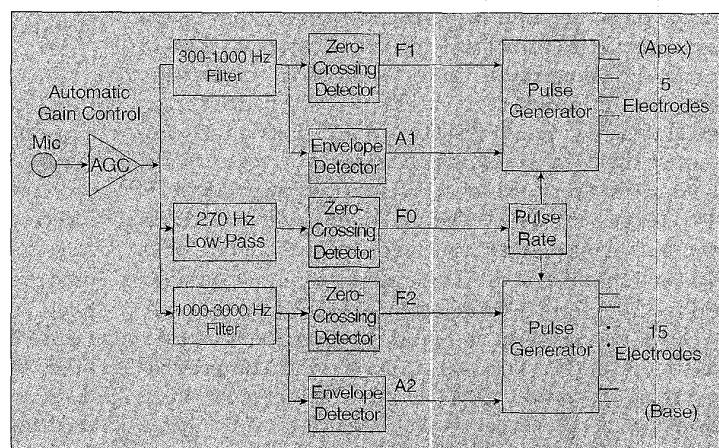
The F0/F2 strategy was later modified to include information about the first formant frequency (F1) [50] and became available in 1985 with the Nucleus wearable speech processor (WSP). An additional zero-crossing detector was included to estimate F1 from the output of a 280-1,000 Hz bandpass filter. The block diagram of the F0/F1/F2 processor is shown in Fig. 20. The processor selects two electrodes for stimulation, one corresponding to the F1 frequency, and one corresponding to the F2 frequency. The five most apical electrodes were dedicated to F1, since they can represent frequencies up to 1,000 Hz, while the remaining 15 electrodes were dedicated to F2 since they can represent frequencies above 1,000 Hz. For voiced segments, two electrical pulses were produced. One pulse was applied to an electrode pair chosen according to F2, and the second pulse was applied to an electrode pair chosen according to F1. The pulses were biphasic with each phase lasting 200  $\mu$ secs. A 800  $\mu$ sec spacing between pulses was used to avoid any interaction that might occur due to temporal channel interactions. The amplitudes of the two pulses were set in proportion to the corresponding amplitudes of F1 and F2. As in the F0/F2 processor, the electrodes were stimulated at a rate of F0 pulses/sec for voiced segments and at an average rate of 100 pulses/sec for unvoiced segments.

The addition of F1 information improved the speech-recognition performance of patients wearing the Nucleus cochlear implant. This was not surprising given the importance of F1-F2 for normal-hearing listeners on speech recognition. Within-patient comparisons between the F0/F2 and F0/F1/F2 strategies demonstrated improvements in speech understanding with the F0/F1/F2 strategy. Dowell et al. [51] found that the average scores on word recognition increased from 30% correct with the F0/F2 processor to 63% correct with the F0/F1/F2 processor. Tye-Murray et al. [52] also reported that the mean scores on monosyllabic word identification (NU-6) improved from 8% correct with the F0/F2 processor to 28% correct with the F0/F1/F2 processor. No significant difference was

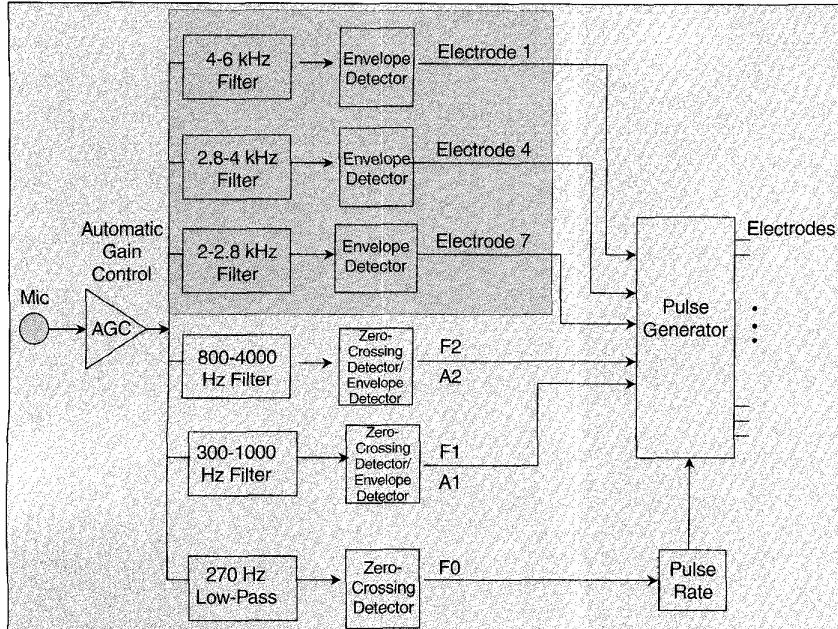
found between the F0/F2 and F0/F1/F2 processors on tests of consonant recognition in the hearing-only condition. Significant improvements with the F0/F1/F2 strategy were found, however, in the visual-plus-hearing condition. The finding that the F0/F1/F2 strategy did not yield significant improvements on consonant-recognition scores was not surprising given that the F0/F1/F2 strategy emphasizes low-frequency information, which is required for vowel recognition. However, the majority of the consonants contain high-frequency information, and this has motivated the refinement of the F0/F1/F2 strategy to the MPEAK strategy.

### MPEAK

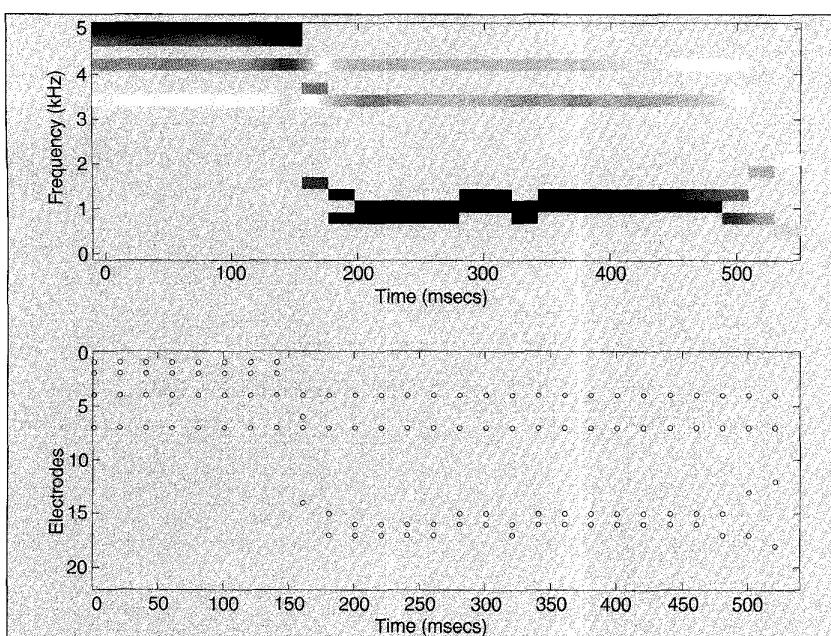
Further improvements to the F0/F1/F2 processor were made in the late 1980s by Cochlear Pty. Limited (a subsidiary of Nucleus Limited) in collaboration with the University of Melbourne [53, 54]. The improvements included new hardware as well as refinement of the F0/F1/F2 strategy to include high-frequency information. A custom integrated circuit for digital signal processing was used that considerably reduced the size and weight of the new processor, now called the Miniature Speech Processor (MSP). A new coding strategy, called MULTIMEAK (or MPEAK), was used that extracted from the speech signal, in addition to formant information, high-frequency information. The block diagram of the MPEAK strategy is shown in Fig. 21. Similar to the F0/F1/F2 strategy, the extraction of the formant frequencies F1 and F2 was performed using zero-crossing detectors, and the amplitudes of F1 and F2 were computed using envelope detectors. The frequency range for F2 was refined in the MPEAK strategy to 800-4,000 Hz. Additional high-frequency information was ex-



▲ 20. Block diagram of the F0/F1/F2 strategy. The fundamental frequency (F0), the first formant (F1), and the second formant (F2) are extracted from the speech signal using zero crossing detectors. Two electrodes are selected for pulsatile stimulation, one corresponding to the F1 frequency, and one corresponding to the F2 frequency. The electrodes are stimulated at a rate of F0 pulses/sec for voiced segments and at a quasi-random rate (with an average rate of 100 pulses/sec) for unvoiced segments.



▲ 21. Block diagram of the MPEAK strategy. Similar to the F0/F1/F2 strategy, the formant frequencies ( $F_1$ ,  $F_2$ ), and fundamental frequency ( $F_0$ ) are extracted using zero-crossing detectors. Additional high-frequency information is extracted using envelope detectors from three high-frequency bands (shaded blocks). The envelope outputs of the three high-frequency bands are delivered to fixed electrodes as indicated. Four electrodes are stimulated at a rate of  $F_0$  pulses/sec for voiced sounds, and at a quasi-random rate for unvoiced sounds.



▲ 22. An example of the MPEAK strategy using the syllable "sa." The bottom panel shows the electrodes stimulated and the top panel shows the corresponding amplitudes of stimulation.

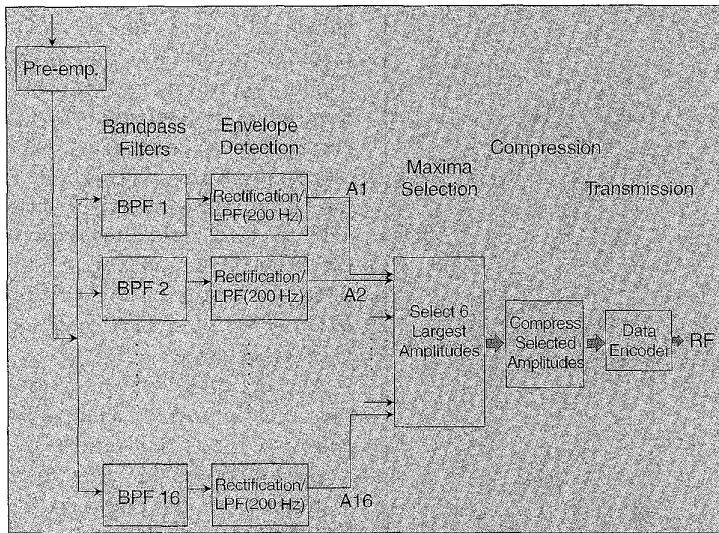
tracted, using envelope detectors, from the frequency bands 2,000-2,800 Hz, 2,800-4,000 Hz, and 4,000-6,000 Hz. The motivation for using the three additional bandpass filters ( $> 2$  kHz) was twofold: (1) to

enhance the representation of the second formant ( $F_2$ ), and (2) to include high-frequency information, which is important for the perception of consonants. The estimated envelope amplitudes of the three bandpass filters were delivered to fixed electrodes. Electrodes 7, 4, and 1 were allocated to the outputs of the filters 2-2.8 kHz, 2.8-4 kHz, and 4-6 kHz respectively.

The MPEAK strategy stimulates four electrodes at a rate of  $F_0$  pulses/sec for voiced sounds, and at quasi-random intervals with an average rate of 250 pulses/sec for unvoiced sounds. For voiced sounds, stimulation occurs on the  $F_1$  and  $F_2$  electrodes and on the high-frequency electrodes 4 (2,000-2,800 Hz) and 7 (2,800-4,000 Hz). The high-frequency electrode 1 does not get stimulated because there is generally little energy in the spectrum above 4 kHz for voiced sounds. For unvoiced sounds, stimulation occurs on high-frequency electrodes 1, 4 and 7 as well as on the electrode corresponding to  $F_2$ . The electrode corresponding to  $F_1$  does not get stimulated because there is generally little energy below 1,000 Hz for unvoiced sounds (e.g., /s/). Figure 22 shows, as an example, a simplified implementation of the MPEAK strategy using the syllable "sa."

Given the addition of high-frequency information, one would expect that the MPEAK strategy would perform better than the F0/F1/F2 strategy on consonant identification. Indeed, Wallenberger and Battmer [55] found that MPEAK strategy yielded a mean improvement of 17% on consonant identification using a group of five patients. A mean improvement of 28% on open-set sentence recognition

was also found with the MPEAK strategy for six patients. Several other studies (e.g., [56] and [57]) were conducted that compared the performance between the MPEAK and the F0/F1/F2 strategies. These studies con-



▲ 23. Block diagram of the Spectral Maxima (SMS) strategy. The signal is first pre-emphasized and then processed through a bank of 16 bandpass filters spanning the frequency range 250 to 5,400 Hz. The envelopes of the filtered waveforms are computed by full-wave rectification and low-pass filtering at 200 Hz. The six (out of 16) largest envelope outputs are then selected for stimulation in 4 msec intervals.

firmed that the MPEAK strategy achieved significant improvements over the F0/F1/F2 strategy on open-set speech recognition.

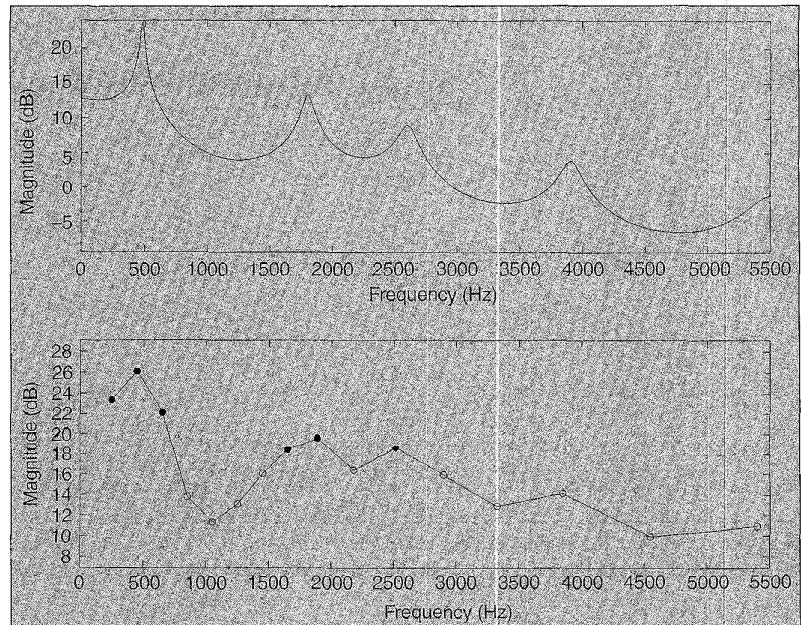
Although the MPEAK strategy has proved to be an efficient strategy for extracting important information from the speech signal, it has one major limitation. The MPEAK strategy, as well as the F0/F2 and F0/F1/F2 strategies, tend to make errors in formant extraction, especially in situations where the speech signal is embedded in noise. This limitation, which is inherent in feature-extraction algorithms, motivated the development of the next-generation processor for the Nucleus multi-electrode cochlear implant.

### SMS

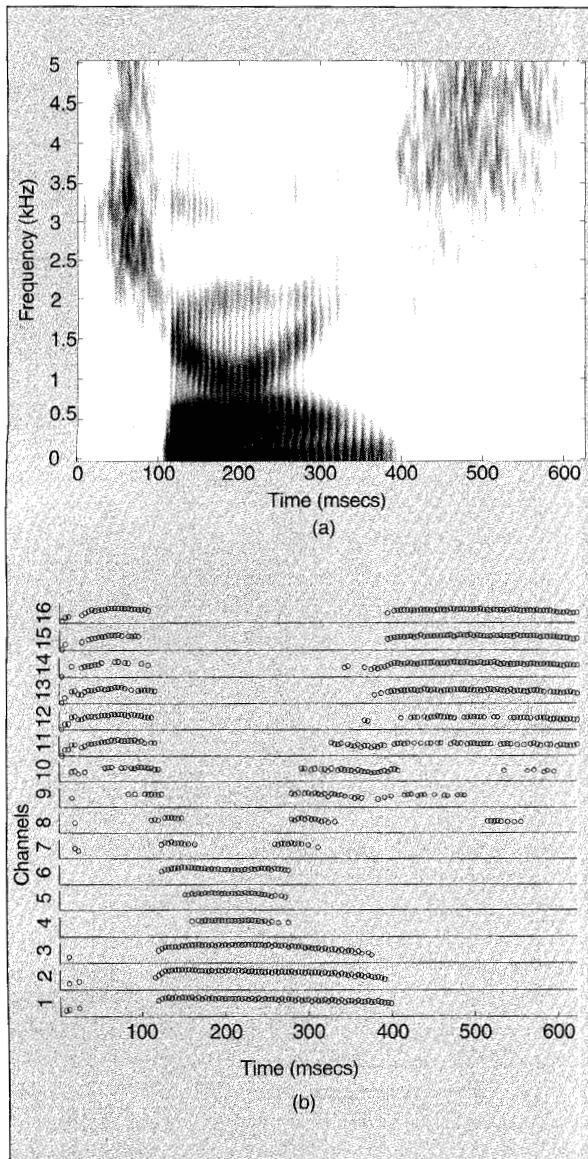
A new processor, called the Spectral Maxima Sound Processor (SMS), was developed in the early 1990s for the University of Melbourne/Nucleus multi-electrode cochlear implant [58]. Unlike previous processors developed for the Nucleus implant, the SMS processor did not extract any features (e.g., F1, F2) from the speech waveform. Instead, it analyzed the speech signal using a bank of 16 bandpass filters and a spectral maxima detector. Figure 23 shows the block diagram of the SMS processor. The signal from

the microphone is first preamplified and then sent through a bank of 16 bandpass filters with center frequencies ranging from 250 to 5,400 Hz. The output of each filter is rectified and low-pass filtered with a cutoff frequency of 200 Hz. After computing all 16 filter outputs, the SMS processor selects, at 4 msec intervals, the six largest filter outputs. The six amplitudes of the spectral maxima are finally logarithmically compressed, to fit the patient's electrical dynamic range, and transmitted to the six selected electrodes through a radio-frequency link. Note that the term "maxima" refers to the largest filter amplitudes, which are not necessarily the spectral peaks. As illustrated in Fig. 24, several maxima may come from a single spectral peak.

One electrode is allocated for each of the 16 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. Only the 16 most-apical electrodes are activated; the remaining basal electrodes in the 22-electrode implant are left inactive. Six biphasic pulses are delivered to the selected electrodes in an interleaved (i.e., nonsimultaneous) fashion at a rate of 250 pulses/sec. Unlike the F0/F1/F2 and MPEAK processors, the SMS processor delivers biphasic pulses to the elec-



▲ 24. An example of spectral maxima selection in the SMS strategy. The top panel shows the LPC spectrum of the vowel /eh/ (as in "head"), and the bottom panel shows the 16 filterbank outputs obtained by bandpass filtering and envelope detection. The filled circles indicate the six largest filterbank outputs selected for stimulation. As shown, more than one maximum may come from a single spectral peak.



▲ 25. Example of the SMSP 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.

trodes at a constant rate of 250 pps for both voiced and unvoiced sounds. Figure 25 illustrates the pattern of electrical stimulation for the word "choice." As can be seen, the electrodes selected for stimulation vary depending upon the spectral content of the signal.

Initial comparisons of the SMSP and the MPEAK strategy using a single patient showed significant improvements with the SMSP strategy on word, consonant, and vowel recognition [58]. The SMSP strategy was later refined and incorporated in the Nucleus Spectra 22 processor.

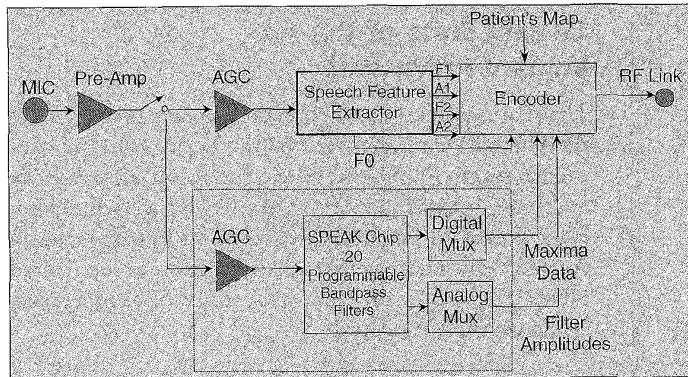
## Current State-of-the-Art Implant Processors

There are currently two cochlear implant processors in the United States approved by the Food and Drug Administration (FDA), the Nucleus Spectra 22 processor and the Clarion processor. There is also a cochlear implant processor, manufactured by Med-El Corporation, Austria, which is currently in clinical trials in the United States. This section provides an overview of commercially available implant processors.

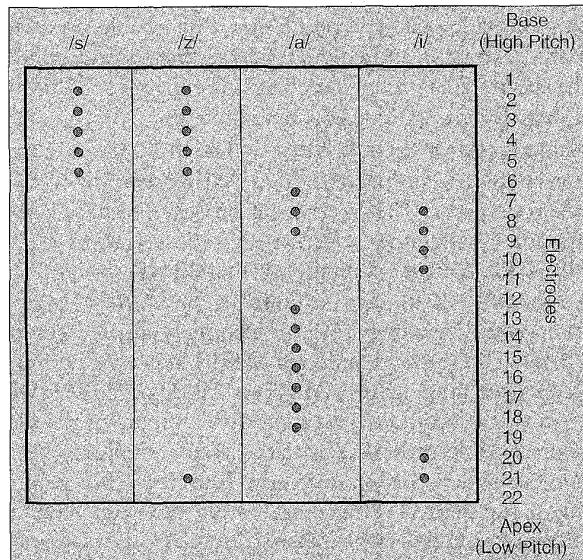
### Nucleus Spectra 22 Processor

The Spectra 22 is the latest speech processor of the Nucleus 22-channel implant system manufactured by Cochlear Pty. Limited, Australia. It includes the functions of previous speech processors (MSP) and also incorporates new circuitry for the Spectral Peak (SPEAK) speech strategy [59]. Two custom integrated circuits are used that perform most of the signal processing needed to convert the speech signal into electrical pulses. The two custom chips provide analog preprocessing, a filterbank, a speech feature extractor, and a digital encoder that encodes either the spectral maxima or speech features (e.g., F1, F2) into signals for the radio-frequency link (Fig. 26). An implanted receiver decodes these signals and presents electrical pulses, according to the decoded instructions, to the electrode array. The Spectra 22 processor can be programmed with either a feature-extraction strategy (e.g., F0/F1/F2, MPEAK strategy) or the SPEAK strategy.

The SPEAK strategy is similar to the SMSP strategy. In the SPEAK strategy [59] the incoming signal is sent to a bank of 20 (rather than 16 in SMSP) filters with center frequencies ranging from 250 Hz to 10 kHz. 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 six maxima. Figure 27 shows examples of electrical stimulation patterns for four different sounds using the SPEAK strategy. Five maxima were selected for /s/, while 10 maxima were selected for /a/. 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 patients' individual parameters. For broadband spectra, more maxima are selected and the stimulation rate is slowed down. For spectra with limited spectral content, fewer maxima are selected and the stimulation rate increases to provide more temporal information. The SPEAK strategy provides more information than any of the previous strategies developed for the Nucleus implant because: (1) it uses up to 20 filters that span a wider frequency range, (2) it stimulates as many as 10 electrodes in a cycle, and (3) it uses an adaptive stimulation rate in order to preserve spectral as well as temporal information.



▲ 26. The architecture of the Spectra 22 processor. The processor consists of two custom monolithic integrated circuits that perform the signal processing required for converting the speech signal to electrical pulses. The two chips provide analog preprocessing of the input signal, a filterbank (20 programmable bandpass filters), a speech feature detector, and a digital encoder that encodes either the spectral maxima or speech features for stimulation. The Spectra 22 processor can be programmed with either a feature-extraction strategy (e.g., F0/F1/F2, MPEAK strategy) or the SPEAK strategy.



▲ 27. Patterns of electrical stimulation for four different sounds, /s/, /z/, /a/ and /i/ using the SPEAK strategy. The filled circles indicate the activated electrodes.

Comparison of the SPEAK strategy and the MPEAK strategy by Skinner et al. [60] using 60 patients showed that the SPEAK strategy performed better than the MPEAK strategy on vowel, consonant, monosyllabic word, and sentence recognition (see Fig. 28). Especially large improvements in performance were found with tests in noise. This finding was not surprising given that the MPEAK strategy is based on feature-extraction algorithms that are known to be susceptible to errors, especially in noisy environments. In contrast, the SPEAK strategy is based on a filterbank approach that does not extract any features from the speech signal.

The SPEAK strategy is now incorporated in a new investigational device that is currently in clinical trials at 13 clinics in the United States. The new cochlear implant system, called Nucleus 24 (CI24M), is available in two sizes: the regular size, which can be worn by the waist, and the ear-level size, which can be worn behind the ear. The ear-level processor is the size of a behind-the-ear hearing aid. Some of the features of the Nucleus 24 implant system include: (1) high-rate stimulation strategies including the CIS strategy, (2) two additional electrodes to be placed outside the inner ear to allow different modes of stimulation, and (3) a removable internal magnet for future MRI compatibility.

### Clarion Processor

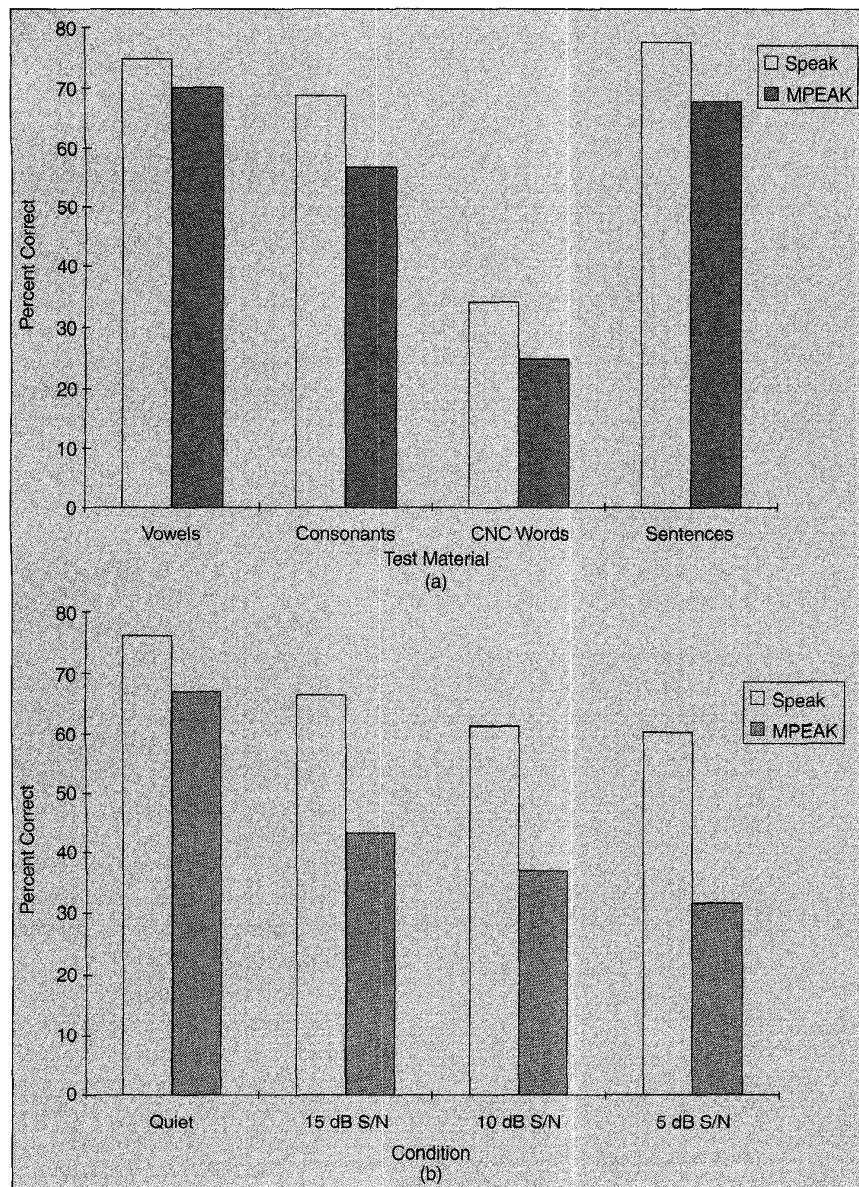
The Clarion cochlear implant system [61, 62] is the result of cooperative efforts among the University of California at San Francisco (UCSF), Research Triangle Institute (RTI), and the device manufacturer, Advanced Bionics Corporation (evolved from MiniMed Technologies). The Clarion implant supports a variety of speech-processing options and stimulation patterns. The stimulating waveform can be either analog or pulsatile, the stimulation can be either simultaneous or sequential, and the stimulation mode can be either monopolar or bipolar. The Clarion processor can be programmed with either the CA strategy or the CIS strategy. In the CA mode, the acoustic signal is processed through eight filters, compressed and delivered simultaneously to eight electrode pairs. Analog waveforms are delivered to each electrode at a rate of 13,000 samples/sec per channel. The CA strategy emphasizes detailed temporal information at the expense of reduced spatial selectivity due to simultaneous stimulation. For some patients, use of simultaneous stimulation results in a loss of speech discrimination due to channel interaction. This problem is alleviated in the CIS mode, which delivers biphasic pulses to all eight channels in an interleaved manner. In the CIS mode, the signal is first pre-emphasized and passed through a bank of eight bandpass filters. The envelopes of the filtered waveforms are then extracted by full-wave rectification and low-pass filtering. The envelope outputs are finally compressed to fit the patient's dynamic range and then used to modulate biphasic pulses. Pulses are delivered to eight electrodes at a maximum rate of 833 pulses/sec per channel in an interleaved fashion.

The Clarion processor (version 1.0) was recently approved by the FDA, and the initial results on open-set speech recognition were very encouraging. In a recent study by Loeb and Kessler [63], 32 of the first 46 patients fitted with the Clarion implant obtained moderate to excellent open-set speech recognition scores (30%-100% on CID sentence test) at 12 months. Preliminary studies by Tyler et al. [64] showed that the pulsatile version (CIS) of the Clarion processor (version 1.0) obtained su-

perior performance over the analog (CA) version of the processor. This was found to be true with six patients (one-third of the patients considered in the Tyler et al. study) who could be fitted satisfactorily with the analog version. Tyler et al. [64] also found that Clarion patients with nine months of experience with the device performed better than Ineraid patients (using the CA strategy) and Nucleus patients (using the F0/F1/F2 strategy) with comparable experience (Fig. 29).

Several changes were recently made to the Clarion implant system (ver. 1.0) that produced Clarion 1.2. Some of those changes include: (1) smaller speech processor,

(2) improved filter implementation using bandpass filters with 30 dB/octave rolloffs, and (3) enhanced preprocessing. Preliminary data obtained six months postimplantation showed that these changes produced an improvement in performance. Although both Clarion 1.0 and 1.2 support simultaneous analog stimulation strategies, only a small number of patients were successfully programmed with a fully simultaneous strategy via the standard bipolar electrode configuration. The electrode coupling configuration was changed in the new Clarion implant system, called Clarion S-Series, to include an enhanced bipolar coupling mode. Preliminary results showed that more than 90% of the Clarion S-Series users can be successfully programmed with an analog strategy via the enhanced bipolar coupling mode, and that about 50% of the users preferred the CA strategy over the CIS strategy in the enhanced bipolar mode. In addition to the enhanced bipolar coupling mode, the new Clarion S-Series processor provides the option for composite simultaneous and sequential stimulation through the use of a new stimulation strategy. The new strategy, called Paired Pulsatile Sampler (currently under investigation), can deliver pulses simultaneously on two channels, thereby increasing the maximum pulse rate per channel to 1666 pulses/sec.



▲ 28. Comparative results between the SPEAK and the MPEAK strategy in quiet (a) and in noise (b) for 63 implant patients (Skinner et al. [60]). Bottom panel shows the mean scores on CUNY sentences presented at different S/N in eight-talker babble using the MPEAK and SPEAK strategies.

### Med-El Processor

The Med-El cochlear implant processor, manufactured by Med-El Corporation, Austria, is currently in clinical trials in the United States. The implant processor [65] is based on the Motorola 56001 DSP and can be programmed with either a high-rate CIS strategy or a high-rate SPEAK-type strategy. The Med-El cochlear implant (also referred to as COMBI-40[66]) uses a very soft electrode carrier specially designed to facilitate deep electrode insertion into the cochlea. Because of the capability of deep electrode insertion (up to 30 mm), the electrodes are

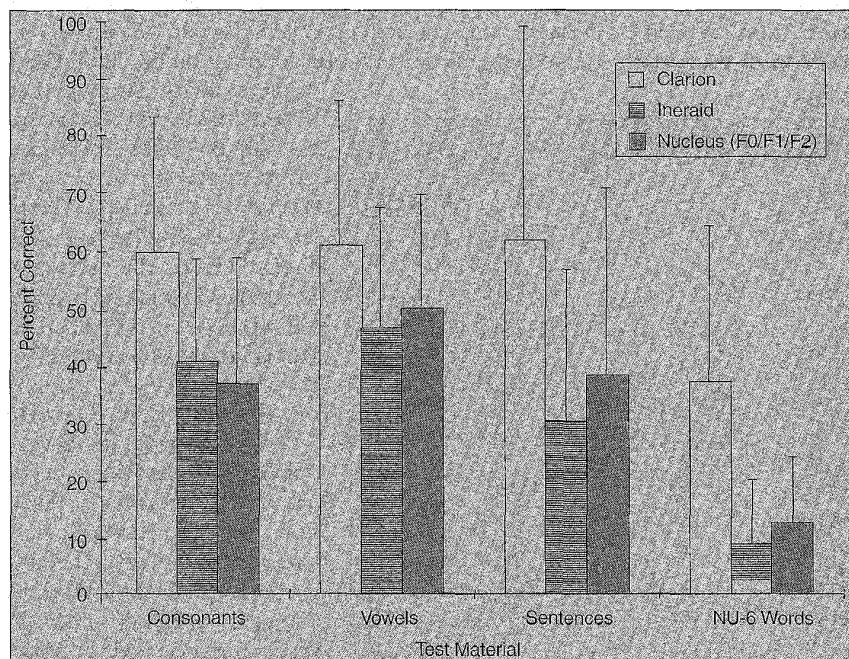
spaced 2.8 mm apart spanning a considerably larger distance (20.6 mm) in the cochlea than any other commercial cochlear implant. The motivation for using wider spacing between electrode contacts is to increase the number of perceivable channels.

The Med-El processor has the capability of generating 12,500 pulses/sec for a high-rate implementation of the CIS strategy. The amplitudes of the pulses are derived as follows. The signal is first pre-emphasized and then ap-

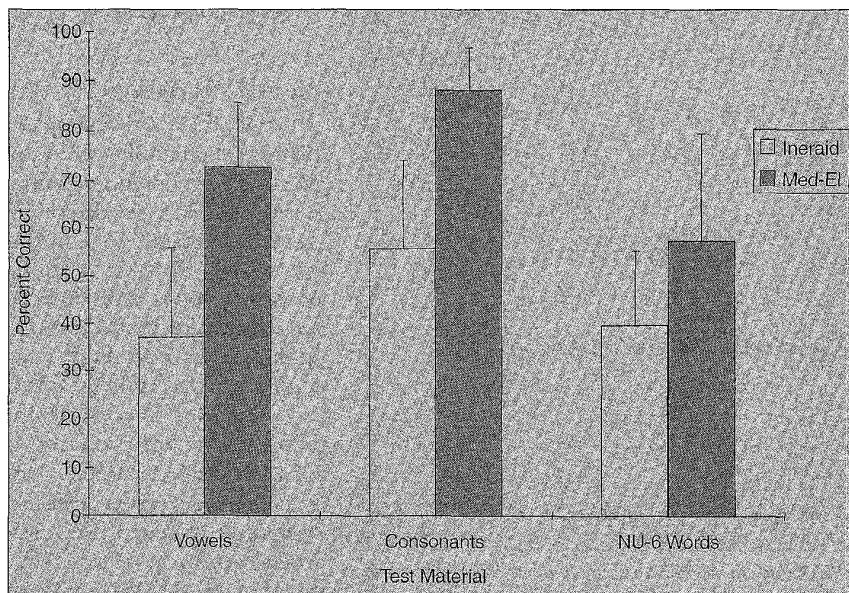
plied to a bank of eight (logarithmically-spaced) bandpass filters of Butterworth type and of sixth-order. The bandpass filter outputs are full-wave rectified and low-pass filtered with a cutoff of 400 Hz. The low-pass filter outputs are finally mapped, using a logarithmic-type compression function, to the patient's dynamic range. Biphasic pulses, with amplitudes set to the mapped filter outputs, are delivered in an interleaved fashion to eight monopolar electrodes at a maximum rate of 1,515 pulses/sec per channel. The pulses are transmitted transcutaneously through a radio-frequency link [67]. A new 12-electrode cochlear implant, called COMBI-40+, is now available from Med-El and is currently in clinical trials. The COMBI-40+ processor can generate up to 18,000 pulses/sec for a high-rate implementation of the CIS strategy.

The Med-El processor can also be programmed with a high-rate "n-of-m" strategy (SPEAK-type). In this strategy, the signal is filtered into  $m$  frequency bands, and the processor selects, out of  $m$  envelope outputs, the  $n$  ( $n < m$ ) envelope outputs with the largest energy. Only the electrodes corresponding to the  $n$  selected outputs are stimulated at each cycle. For example, in a 4-of-8 strategy, from a maximum of eight channel outputs, only the four channel outputs with the largest amplitudes are selected for stimulation at each cycle. The "n-of-m" strategy is very similar to the SPEAK strategy used in the Nucleus Spectra 22 processor. The main difference is that the selected channels are stimulated at a considerably higher pulse rate.

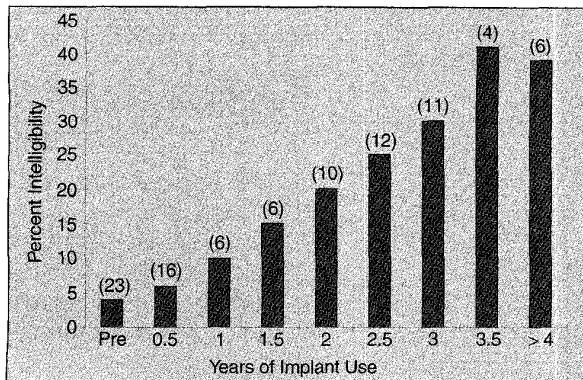
The Med-El implant processor is widely used in Europe [66]. A percutaneous version of the Med-El implant processor is currently being used successfully in the United States by a number of Ineraid patients. Results on consonant and vowel recognition with Ineraid patients fit with the Med-El processor were re-



▲ 29. Comparative results between patients wearing the Clarion (1.0) device, the Ineraid device (CA) and the Nucleus (F0/F1/F2) device (Tyler et al. [64]) after 9 months of experience.



▲ 30. Mean speech-recognition performance of seven Ineraid patients obtained before and after they were fitted with the Med-El processor and worn their device for more than 5 months.



▲ 31. Mean speech-intelligibility scores of prelingually deafened children (wearing the Nucleus implant) as a function of number of years of implant use (Osberger et al. [71]). Numbers in parenthesis indicate the number of children used in the study.

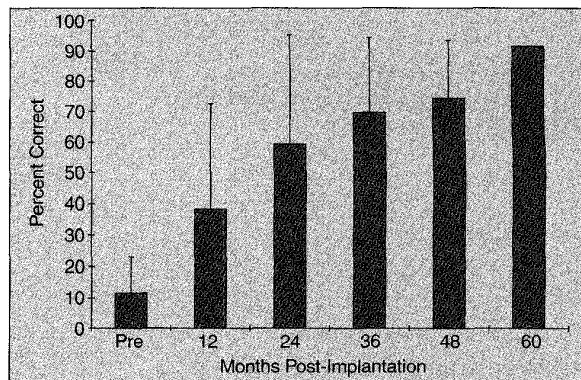
ported by Dormant and Loizou [43, 44, 68, 69]. Figure 30 compares the speech-recognition performance of seven Ineraid patients obtained before and after they were fitted with the Med-El processor and worn the device for more than five months. (The Ineraid patients have six implanted monopolar electrodes, so only the six most-apical channels of the Med-El processor are used.) As can be seen, significant improvements were obtained on all test materials.

### Cochlear Implants in Children

Postlingually deafened adults are not the only recipients of cochlear implants. Children age 2 or older have also received and continue to receive cochlear implants. The implications of a successful implant in a young child are far greater than those of an adult. This is because the child is at an age that he/she needs to develop spoken-language skills. That age is therefore extremely crucial for the child's language and cognitive development (e.g., see [18] and [70]). The implant may help a child in two important aspects of his/her development: (1) speech-production skill, i.e., the ability to speak clearly, and (2), speech-perception skill, i.e., the ability to understand speech.

### Speech-Production Skills of Children with Implants

The ability to speak is closely related to the ability to hear. If the child is not able to hear, then the child will have difficulty learning how to speak correctly. Hearing provides feedback which is used by the child to correct or improve his/her speech production skills [70]. Auditory feedback is therefore very important for learning how to speak, and cochlear implants can provide that. Research (e.g., [70], [71], and [72]) has shown that the intelligibility of speech produced by children with cochlear implants improves over time. Osberger et al. [71] measured the intelligibility of 29 prelingually deafened children (i.e., deafened before or during the development of speech and language skills) over a period of four years after implantation. Each child

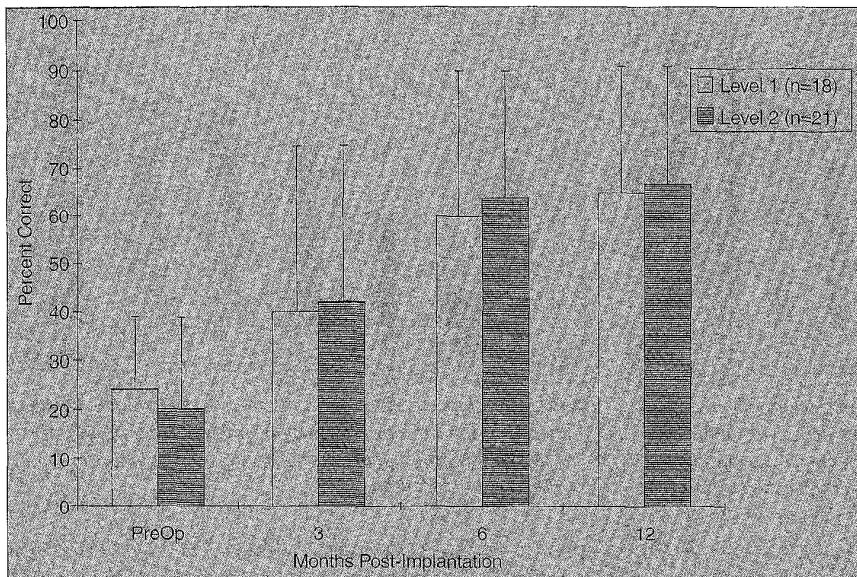


▲ 32. Speech-perception scores of prelingually deafened children (wearing the Nucleus implant) on word recognition (MTS test [18]) as a function of number of months of implant use (Miyamoto et al. [73]).

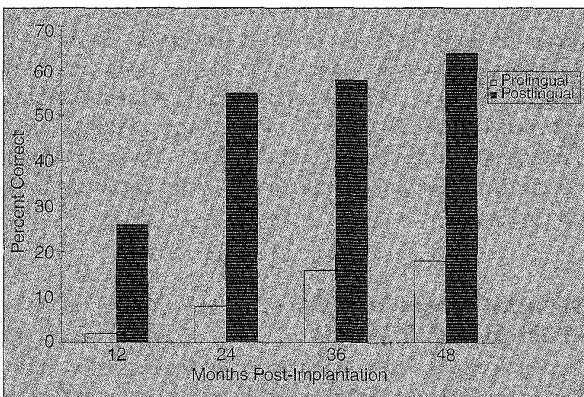
produced 10 sentences, which were evaluated for intelligibility by three expert listeners. Intelligibility was measured in terms of percentage of words correctly understood by the expert listeners. The results are shown in Fig. 31. As can be seen, intelligibility improves gradually over time. The largest changes in speech intelligibility were not observed until after the children had worn their cochlear implant device for two or more years. In fact, the mean intelligibility score of children with implants after 2.5 years of use was found to be higher than the mean score of children wearing hearing aids (with thresholds between 100 to 110 dB HL) for the same period of time. These results suggest that some children might get more benefit from a cochlear implant than from a conventional hearing aid.

### Speech-Perception Skills of Children with Implants

Research has also shown that the speech-perception abilities of children with implants improve steadily over time (e.g., [18], [73], and [74]). Figure 32 shows a longitudinal study [73] on the perception abilities of 39 prelingually deafened children using the Nucleus implant. The children were tested on the Monosyllable-Trochee-Spondee (MTS) test [75], which uses 12 pictures of nouns. In this test, the children were asked to point to the picture corresponding to the word they hear. A word score was determined by counting the number of words identified correctly. As can be seen in Fig. 32, the mean scores improved over time. Similar improvements were also found with children wearing the Clarion implant (Fig. 33). These results demonstrate a steady improvement of speech-recognition performance for prelingually deafened children over a three to four year period of implant use. In contrast, postlingually deafened children (deafened after the development of speech and language skills) have been found to attain rapid improvement in performance over the first six months of use of their implant device [76]. In addition, postlingually deaf-



▲ 33. Performance of children with the Clarion implant on monosyllabic word (ESP test [18]) identification as a function of number of months of implant use. Two levels of test difficulty were used. Level 1 tests were administered to all children 3 years of age and younger, and level 2 tests were administered to all children 7 years of age and older.



▲ 34. Comparison in performance between prelingually deafened and postlingually deafened children on open-set word recognition (Gantz et al. [76]). The postlingually deafened children obtained significantly higher performance than the prelingually deafened children.

ened children have been found to perform better on tests of open-set speech understanding compared to prelingually deafened children (Fig. 34).

In summary, both prelingually and postlingually deafened children obtain significant benefit from cochlear implants as demonstrated by significant improvements in speech perception and speech production skills. Prelingually deafened children, including congenitally deaf children, acquire these skills at a slower rate than the postlingually deafened children. Speech perception and speech production abilities of children with cochlear implants continue to improve over a four-year period following implantation.

## 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 [77-80]. 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 strong negative effect on auditory performance. Individuals with shorter duration of auditory deprivation tend to achieve better auditory performance than individuals with longer duration 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 or adults with postlingual deafness perform better than children or adults with prelingual or congenital deafness.

▲ **Age at implantation.** Prelingually deafened persons who were implanted in adolescence have been found to obtain different levels of auditory performance than those implanted in adulthood. People implanted at an early age seem to perform better than people implanted in adulthood. It still remains unclear, however, whether children should be implanted at a minimum age of 2 years for maximum auditory performance.

▲ **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. The speech perception and speech production skills of children continue to improve over a four-year period following implantation.

▲ **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) sig-

nal-processing strategy. There are also factors, such as patient's level of intelligence and communicativeness, which are unrelated to deafness but may also affect auditory performance. Aural rehabilitation, commitment from the cochlear implant patient in terms of time and effort, and support from family, friends, and workplace also play an important role.

Taking the above factors into account, Blamey [80] developed a three-stage model of auditory performance for postlingually deafened adults (Fig. 35). Stage 1 begins after normal language development. During stage 1, the patient has normal hearing abilities and the level of auditory performance is close to 100%. Stage 2 begins at the onset of deafness. A drop in auditory performance immediately occurs at the onset of deafness by an amount that varies among patients and may depend on the etiology of the hearing loss. The auditory performance keeps decreasing, due to auditory deprivation, until implantation. Stage 3 begins with implantation, and the implant patient immediately attains improvements in auditory performance depending on the duration of deafness. As the patient's experience with the implant device increases, the level of auditory performance rises as a result of learning.

### Acoustic Simulations of Cochlear Implants

It is not surprising that there is a large variance in speech performance among implant patients given the factors above that may affect performance. Unfortunately, it is not easy to assess the significance of individual factors on speech perception due to the interaction among factors. For example, in assessing meningitis as a factor that affects auditory performance, one needs to bear in mind that meningitis is commonly associated with bone growth in the cochlea, and this bone growth can obstruct the insertion of intracochlear electrodes. So, the etiology (in this case meningitis) is confounded with electrode insertion depth and as a consequence we do not know if a

---

**It remains a puzzle how some single-channel patients can perform so well given the limited spectral information they receive.**

patient performs poorly on speech-recognition tasks because of the etiology of hearing loss or because of shallow electrode insertion. How can we isolate the effect of electrode insertion depth, or any other factor, on speech performance assuming that all other factors are held equal?

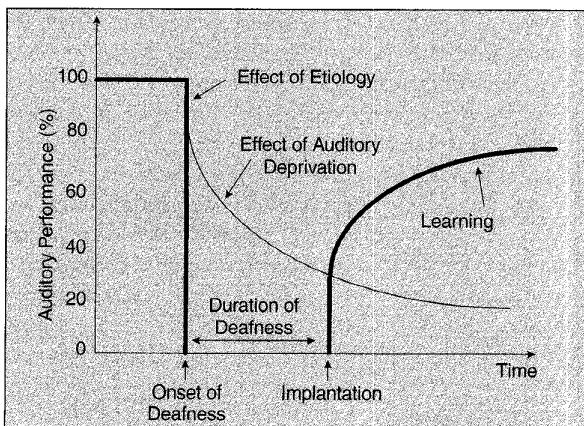
As a step towards assessing the effect of factors, such as number of channels, on auditory performance, Dorman and Loizou [36, 81] and Shannon et al. [35, 82] used acoustic simulations of cochlear implants. In these simulations, speech was processed in a manner similar to the implant processor and output either as a sum of sinusoids or as a sum of noise bands (listening demonstrations of the acoustic simulations can be accessed from our web site at <http://giles.ulr.edu/asd/cimplants>). The reconstructed speech was presented to normal-hearing listeners for identification. In the following sections, we describe some of our simulations that examined (1) the number of channels necessary for achieving high-levels of speech understanding and (2) the effect of electrode insertion depth on auditory performance.

### Number of Channels

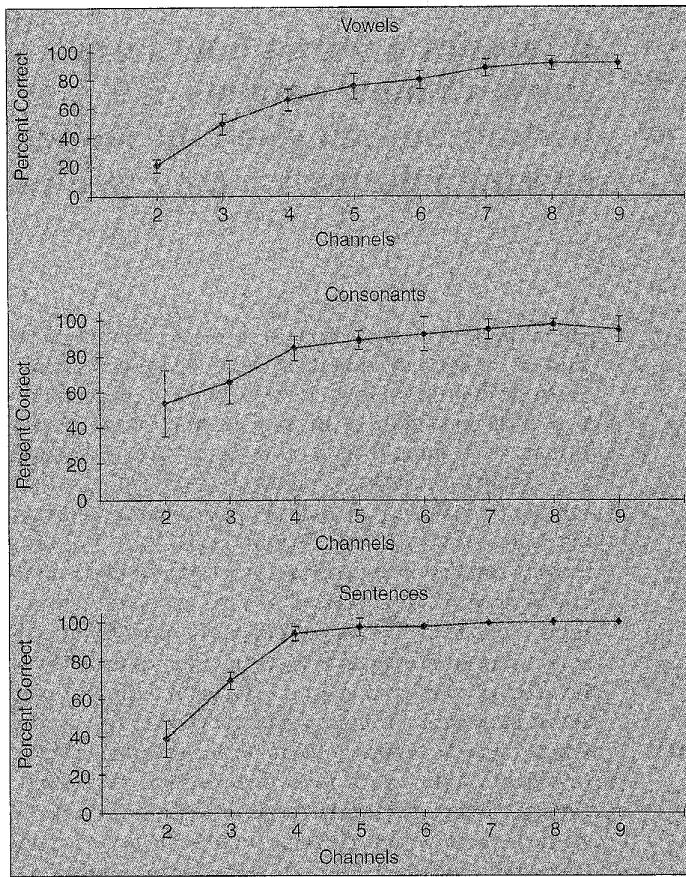
How many independent channels are needed to achieve high levels of speech understanding? It is difficult to answer this question using implant patients because of the confounding factors (e.g., number of surviving ganglion cells) that may affect their performance. For example, if a patient obtains poor auditory performance using four channels of stimulation, we do not know if it is because of the small number of channels or because there are not enough surviving ganglion cells near the stimulating electrodes. Acoustic simulations can be used to unconfound the effect of surviving ganglion cells and therefore determine how many independent channels are needed to achieve high auditory performance, assuming that all other factors are held equal.

The acoustic simulations [36] mimic the front-end processing of the implant processor and represent speech as a sum of sinusoids. More specifically, speech is reconstructed as a sum of sinusoids with time-varying amplitudes and fixed frequencies, i.e.,

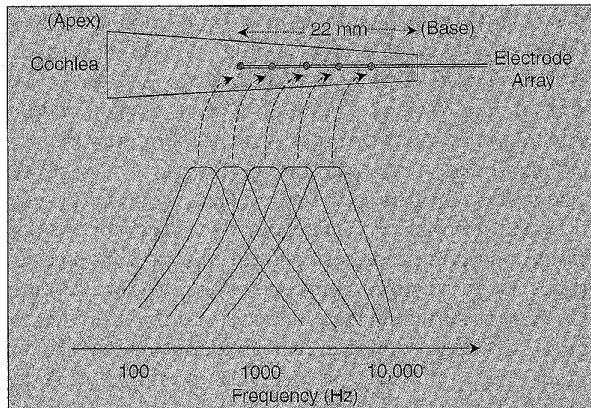
$$s^{(t)}(n) = \sum_{i=1}^L A_i^{(t)} \cos(2\pi f_i n) \quad (3)$$



▲ 35. A three-stage model of auditory performance for postlingually deafened adults (Blamey et al. [80]). The thick lines show measurable auditory performance and the thin line shows potential auditory performance.



▲ 36. Mean scores of normally hearing listeners on recognition of vowels, consonants, and sentences as a function of number of channels [36]. Error bars indicate standard deviations.



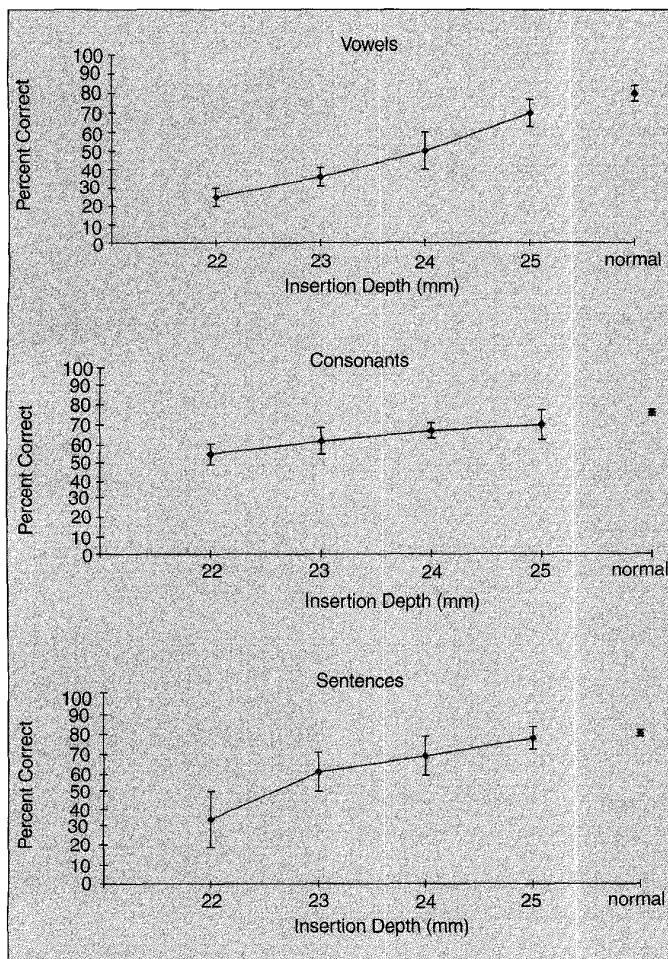
▲ 37. Diagram showing the analysis filters used in a five-channel cochlear prosthesis and a 5-electrode array (with 4 mm electrode spacing) inserted 22 mm into the cochlea. Due to shallow electrode insertion, there is a frequency mismatch between analysis frequencies and stimulating frequencies. As shown, the envelope output of the first analysis filter (centered at 418 Hz) is directed to the most apical electrode, which is located at the 831 Hz place in the cochlea. Similarly, the outputs of the other filters are directed to electrodes located higher in frequency-place than the corresponding analysis frequencies. As a result, the speech signal is up-shifted in frequency.

where  $s^{(t)}(n)$  is the synthesized speech signal at frame  $t$ ,  $A_i^{(t)}$  is the  $i$ -th amplitude at frame  $t$ ,  $L$  is the number of channels, and  $f_i$  is the center frequency of the  $i$ -th analysis filter. The amplitudes,  $A_i$ , of the sinusoids are computed in a manner similar to the CIS strategy. The speech signal is first processed through a pre-emphasis filter (low-passed below 1,200 Hz with  $-6$  dB/octave rolloff) and then bandpassed into  $L$  ( $2 \leq L \leq 9$ ) logarithmic frequency bands using sixth-order Butterworth filters. The envelope of the signal is then extracted by full-wave rectification and low-pass filtering (second-order Butterworth) with a 400 Hz cutoff frequency. The amplitudes of the sinusoids are computed by estimating the root-mean-square (rms) energy of the envelopes every 4 msec. The sinusoids are finally summed up and presented to normal-hearing listeners for identification. The results are shown in Fig. 36. As it can be seen, the number of channels needed to reach asymptotic performance depended on the test material. For the most difficult test, i.e., vowel recognition, eight channels were needed, while for the least difficult test, i.e., sentence recognition, five channels were needed. These results suggest that high levels of speech understanding could be obtained with 5-8 independent channels of stimulation.

### Electrode Insertion Depth and Frequency Up-Shifting

Electrode arrays are inserted only partially into the cochlea, typically 22-30 mm, depending on the state of the cochlea. The fact that the electrode array is not fully inserted into the cochlea creates a frequency mismatch between the analysis frequency and the stimulating frequency. Consider for example an electrode array, consisting of five electrodes, inserted 22 mm into the cochlea. The output of the first analysis filter, which is centered at 418 Hz, is directed to the most apical electrode, which resides in the 831 Hz place in the cochlea (see Fig. 37). Similarly, the outputs of the other filters are directed to electrodes that lie higher in frequency-place in the cochlea than the corresponding analysis frequencies. As a result, the speech signal is up-shifted in frequency and is therefore less intelligible. This is consistent with patients' reports that speech sounds unnatural and "high-pitched" or "Donald-duck like" when their device is first activated.

Although we would expect that speech understanding would be best in the situation in which the envelope outputs from the analysis filters are directed to the "correct" place in the cochlea, clinical data do not provide overwhelming evidence in support of this position. Some patients with shallow insertions (and therefore larger frequency mismatch between analysis and output frequencies) perform as well as patients with deep insertions (and therefore smaller frequency mismatch). This situa-



▲ 38. Percent correct recognition of vowels, consonants, and sentences as a function of simulated insertion depth [81]. The normal condition corresponds to the situation in which the analysis frequencies and output frequencies match exactly.

tion exists because of all the coexisting conditions that affect patient's performance, and therefore make it difficult to assess the effect of insertion depth alone. Acoustic simulations were used by Dorman and Loizou [81] to determine the effect of electrode insertion depth on speech understanding for a five-channel cochlear prosthesis. Different insertion depths were simulated ranging from 22 mm to 25 mm insertion. Greenwood's frequency-to-place equation [83] was used to determine the sinewave output frequencies that simulated different electrode depths. For example, to simulate the 22 mm insertion into the cochlea with 4 mm electrode spacing, sinewaves were generated with output frequencies of 831, 1,566, 2,844, 5,056, and 8,924 Hz. The corresponding sinewave amplitudes were computed as in Eq. (3) using analysis filters with center frequencies of 418, 748, 1,339, 2,396, and 4,287 Hz, respectively (Fig. 37). The results on consonant, vowel, and sentence recognition are shown in Fig. 38 for different insertion depths. As can be seen, there was a significant effect of insertion

depth for all test materials. Performance in the 22 mm and 23 mm conditions differed significantly from the normal condition (i.e., the condition in which the analysis and output frequency matched exactly) for all test materials. Performance in the 25 mm condition, however, did not differ significantly from the normal condition. These results suggest that relatively shallow insertions should result in relatively poor speech understanding assuming all other factors are held equal.

## Conclusions and Future Directions

Cochlear implants have been very successful in restoring partial hearing to profoundly deaf people. Many individuals with implants are now able to communicate and understand speech without lip-reading, and some are able to talk over the phone. Children with implants can develop spoken-language skills and attend normal schools (i.e., schools with normal-hearing children). The greatest benefits with cochlear implantation have occurred in postlingually deafened adults. Gradual, but steady, improvements in speech production and speech perception have also occurred in prelingually deafened adults or children. Auditory performance has been found to be better in patients who (1) acquired speech and language before their hearing loss (i.e., postlingually deafened) and (2) have shorter duration of deafness.

Much of the success of cochlear implants was due to the advancement of signal processing techniques developed over the years (see summary in Table 2). While this success is very encouraging, there is still a great deal to be learned about electrical stimulation of the auditory nerve, and many questions to be answered. Future research in cochlear prosthesis should:

- ▲ Continue investigating the strengths and limitations of present signal-processing strategies including CIS-type and SPEAK-type strategies. The findings of such investigations may lead to the development of signal-processing techniques capable of transmitting more information to the brain.
- ▲ Develop noise-reduction algorithms that will help implant patients better communicate in noisy environments. It would be desirable to develop signal-processing algorithms that are robust in competing noise.
- ▲ Identify factors that contribute to the variability in performance among patients [84]. Knowing these factors may help us develop signal-processing techniques that are patient specific. Patients will then be optimally fit with specific signal processors, much like people are fit with new eye glasses by an optometrist. The success of the new signal processors will ultimately narrow the gap in auditory performance between "poorly performing" and "better-performing" patients.

**Table 2. Strategies used in multichannel cochlear implants.**

Strategy	Signal Representation	Stimulation	Channels	Rate per Channel	Device
CA	Bandpassed waveforms	Analog	4	Continuous waveform	Ineraid
CA	Bandpassed waveforms	Analog	8	13,000 samples/sec	Clarion1.0
CIS	Envelope signals	Pulsatile	8	833 pps	Clarion1.0
CIS	Envelope signals	Pulsatile	8	1,515 pps	Med-El/COMBI-40
CIS	Envelope signals	Pulsatile	8	1,515 pps	Med-El/COMBI-40+
F0/F2	Second formant, voicing features	Pulsatile	1	F0 or random rate	Nucleus
F0/F1/F2	First and second formant, voicing features	Pulsatile	2	F0 or random rate	Nucleus
MPEAK	First and second formant, envelope signals	Pulsatile	4	F0 or random rate	Nucleus
SMSP	Envelope signals, spectral maxima	Pulsatile	6	250 pps	Nucleus
SPEAK	Envelope signals, spectral maxima	Pulsatile	5-10	180-300 pps	Nucleus

▲ Develop preoperative procedures that can predict how well a patient will perform with a cochlear implant.

▲ Continue investigating the effects of electrical stimulation on encoding of speech in the auditory nerve. Such investigations may help us design better electrodes as well as develop new signal-processing strategies.

▲ Design electrode arrays capable of providing a high degree of specificity. Such electrode arrays will provide channel selectivity, which is now considered to be one of the limiting factors in performance.

▲ Investigate the effect of high-rate pulsatile stimulation (> 3000 pulses/sec) on speech perception as well as on music appreciation using more than eight channels.

It is hoped that future research in cochlear prostheses will mature to a level that will enable all implant patients to be "better-performing" patients.

## Acknowledgments

The author would like to thank Blake Wilson, Michael Dorman, and Mary Barker for providing valuable suggestions on earlier drafts of this manuscript. This work was supported in part by grant No. 98-B-07 from the Arkansas Science and Technology Authority.

*Philipos Loizou* is an assistant professor at the University of Arkansas at Little Rock (e-mail: loizou@ualr.edu).

## References

1. B. Wilson, "Signal processing," in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), pp. 35-86, Singular Publishing Group, Inc, 1993.
2. G. Loeb, "Cochlear prosthetics," *Annual Review in Neuroscience*, vol. 13, pp. 357-371, 1990.
3. J. Millar, Y. Tong, and G. Clark, "Speech processing for cochlear implant prostheses," *Journal of Speech and Hearing Research*, vol. 27, pp. 280-296, 1984.
4. C. Parkins and S. Anderson (eds.), *Cochlear prostheses: An international symposium*. New York: New York Academy of Sciences, 1983.
5. G. Fant, *Acoustic Theory of Speech Production*. The Hague, Netherlands: Mouton, 1970.
6. G. Borden, K. Harris, and L. Raphael, *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Baltimore, Maryland: Williams and Wilkins, 1994.
7. F. Cooper, P. Delattre, A. Liberman, J. Borst, and L. Gerstman, "Some experiments on the perception of synthetic speech sounds," *Journal of the Acoustical Society of America*, vol. 24, no. 6, pp. 597-606, November 1952.
8. W. Yost and D. Nielsen, *Fundamentals of Hearing: An Introduction*. New York: Holt, Rinehart and Winston, 1985.
9. R. Hinojosa and M. Marion, "Histopathology of profound sensorineural deafness," *Annals of New York Academy of Sciences*, vol. 405, pp. 459-484, 1983.
10. I. Hochmair-Desoyer, E. Hochmair, and K. Burian, "Design and fabrication of multiwire scala tympani electrodes," *Annals of New York Academy of Sciences*, vol. 405, pp. 173-182, 1983.

11. G. Clark, R. Shepherd, J. Patrick, R. Black, and Y. Tong, "Design and fabrication of the banded electrode array," *Annals of New York Academy of Sciences*, vol. 405, pp. 191-201, 1983.
12. C. van den Honert and P. Stypulkowski, "Single fiber mapping of spatial excitation patterns in the electrically stimulated nerve," *Hearing Research*, vol. 29, pp. 195-206, 1987.
13. M. Merzenich and M. White, "Cochlear implant - the interface problem," in *Functional electrical stimulation: Applications in neural prostheses* (F. Hambrecht and J. Reswick, eds.), Marcel Dekker, 1977, pp. 321-340.
14. B. Wilson, D. Lawson, and M. Zerbi, "Advances in coding strategies for cochlear implants," *Advances in Otolaryngology - Head and Neck Surgery*, vol. 9, pp. 105-129, 1995.
15. A. Boothroyd, "Profound deafness," in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), Singular Publishing Group, Inc, 1993, pp. 1-34.
16. E. Owens, D. Kessler, and E. Schubert, "The minimal auditory capabilities (MAC) battery," *Hearing Aid Journal*, vol. 34, pp. 9-34, 1981.
17. R. Tyler, J. Preece, and N. Tye-Murray, "The Iowa phoneme and sentence tests," Tech. Rep., Department of Otolaryngology-Head and Neck Surgery, University of Iowa, 1986.
18. R. Tyler, "Speech perception by children," in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), Singular Publishing Group, Inc, 1993, pp. 191-256.
19. W. House, "A personal perspective on cochlear implants," in *Cochlear Implants* (R. Schindler and M. Merzenich, eds.), New York: Raven Press, 1985, pp. 13-16.
20. W. House and J. Urban, "Long term results of electrode implantation and electronic stimulation of the cochlea in man," *Annals of Otology, Rhinology and Laryngology*, vol. 82, pp. 504-517, 1973.
21. W. House and K. Berliner, "Cochlear implants: Progress and perspectives," *Annals of Otology, Rhinology and Laryngology*, vol. (Suppl. 91), pp. 1-124, 1982.
22. B. Edgerton and J. Brimacombe, "Effects of signal processing by the House-3M cochlear implant on consonant perception," *Acta Otolaryngologica*, pp. 115-123, 1984.
23. B. Gantz, R. Tyler, J. Knutson, G. Woodworth, P. Abbas, et al., "Evaluation of five different cochlear implant designs: Audiologic assessment and predictors of performance," *Laryngoscope*, vol. 98, pp. 1100-1106, 1988.
24. S. Rosen, J. Walliker, J. Brimacombe, and B. Edgerton, "Prosodic and segmental aspects of speech perception with the House/3M single-channel implant," *Journal of Speech and Hearing Research*, vol. 32, pp. 93-111, 1989.
25. J. Danhauer, F. Ghadiali, D. Eskwitt, and L. Mendel, "Performance of 3M/House cochlear implant users on tests of speech perception," *Journal of the American Academy of Audiology*, vol. 1, pp. 236-239, 1990.
26. I. Hochmair-Desoyer and E. Hochmair, "Percepts elicited by different speech-coding strategies," *Annals of New York Academy of Sciences*, vol. 405, pp. 268-279, 1983.
27. R. Shannon, "Multichannel electrical stimulation of the auditory nerve in man: I. Basic psychophysics," *Hearing Research*, vol. 11, pp. 157-189, 1983.
28. R. Shannon, "Psychophysics," in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), Singular Publishing Group, Inc, 1993, pp. 357-388.
29. R. Tyler, "Open-set recognition with the 3M/Vienna single-channel cochlear implant," *Archives of Otolaryngology, Head and Neck Surgery*, vol. 114, pp. 1123-1126, 1988.
30. I. Hochmair-Desoyer, E. Hochmair, and H. Stiglbrunner, "Psychoacoustic temporal processing and speech understanding in cochlear implant patients," in *Cochlear Implants* (R. Schindler and M. Merzenich, eds.), New York: Raven Press, 1985, pp. 291-304.
31. P. Abbas, "Electrophysiology," in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), Singular Publishing Group, Inc, 1993, pp. 317-356.
32. E. Burns and N. Viemester, "Nonspectral pitch," *Journal of the Acoustical Society of America*, vol. 60, pp. 863-869, 1976.
33. S. Stevens and J. Volkmann, "The relation of pitch to frequency: A revised scale," *The American Journal of Psychology*, vol. LIII, no. 3, pp. 329-353, July 1940.
34. M. Dorman, "Speech perception by adults," in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), Singular Publishing Group, Inc, 1993, pp. 145-190.
35. R. Shannon, F. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, "Speech recognition with primarily temporal cues," *Science*, vol. 270, pp. 303-304, 1995.
36. M. Dorman, P. Loizou, and D. Rainey, "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs," *Journal of the Acoustical Society of America*, vol. 102, pp. 2403-2411, 1997.
37. D. Eddington, "Speech discrimination in deaf subjects with cochlear implants," *Journal of the Acoustical Society of America*, vol. 68, no. 3, pp. 885-891, 1980.
38. M. Merzenich, S. Rebscher, G. Loeb, C. Byers, and R. Schindler, "The UCSF cochlear implant project: State of development," *Advances in Audiology*, vol. 2, pp. 119-144, 1984.
39. M. Dorman, M. Hannley, K. Dankowski, L. Smith, and G. McCandless, "Word recognition by 50 patients fitted with the Symbion multichannel cochlear implant," *Ear and Hearing*, vol. 10, pp. 44-49, 1989.
40. M. White, M. Merzenich, and J. Gardi, "Multichannel cochlear implants: Channel interactions and processor design," *Archives of Otolaryngology*, vol. 110, pp. 493-501, 1984.
41. B. Wilson, C. Finley, D. Lawson, R. Wolford, D. Eddington, and W. Rabinowitz, "Better speech recognition with cochlear implants," *Nature*, vol. 352, pp. 236-238, July 1991.
42. C. Bocx, M. Pelizzzone, and P. Montandon, "Improvements in speech recognition with the CIS strategy for the Ineraid multichannel intracochlear implant," in *Advances in cochlear implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, 1994, pp. 136-140.
43. M. Dorman and P. Loizou, "Changes in speech intelligibility as a function of time and signal processing strategy for an Ineraid patient fitted with Continuous Interleaved Sampling (CIS) processors," *Ear and Hearing*, vol. 18, pp. 147-155, 1997.
44. M. Dorman and P. Loizou, "Mechanisms of vowel recognition for Ineraid patients fit with continuous interleaved sampling processors," *Journal of the Acoustical Society of America*, vol. 102, pp. 581-587, 1997.
45. D. Eddington, W. Dobelle, D. Brachman, M. Mladovsky, and J. Parkin, "Auditory prosthesis research using multiple intracochlear stimulation in man," *Annals of Otology, Rhinology and Laryngology*, vol. 87 (Suppl. 53), pp. 1-39, 1978.
46. F. Zeng and R. Shannon, "Loudness balance between acoustic and electric stimulation," *Hearing Research*, vol. 60, pp. 231-235, 1992.
47. G. Clark, "The University of Melbourne-Nucleus multi-electrode cochlear implant," *Advances in Oto-Rhino-Laryngology*, vol. 38, pp. 1-189, 1987.
48. P. Seligman, J. Patrick, Y. Tong, G. Clark, R. Dowell, and P. Crosby, "A signal processor for a multiple-electrode hearing prosthesis," *Acta Otolaryngologica*, pp. 135-139, (Suppl. 411), 1984.
49. R. Dowell, D. Mecklenburg, and G. Clark, "Speech recognition for 40 patients receiving multichannel cochlear implants," *Archives of Otolaryngology, Head and Neck Surgery*, vol. 112, pp. 1054-1059, 1986.
50. P. Blamey, R. Dowell, and G. Clark, "Acoustic parameters measured by a formant-estimating speech processor for a multiple-channel cochlear implant," *Journal of the Acoustical Society of America*, vol. 82, pp. 38-47, 1987.
51. R. Dowell, P. Seligman, P. Blamey, and G. Clark, "Evaluation of a two-formant speech processing strategy for a multichannel cochlear prosthesis," *Annals of Otology, Rhinology and Laryngology*, vol. 96 (Suppl. 128), pp. 132-134, 1987.

52. N. Tye-Murray, M. Lowder, and R. Tyler, "Comparison of the F0/F2 and F0/F1/F2 processing strategies for the Cochlear Corporation cochlear implant," *Ear and Hearing*, vol. 11, pp. 195-200, 1990.
53. J. Patrick, P. Seligman, D. Money, and J. Kuzma, "Engineering," in *Cochlear prostheses* (G. Clark, Y. Tong, and J. Patrick, eds.), Edinburgh: Churchill Livingstone, 1990, pp. 99-124.
54. J. Patrick and G. Clark, "The Nucleus 22-channel cochlear implant system," *Ear and Hearing*, vol. 12, pp. 3-9, (Suppl. 1), 1991.
55. E. Wallenberger and R. Battmer, "Comparative speech recognition results in eight subjects using two different coding strategies with the Nucleus 22 channel cochlear implant," *British Journal of Audiology*, vol. 25, pp. 371-380, 1991.
56. R. Dowell, P. Dawson, S. Dettman, R. Shepherd, L. Whitford, et al., "Multichannel cochlear implantation in children: A summary of current work at the University of Melbourne," *American Journal of Otology*, vol. 12, pp. 137-143, (Suppl. 1), 1991.
57. M. Skinner, L. Holden, T. Holden, R. Dowell, P. Seligman, et al., "Performance of postlinguistically deaf adults with the Wearable Speech Processor (WSP III) and Mini Speech Processor (MSP) of the Nucleus multi-electrode cochlear implant," *Ear and Hearing*, vol. 12, pp. 3-22, 1991.
58. H. McDermott, C. McKay, and A. Vandali, "A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant," *Journal of the Acoustical Society of America*, vol. 91, pp. 3367-3371, 1992.
59. P. Seligman and H. McDermott, "Architecture of the Spectra 22 speech processor," *Annals of Otology, Rhinology and Laryngology*, pp. 139-141, (Suppl. 166), 1995.
60. M. Skinner, G. Clark, L. Whitford, P. Seligman, S. Staller, et al., "Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system," *American Journal of Otology*, vol. 15, pp. 15-27, (Suppl. 2), 1994.
61. R. Schindler and D. Kessler, "Preliminary results with the Clarion cochlear implant," *Laryngoscope*, vol. 102, pp. 1006-1013, 1992.
62. D. Kessler and R. Schindler, "Progress with a multi-strategy cochlear implant system: The Clarion," in *Advances in cochlear implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, 1994, pp. 354-362.
63. G. Loeb and D. Kessler, "Speech recognition performance over time with the Clarion cochlear prosthesis," *Annals of Otology, Rhinology and Laryngology*, vol. 104, pp. 290-292, (Suppl. 166), 1995.
64. R. Tyler, B. Gantz, G. Woodworth, A. Parkinson, M. Lowder, and L. Schum, "Initial independent results with the Clarion cochlear implant," *Ear and Hearing*, vol. 17, pp. 528-536, 1996.
65. C. Zierhofer, O. Peter, S. Bril, P. Pohl, I. Hochmair-Desoyer, and E. Hochmair, "A multichannel cochlear implant system for high-rate pulsatile stimulation strategies," in *Advances in cochlear implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, 1994, pp. 204-207.
66. J. Helms, J. Muller, F. Schon et al., "Evaluation of performance with the COMBI 40 cochlear implant in adults: A multicentric clinical study," *Oto-Rhino-Laryngology*, vol. 59, pp. 23-35, 1997.
67. C. Zierhofer, I. Hochmair-Desoyer, and E. Hochmair, "Electronic design of a cochlear implant for multichannel high-rate pulsatile stimulation strategies," *IEEE Transactions on Rehabilitation Engineering*, vol. 3, no. 1, pp. 112-116, March 1995.
68. P. Loizou, M. Dorman, and V. Powell, "The recognition of vowels produced by men, women, boys and girls by cochlear implant patients using a 6 channel CIS processor," *Journal of the Acoustical Society of America*, vol. 103, pp. 1141-1149, 1998.
69. M. Dorman and P. Loizou, "Improving consonant intelligibility for Deaf patients fit with Continuous Interleaved Sampling (CIS) processors by enhancing contrast among channel outputs," *Ear and Hearing*, vol. 17, pp. 308-313, 1996.
70. E. Tobey, "Speech production," in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), Singular Publishing Group, Inc, 1993, pp. 257-316.
71. M. Osberger, A. Robbins, S. Todd, A. Riley, and R. Miyamoto, "Speech production skills of children with multichannel cochlear implants," in *Advances in cochlear implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, 1994, pp. 503-507.
72. S. Pancamo and E. Tobey, "Effects of multichannel cochlear implant upon sound production in children," *Proc. of the 2nd Annu. Southeastern Allied Health Research Symp.*, pp. 319-330, 1989.
73. R. Miyamoto, M. Osberger, S. Todd, and A. Robbins, "Speech perception skills of children with multichannel cochlear implants," in *Advances in cochlear implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, 1994, pp. 498-502.
74. S. Staller, R. Dowell, A. Beiter, and J. Brimacombe, "Perceptual abilities of children with the Nucleus 22-channel cochlear implant," *Ear and Hearing*, vol. 12, pp. 34S-47S, 1991.
75. N. Erber and C. Alencewicz, "Audiologic evaluation of deaf children," *Journal of Speech and Hearing Disorders*, vol. 41, pp. 256-267, 1976.
76. B. Gantz, R. Tyler, N. Tye-Murray, and H. Fryauf-Bertschy, "Long term results of multichannel cochlear implants in congenitally deaf children," in *Advances in Cochlear Implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, 1994, pp. 528-533.
77. D. Shipp and J. Nedzelski, "Prognostic value of round window psychophysical measurements with adult cochlear implant candidates," in *Advances in Cochlear Implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, 1994, pp. 79-81.
78. B. Gantz, G. Woodworth, P. Abbas, J. Knutson, and R. Tyler, "Multivariate predictors of audiological success with multichannel cochlear implants," *Annals of Otology, Rhinology and Laryngology*, vol. 102, pp. 909-916, 1993.
79. A. Summerfield and D. Marshall, "Preoperative predictors of outcomes from cochlear implantation in adults: Performance and quality of life," *Annals of Otology, Rhinology and Laryngology*, pp. 105-108, (Suppl. 166), 1995.
80. P. Blamey, P. Arndt, F. Bergeron, G. Bredberg, J. Brimacombe, et al., "Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants," *Audiology and Neuro-Otology*, vol. 1, pp. 293-306, 1996.
81. M. Dorman, P. Loizou, and D. Rainey, "Simulating the effect of cochlear implant electrode insertion-depth on speech understanding," *Journal of the Acoustical Society of America*, vol. 102, pp. 2993-2996, 1997.
82. R. Shannon, F. Zeng, and J. Wygonski, "Speech recognition with altered spectral distribution of envelope cues," *Journal of the Acoustical Society of America*, vol. 100, Pt.2, p. 2692, 1996.
83. D. Greenwood, "A cochlear frequency-position function for several species - 29 years later," *Journal of the Acoustical Society of America*, vol. 87, pp. 2592-2605, 1990.
84. "Cochlear implants in adults and children," *NIH Consensus Statement*, vol. 13, pp. 1-30, May 1995.
85. B. Wilson, C. Finley, D. Lawson, and R. Wolford, "Speech processors for cochlear prostheses," *Proceedings of IEEE*, vol. 76, pp. 1143-1154, September 1988.