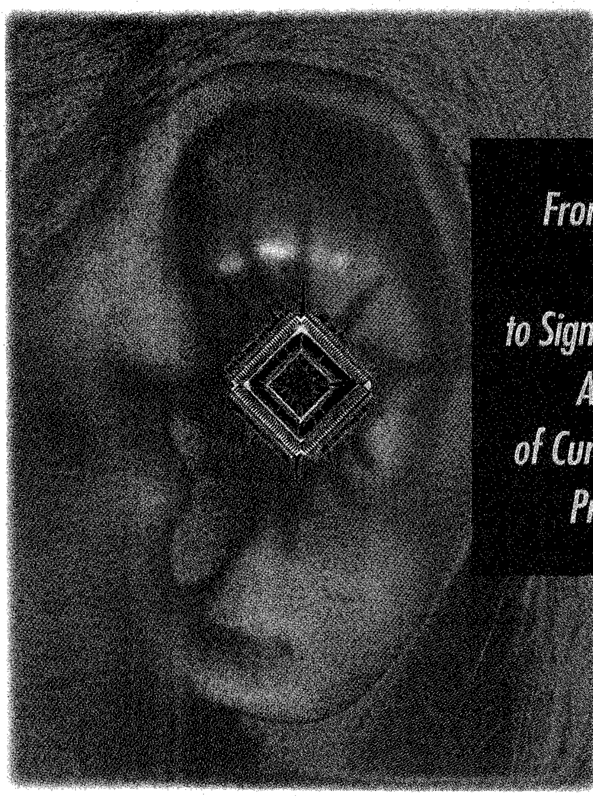
Philipos C. Loizou
Department of Applied Science,
University of Arkansas at Little Rock

Introduction to Cochlear Implants

From Electrode Design to Signal Processing, A Review of Current Implant Processors



For 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 the *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. This article is a tutorial review of cochlear implants.

The Elements of Hearing

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.

According to the source-filter model of speech production [1, 2], 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. If the input excitation is periodic, then voiced sounds like /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* (F0).

The vocal tract shape, defined in terms of tongue, velum, lip and jaw position, acts as a "filter" of 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, 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 F1, F2, and F3) contain sufficient information for the recognition of vowels as well as other voiced sounds. Formant

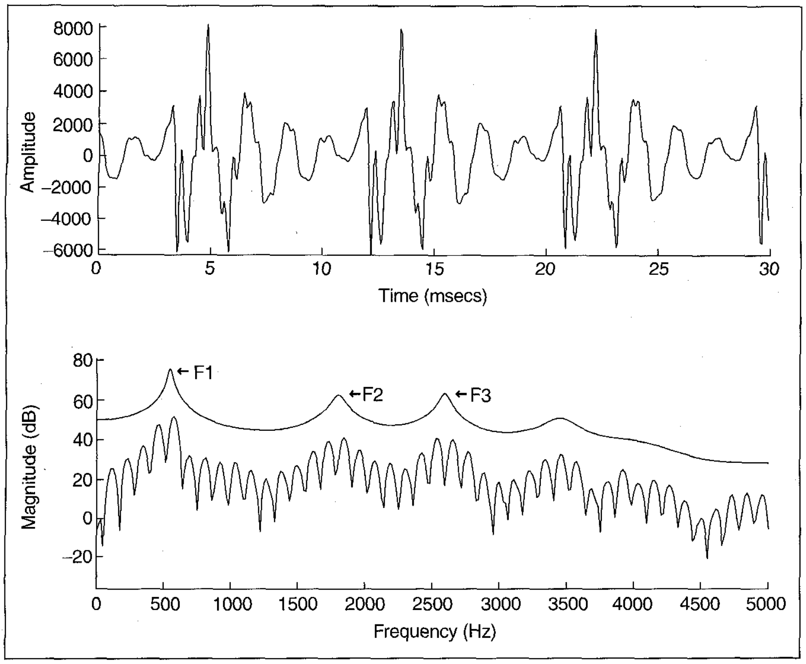
Editor's Note: This article is adapted from a version that appeared in the September 1998 issue of IEEE Signal Processing Magazine, *(vol. 15, no. 5, pp. 101-130), which included more in-depth information on signal-processing techniques.*

 0739-5175/99/$10.00©1999 IEEE

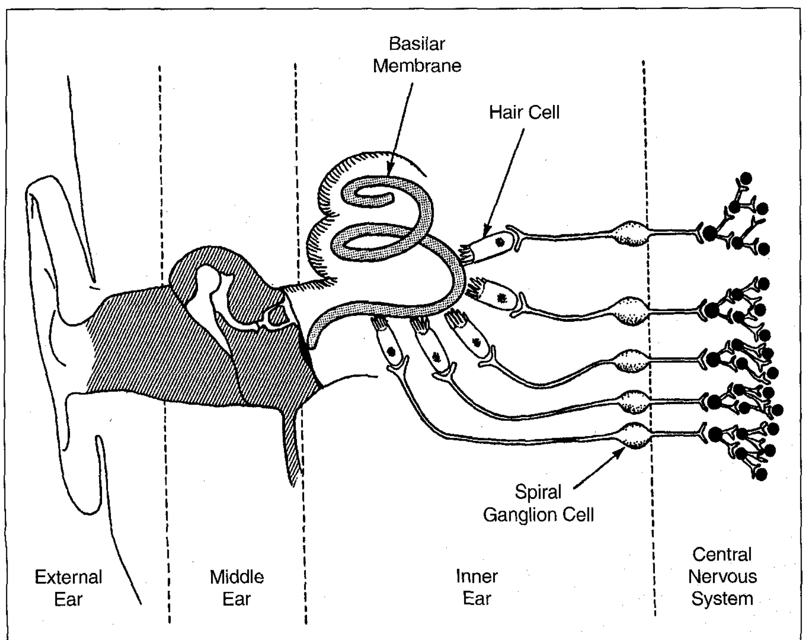
movements have also been found to be extremely important for the perception of unvoiced sounds (i.e., consonants) (e.g., [2, 3]). In summary, the formants carry some information about the speech signal, and thus some of the early cochlear implants conveyed formant information to the electrodes.

Normal Hearing

Figure 2 shows a simplified diagram of the human ear, consisting of the outer, middle, and inner ear [4]. Sound undergoes a series of transformations as it travels through the outer ear, middle ear, inner ear, auditory nerve, and into the brain. The outer ear picks up acoustic pressure waves, which are converted to mechanical vibrations by a series of small bones in the middle ear. In the inner ear, the cochlea, a snail-shaped cavity filled with fluid, transforms the mechanical vibrations to vibrations in fluid. Pressure variations within the fluid of the cochlea lead to displacements of a flexible membrane, called the basilar membrane. These displacements contain information about the frequency of the acoustic signal. Attached to the basilar membrane are hair cells, which are bent according to the displacements of the basilar membrane. The bending of the hairs releases 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.



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 F1, F2, and F3.



2. A diagram (not to scale) of the human ear. (Reprinted with permission from [32]).

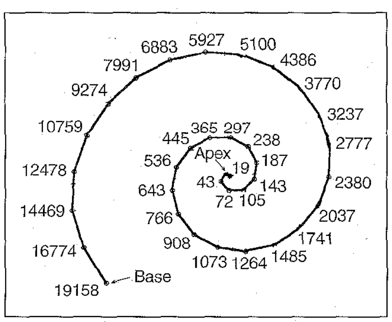
Deafness

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, resulting in hearing impairment. 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. If a large number of hair cells or auditory neurons are damaged, then the condition is called profound deafness. Research [5] 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 how to stimulate (electrically) auditory neurons so that meaningful information about speech is conveyed to the brain. For example, information about the amplitude and the frequency of the acoustic signal should be conveyed.

Encoding Frequency

The question then arises: "How does the auditory system encode frequencies?" The pioneering work of Georg von 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 the 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 the largest displacement at the base (near the stapes) of 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 thereby acts as a spectrum analyzer, decomposing complex sounds into their frequency components.

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 this theory, the auditory nerve fibers fire at rates proportional to the period of the input signal, up to frequencies of 5000 Hz. At low frequencies, individual nerve fibers fire at each cycle of the stimulus; i.e., they are "phase locked" with the stimulus. High frequencies are indicated by the organized firing of groups of nerve fibers.



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.

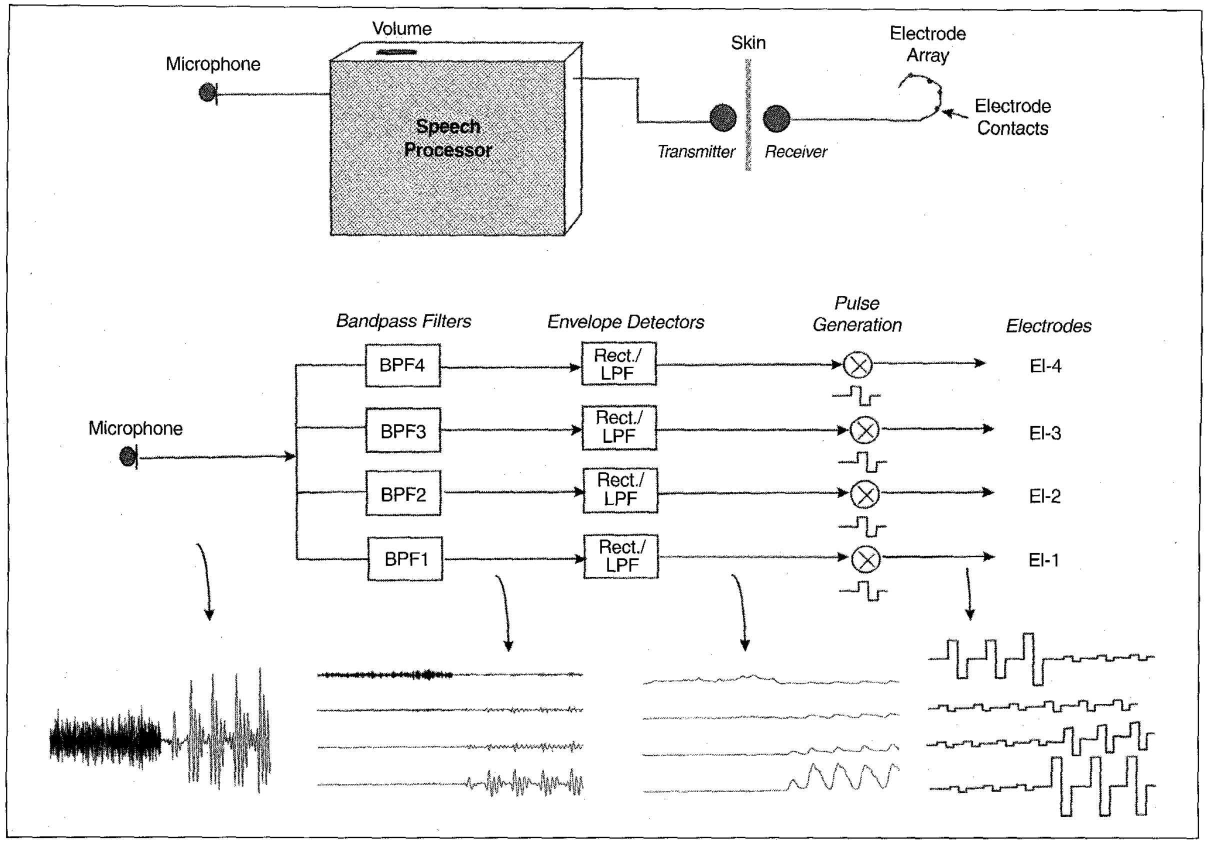
The design of electrodes for cochlear prosthesis has been the focus of research for over two decades.

Cochlear Implants

Several cochlear implant devices have been developed over the years [6]. 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 implanted into the cochlea (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, thereby exploiting the place mechanism for coding frequencies. Different electrodes are stimulated, depending on the frequency of the signal. Electrodes near the base of the cochlea are stimulated with high-frequency signals, while electrodes near the apex are stimulated with low-frequency signals. The signal processor is responsible for breaking the input signal into different frequency bands or channels and delivering the filtered signals to the appropriate electrodes. The main function of the signal processor is to decompose the input signal into its frequency components, much like a healthy cochlea. The designers of cochlear prosthesis are faced with the challenge of developing signal-processing techniques that mimic the function of a healthy cochlea.

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

Figure 4 shows, as an example, 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, which divide the acoustic signal into four channels. Current pulses are generated with amplitudes proportional to the energy in each channel, and transmitted to the four electrodes through a radio-frequency link. The relative amplitudes of the current pulses delivered to the electrodes reflect the spectral content of the input signal (Fig. 4). For instance, if the speech signal contains mostly high-frequency information (e.g., /s/), then the pulse amplitude of channel 4 will be large relative to the pulse amplitudes of channels 1-3. Similarly, if the speech signal contains mostly low-frequency information (e.g., vowel /a/) then the pulse amplitude of channels 1 and 2 will be large relative to the amplitudes of channels 3 and 4 (Fig. 4).



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." 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 lowpass 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.

Implant Characteristics

Figure 4 showed one type of cochlear implant that is being used. Other types of implant devices have been developed over the years [6]. These devices differ in the following characteristics:

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

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

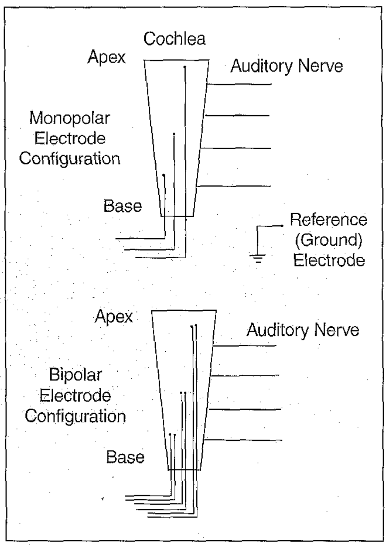
Electrode Design

The design of electrodes for cochlear prosthesis has been the focus of research for over two decades [7, 8]. Some of the issues associated with electrode design are (1) placement, (2) number and spacing of contacts, (3) orientation with respect to the excitable tissue, and (4) 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 them in close proximity with auditory neurons that lie along the length of the cochlea. This electrode placement is preferred because it preserves the "place" mechanism of the normal cochlea for coding frequencies. That is, auditory neurons that are "tuned" for high frequencies are stimulated whenever the electrodes near the base are stimulated, whereas auditory neurons that are "tuned" for low frequencies are stimulated whenever the electrodes near the apex are stimulated. In most cases, the electrode arrays can be inserted in the scala tympani to depths of 22-30 mm within the cochlea.

The number of electrodes as well as the spacing between them affects the place resolution for coding frequencies. In principle, the larger the number of electrodes, the finer the place resolution for coding. Frequency coding is constrained, however, by two inherent factors: (1) number of surviving auditory neurons that can be stimulated at a particular site in the cochlea, and (2) spread of excitation associated with electrical stimulation. Unfortunately, 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 surviving 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 along the cochlea. So, using a large number of electrodes will not necessarily result in better performance, because frequency coding is constrained by the number of surviving auditory neurons that can be stimulated.

In addition, frequency coding is constrained by the spread of excitation caused by electrical stimulation. Electric current injected into the cochlea tends to spread out symmetrically from the source. As a result, the current does not stimulate just a single (isolated) site of auditory neurons, but several. Such a spread in excitation is most prominent in the monopolar electrode configuration. In this configuration, the active electrode is located far from the reference electrode, which acts as a ground for all electrodes (see Fig. 5). The spread of excitation can be constrained, to a degree, by using a bipolar electrode configuration. In this configuration, the active and the reference (ground) electrodes are placed close to each other (Fig. 5). Bipolar electrodes have been shown to produce a more localized stimulation than monopolar [9, 10]. Although the patterns of electrical stimulation produced by these two configurations are different, it is still not clear which will result in better performance for a particular patient.



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.

Currently, some implant devices employ monopolar electrodes, other devices employ bipolar electrodes, and yet other devices provide both types. Table 1 lists some current implant devices and their characteristics. The Ineraid (also called Symbion) (Smith and Nephew, Inc., UK) device uses six electrodes spaced 4 mm apart. Only the four most apical electrodes are used in monopolar configuration. The Nucleus device (Cochlear Corporation, Australia) uses 22 electrodes spaced 0.75 mm apart. Electrodes that are 1.5 mm apart are used as bipolar pairs. The Clarion device (Advanced Bionics, California, USA) provides both monopolar and bipolar configurations, with eight electrodes (spaced 2.4 mm apart) in a monopolar configuration. The Med-El device (Med-El Corporation, Austria) also uses eight electrodes (spaced 2.8 mm apart) in a monopolar configuration.

Type of Stimulation

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

The cochlear implant is based on the premise that there are sufficient auditory nerve fibers remaining for stimulation in the vicinity of the electrodes.

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. Pulse rate (i.e., the rate at which the pulses are delivered to the electrodes) has been found to affect speech-recognition performance [11]; high pulse rates tend to yield better performance than low pulse rates.

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

The speech-perception abilities of children with implants improve steadily over time.

Transmission Link

There are currently two ways of transmitting the signals from the external processor to the implanted electrodes: either through a transcutaneous connection, or through a percutaneous connection (see Fig. 6).

For the transcutaneous system, an external transmitter encodes the stimulus information for radio-frequency transmission from an external coil to an implanted coil. The internal receiver decodes the signal and delivers the stimuli to the electrodes (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. Its disadvantage is that the implanted electronics (i.e., the receiver circuitry) may fail and would require surgery for replacement. Another disadvantage is that the transcutaneous connector contains magnetic materials, which are incompatible with MRI scanners. Most cochlear implant devices (e.g., Nucleus, Clarion, Med-El) today use transcutaneous connections.

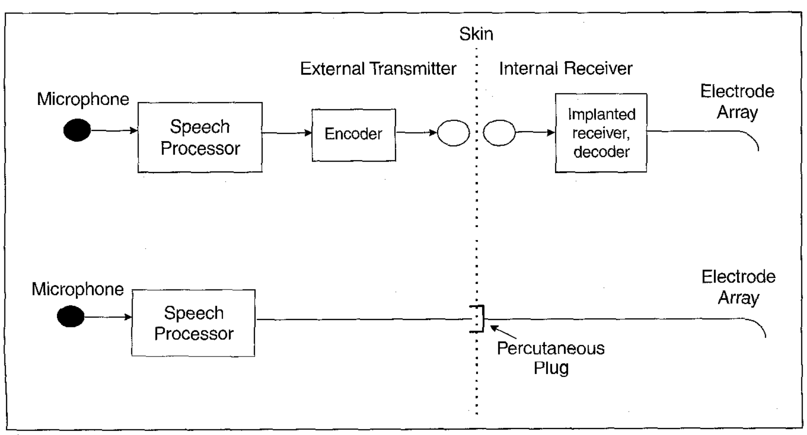
The percutaneous system transmits the stimuli to the electrodes directly through plug connections (Fig. 6). In this system, there are no implanted electronics other than the electrodes. The major advantage of this system is flexibility and signal transparency. The signal transmission is in no way constrained by the implanted receiver circuitry. It is therefore ideal for research purposes such as investigating new signal-processing techniques. Currently, only the Ineraid device 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 original speech signal to electrical stimuli. Several signal-processing techniques have been developed over the past 25 years [6]. Some of these techniques were aimed at preserving waveform information, others were aimed at preserving envelope information, and yet others were aimed at preserving spectral features (e.g., formants). A more detailed discussion on these signal-processing techniques will be found in a forthcoming article.

Who Can be Implanted?

Not all people with hearing impairments are candidates for cochlear implantation. Certain audiological criteria need to be met. First, the hearing loss has to be severe or profound and it must be bilateral. Profound deafness [12] 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, 1000, and 2000 Hz, expressed in dB with reference to normal thresholds. Second, the candidate has to obtain a sentence-recognition score of 30% or less under best aided conditions. Children age 2 years or older with profound (> 90 dB HL) sensorineural loss in both ears are also candidates for cochlear implantation.

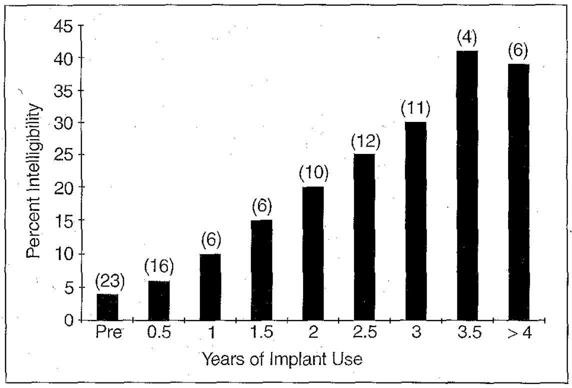


6. Diagram showing two different ways of transmitting electrical stimuli to the electrode array. The top panel shows a transcutaneous (radio-frequency link) connection and the bottom panel shows a percutaneous (direct) connection.

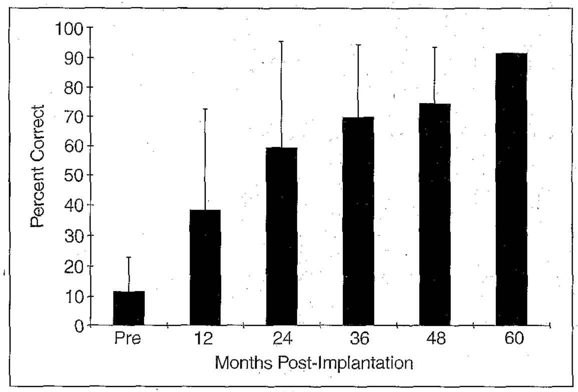
Evaluating Performance

Once a patient has been fit with a cochlear implant, how do we evaluate his/her ability to identify or recognize speech? A patient's speech-perception ability is 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 example, a patient might only perceive 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," and "who'd," which only differ in the middle segment (i.e., the vowel). Vowel and consonant tests are aimed at assessing the patient's ability in resolving spectral and temporal information. The most difficult test, by far, is the recognition of monosyllabic words. One such test, the NU-6 word lists, was developed by Northwestern University and consists of lists of 50 monosyllable words [13]. 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 [14], which consists of sentences, vowels, and consonants recorded on a laserdisc in audio, visual, and audio-visual format.

Other tests are used to evaluate the speech-perception abilities of children.



7. 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. [17]). Numbers in parenthesis indicate the number of children



8. Speech perception scores of prelingually deafened children (wearing the Nucleus implant) on word recognition (MTS test [15]) as a function of number of months of implant use (Miyamoto, et al. [19]).

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 of various tests developed to evaluate the speech-perception abilities of children can be found in [15].

Commercially Available 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 and the Clarion. There is also a cochlear implant processor, manufactured by Med-El Corporation, Austria, that is currently in clinical trials in the United States. A more detailed description on these processors will be given in a forthcoming article.

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 when 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 [15, 16]). The implant may help a child in two important aspects of 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

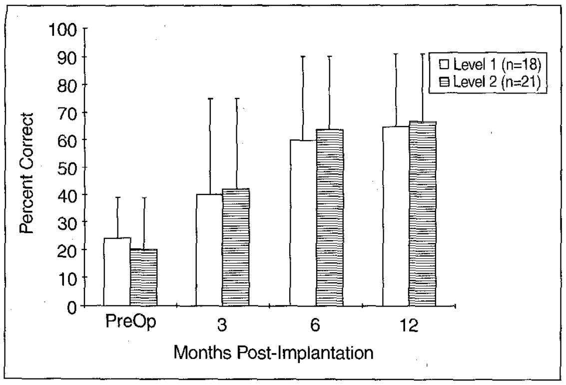
The ability to speak is closely related to the ability to hear. If the child is not able to hear, then the child will have difficulty learning how to speak correctly. Hearing provides feedback that is used by the child to correct or improve his/her speech-production skills [16]. Auditory feedback is therefore very important for learning how to speak, and cochlear implants can provide that. Research (e.g., [16-18]) has shown that the intelligibility of speech produced by children with cochlear implants improves over time. Osberger, et al. [17] 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 produced 10 sentences that 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 (Fig. 7) show that intelligibility improves gradually over time. The largest improvements 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 after 2.5 years of implant use was found to be higher than that 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

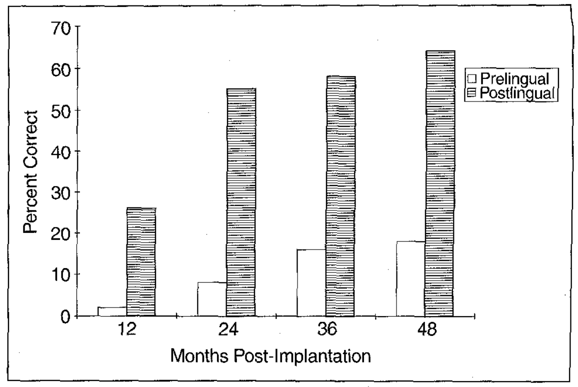
Research has also shown that the speech-perception abilities of children with implants improve steadily over time (e.g., [15, 19, 20]). Figure 8 shows a longitudinal study [19] on the perception abilities of 39 prelingually deafened children using the Nucleus implant. The children were tested on the Monosyllable-Trochee-Spondee (MTS) test [21], which uses 12 pictures of nouns. In this test, the children were asked to point to the picture corresponding to the word they heard. As can be seen in Fig. 8, the mean scores improved over time. Similar improvements were also found with children wearing the Clarion implant (Fig. 9). These results demonstrate a steady improvement in speech-recognition performance for prelingually deafened children over a three- to four-year period of implant use. In contrast, postlingually deafened children (i.e., 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 implants [22]. In addition, postlingually deafened children have been found to perform better on tests of open-set speech understanding as compared to prelingually deafened children (Fig. 10).

Factors Affecting the Performance of Cochlear Implant Patients

There is great variability in the speech-recognition performance of cochlear implant patients. For a given type of implant, auditory performance may



9. Performance of children with the Clarion implant on monosyllabic word (ESP test [15]) 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.



10. Comparison in performance between prelingually deafened and postlingually deafened children on open-set word recognition (Gantz, et al. [22]). The postlingually deafened children obtained significantly higher performance than the prelingually deafened children.

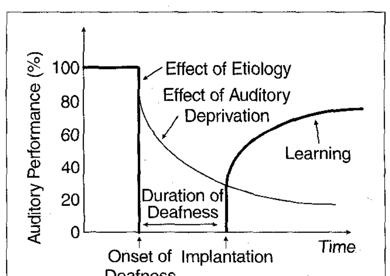
vary from zero to nearly 100% correct. Auditory performance is defined here as the ability to discriminate, detect, identify, or recognize speech (a typical measure of auditory performance is the percent-correct score on open-set speech recognition tests). The factors responsible for such variability in auditory performance have been the focus of research for many years [23-26]. 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 and adults with postlingual deafness perform better than those with prelingual or congenital deafness.
- **Age at implantation**
  Prelingually deafened persons who were implanted in adolescence have been found to obtain higher levels of auditory performance than those implanted in adulthood. People implanted at an early age seem to perform better than people implanted in adulthood. 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.
- **Other factors**
  Other factors that may affect auditory performance include: (1) number of surviving spiral ganglion cells, (2) electrode placement and insertion depth, (3) electrical dynamic range, and (4) signal-processing strategy. There are also factors, such as the patient's level of intelligence and communicativeness, that are unrelated to deafness, but may also affect auditory performance.

Taking the above factors into account, Blamey [26] developed a three-stage model of auditory performance for postlingually deafened adults (Fig. 11). 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 occurs immediately 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 patient immediately attains improvements in auditory performance, depending on the duration of deafness. As the patient's experience with the implant 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 many factors that may affect performance. Unfortunately, it is not easy to assess the significance of individual factors on speech perception, due to the interaction among those factors. For example, in assessing meningitis as a factor that affects auditory performance, one needs to bear in mind that this disease is commonly associated with bone growth in the cochlea. This bone growth can obstruct the insertion of intracochlear electrodes. So, the etiology (in this case meningitis) is confounded



11. A three-stage model of auditory performance for postlingually deafened adults (Blamey, et al. [26]). The thick lines show measurable auditory performance, and the thin line shows potential auditory performance.

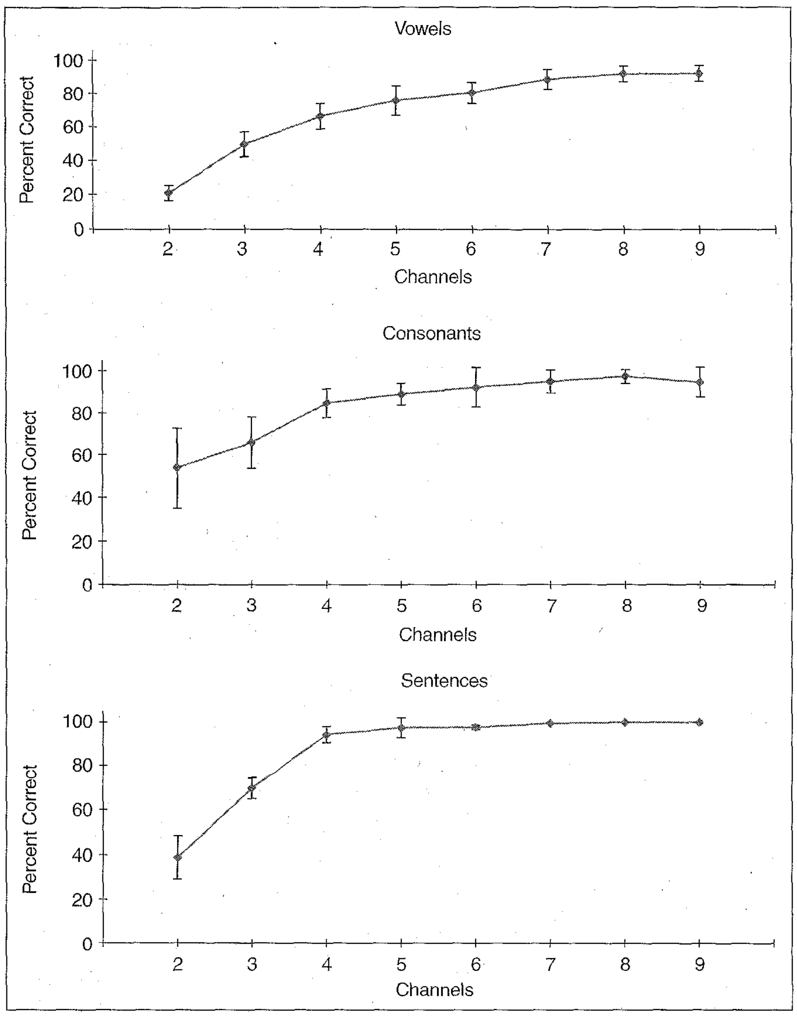
with electrode insertion depth and, as a consequence, we do not know if a patient performs poorly on speech recognition because of the etiology of hearing loss or because of shallow electrode insertion.

As a step toward assessing the effect of factors such as number of channels on auditory performance, Dorman and Loizou [27, 28] and Shannon, et al. [29, 30] 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.ualr.edu/asd/cimplants/). The reconstructed speech was presented to normal-hearing listeners for identification. Below, we describe some of our simulations that examined: (1) number of channels necessary for achieving high levels of speech understanding, and (2) the effect of electrode insertion depth on auditory performance.

There is a great variability in the speech-recognition performance of cochlear implant patients.

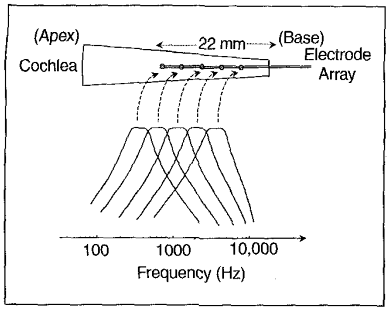
Number of Channels

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 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 the number 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.



12. Mean scores of normal-hearing listeners on recognition of vowels, consonants, and sentences as a function of number of channels [27]. Error bars indicate standard deviations.

The acoustic simulations [27] represent speech as a sum of sinusoids and mimic the front-end processing of the implant processor. More specifically, speech is reconstructed as a sum of sinusoids with time-varying amplitudes and fixed frequencies. The amplitudes of the sinusoids are computed in a manner similar to the continuous interleaved strategy (CIS). The speech signal is first processed through a pre-emphasis filter (lowpass below 1200 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 lowpass 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 and presented to normal-hearing listeners for identification. The results are shown in Fig. 12. As can be seen, the number of channels needed to reach asymptotic performance depends 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 sug-



13. Diagram showing the analysis filters used in a 5-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.

gest that high levels of speech understanding can be obtained with 5-8 independent channels of stimulation.

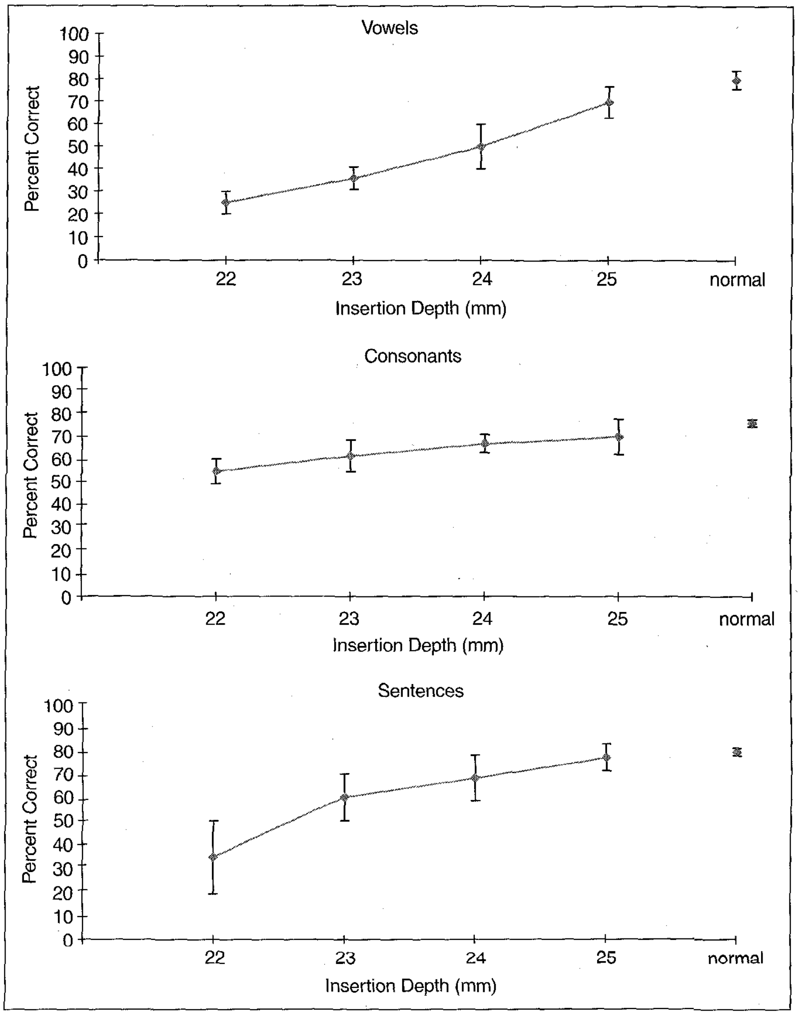
Electrode Insertion Depth and Frequency Up-Shifting

Electrode arrays are inserted only partially into the cochlea, typically 22-30 mm, depending on the state of the cochlea. 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 5 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. 13). 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 implants are first activated.

Although we would expect that speech understanding would be best when 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 hypothesis. 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 situation exists because of all the coexisting conditions that affect the patients' performance, and therefore make it difficult to assess the effect of insertion depth alone.

Acoustic simulations were used by Dorman and Loizou [28] to determine the effect of electrode insertion depth on speech understanding for a 5-channel cochlear prosthesis. Different insertion depths were simulated, ranging from 22 mm to 25 mm. Greenwood's frequency-to-place equation [31] 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 of 831, 1566, 2844, 5056, and 8924 Hz were generated. The corresponding amplitudes were computed using analysis filters with center frequencies of 418, 748, 1339, 2396, and 4287 Hz, respectively (Fig. 13). The results on consonant, vowel, and sentence recognition are shown in Fig. 14 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 dif-



14. Percent-correct recognition of vowels, consonants, and sentences as a function of simulated insertion depth [28]. The normal condition corresponds to the situation in which the analysis frequencies and output frequencies match exactly.

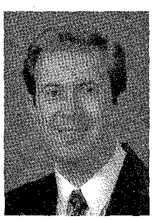
fered significantly from the normal condition (i.e., the condition in which the analysis and output frequency matched exactly) for all test materials. Performance in the 25 mm condition, however, did not differ significantly from the normal condition. These results suggest that relatively shallow insertions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

Summary

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 patients who (1) acquired speech and language before their hearing loss, and (2) have shorter duration of deafness. Gradual, but steady, improvements in speech production and speech perception have also occurred in prelingually deafened adults or children.

Acknowledgments

The author would like to thank Michael Dorman, Blake Wilson, 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 received his Ph.D. in electrical engineering from Arizona State University in 1995. From 1995 to 1996 he was a postdoctoral fellow in the Department of Speech and Hearing Science at Arizona State University, working on research related to cochlear prosthesis. He is now an assistant professor in the Department of Applied Science at the University of Arkansas at Little Rock. His research interests are in the areas of signal processing, speech recognition, spectral analysis, and cochlear implants. Dr. Loizou is a member of IEEE, the Acoustical Society of America, Eta Kappa Nu, and Phi Kappa Phi.

Address for Correspondence: Philipos Loizou, Assistant Professor, Department of Applied Science, University of Arkansas at Little Rock, Little Rock, AR 72204-1099. Tel: (501) 569-9067. Fax: (501) 569-8020. E-mail: loizou@ualr.edu.

References

1. **Fant G:** *Acoustic Theory of Speech Production.* The Hague, Netherlands: Mouton, 1970.
2. **Borden G, Harris K, and Raphael L:** *Speech Science Primer: Physiology, Acoustics, and Perception of Speech.* Baltimore, MD: Williams and Wilkins, 1994.
3. **Cooper F, Delattre P, Liberman A, Borst J, and Gerstman L:** Some experiments on the perception of synthetic speech sounds, *J. Acoust. Soc. Amer.*, 24(6): 597-606, November 1952.
4. **Yost W and Nielsen D:** *Fundamentals of Hearing: An Introduction.* New York: Holt, Rinehart and Winston, 1985.
5. **Hinojosa R and Marion M:** Histopathology of profound sensorineural deafness, *Ann. New York Acad. of Sci.*, 405: 459-484, 1983.
6. **Wilson B:** Signal processing, in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.). Singular Publishing Group, Inc, 1993, pp. 35-86,.
7. **Hochmair-Desoyer I, Hochmair E, and Burian K:** Design and fabrication of multiwire scala tympani electrodes, *Ann. New York Acad. of Sci.*, 405: 173-182, 1983.
8. **Clark G, Shepherd R, Patrick J, Black R, and Tong Y:** Design and fabrication of the banded electrode array, *Ann. New York Acad. of Sci.*, 405: 191-201, 1983.
9. **van den Honert and Stypulkowski P:** Single fiber mapping of spatial excitation patterns in the electrically stimulated nerve, *Hearing Res.*, 29: 1987.
10. **Merzenich M and White M:** 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.
11. **Wilson B, Lawson D, and Zerbi M:** Advances in coding strategies for cochlear implants, *Advances in Otolaryngology - Head and Neck Surgery*, 9: 105-129, 1995.
12. **Boothroyd A:** Profound deafness, in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.). Singular Publishing Group, Inc, 1993, pp. 1-34.
13. **Owens E, Kessler D, and Schubert E:** The minimal auditory capabilities (MAC) battery, *Hearing Aid J.*, 34: 9-34, 1981.
14. **Tyler R, Preece J, and Tye-Murray N:** The Iowa phoneme and sentence tests, Department of Otoloryngology-Head and Neck Surgery Tech. Report, University of Iowa, 1986.
15. **Tyler R:** Speech perception by children, in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.). Singular Publishing Group, Inc, 1993, pp. 191-256.
16. **Tobey E:** Speech production, in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.). Singular Publishing Group, Inc, 1993, pp. 257-316.
17. **Osberger M, Robbins A, Todd S, Riley A, and Miyamoto R:** 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,.
18. **Pancamo S and Tobey E:** Effects of multichannel cochlear implant upon sound production in children, *Proc. Second Annu. Southeastern Allied Health Res. Symp.*, pp. 319-330, 1989.
19. **Miyamoto R, Osberger M, Todd S, and Robbins A:** 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.
20. **Staller S, Dowell R, Beiter A, and Brimacombe J:** Perceptual abilities of children with the Nucleus 22-channel cochlear implant, *Ear and Hearing*, 12: 34S-47S, 1991.
21. **Erber N and Alencewicz C:** Audiologic evaluation of deaf children, *J. of Speech and Hearing Disorders*, 41: 256-267, 1976.
22. **Gantz B, Tyler R, Tye-Murray N, and Fryauf-Bertschy H:** 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.
23. **Shipp D and Nedzelski J:** 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.
24. **Gantz B, Woodworth G, Abbas P, Knutson J, and Tyler R:** Multivariate predictors of audiological success with multichannel cochlear implants, *Ann. Otology, Rhinology and Laryngology*, 102: 909-916, 1993.
25. **Summerfield A and Marshall D:** Preoperative predictors of outcomes from cochlear implantation in adults: Performance and quality of life, *Ann. Otology, Rhinology and Laryngology*, pp. 105-108, (Suppl. 166), 1995.
26. **Blamey P, Arndt P, Bergeron F, Bredberg G, Brimacombe J, et al.:** Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants, *Audiology and Neuro-Otology*, 1: 293-306, 1996.
27. **Dorman M, Loizou P, and Rainey D:** Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs, *J. Acoust. Soci. of Amer.*, 102: 2403-2411, 1997.
28. **Dorman M, Loizou P, and Rainey D:** Simulating the effect of cochlear implant electrode insertion-depth on speech understanding, *J. Acoust. Soci. of Amer.*, 102: 2993-2996, 1997.
29. **Shannon R, Zeng F, Kamath V, Wygonski J, and Ekelid M:** Speech recognition with primarily temporal cues, *Science*, 270: 303-304, 1995.
30. **Shannon R, Zeng F, and Wygonski J:** Speech recognition with altered spectral distribution of envelope cues, *J. Acoust. Soc. of Amer.*, 100, Pt. 2: 2692, 1996.
31. **Greenwood D:** A cochlear frequency-position function for several species—29 years later, *J. Acoust. Soc. of Amer.*, 87: 2592-2605, 1990.
32. **Wilson B, Finley C, Lawson D, and Wolford R:** Speech processors for cochlear prostheses, *Proc. IEEE*, 76: 1143-1154, September 1988.