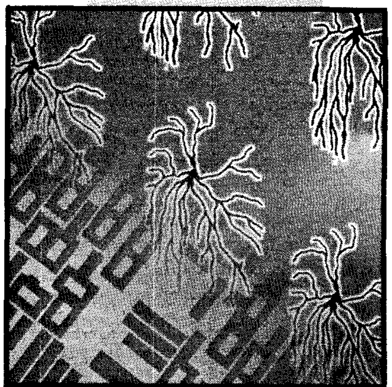


Signal-Processing Techniques for Cochlear Implants

A Review of Progress in Deriving Electrical Stimuli from the Speech Signal

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

Editor's Note: This article is a follow-up to the article that appeared in the January/February 1999 issue of this magazine (vol. 18, no. 1, pp. 32-42) and is adapted from a version that appeared in the September 1998 issue of IEEE Signal Processing Magazine *(vol. 15, no. 5, pp. 101-130).*

Cochlear implants have been successful in restoring partial hearing to profoundly deaf people. 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 cochlear implant design. Signal processing, in particular, played an important role in the development of different techniques for deriving electrical stimuli from the speech signal. The purpose of this article is to present a review of various signal-processing techniques that have been used for cochlear prosthesis over the past 25 years.

Single-Channel Implants

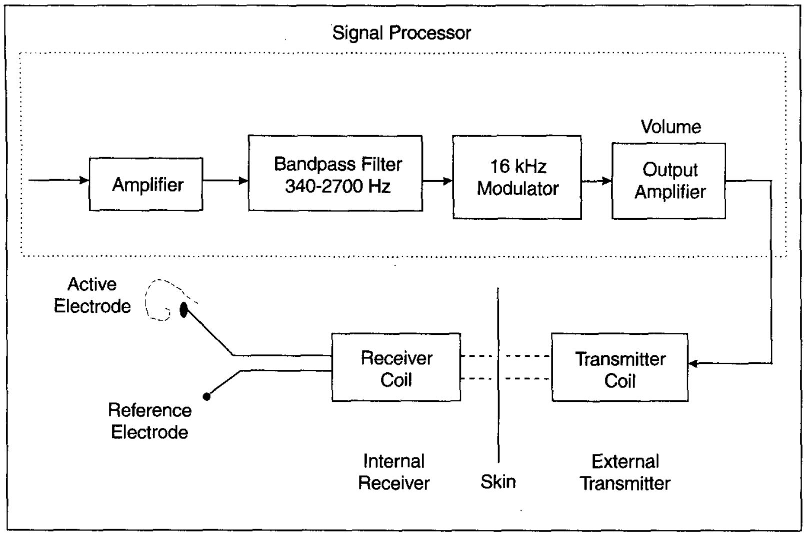
Single-channel implants provide electrical stimulation at a single site in the cochlea using a single electrode. Single-channel implants were first implanted in human subjects in the early 1970s. At that time, there was a lot of skepticism about whether single-channel stimulation could really work [1]. Doctors and scientists argued that electrical stimulation of the auditory nerve could produce nothing but noise. Despite the controversy, researchers in the United States and in Europe kept working on the development of single-channel prosthesis. These early efforts led to, among other devices, the House/3M and the Vienna/3M single-channel implants.

House/3M Device

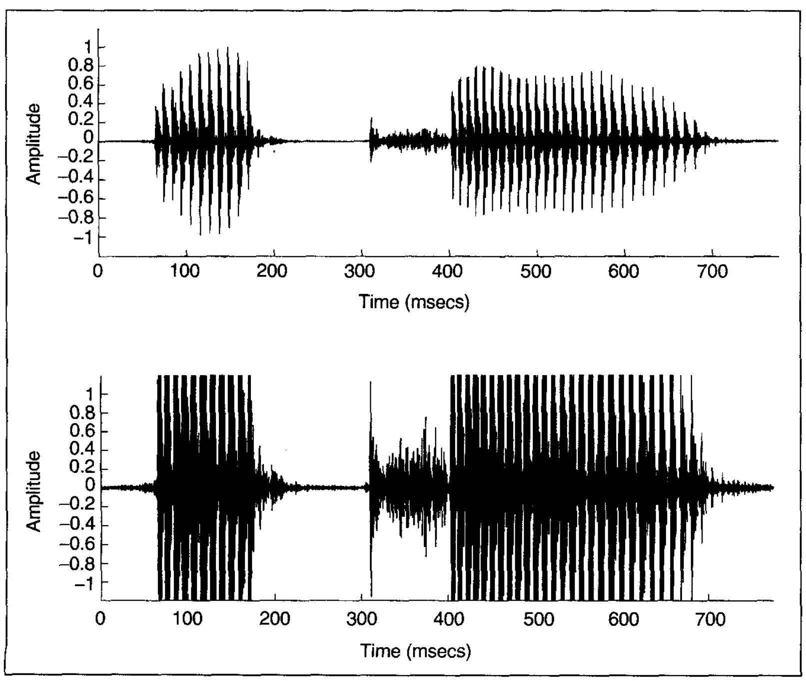
The House single-channel implant was originally developed by William House and his associates in the early 1970s [2, 3]. Improvements to the implant were later undertaken jointly with the 3M company, and the device was henceforth referred to as House/3M. Figure 1 shows the block diagram of this device. It consists of three main components: (1) the signal processor, (2) the signal transmitter/receiver, and (3) the implanted electrodes [4]. The acoustic signal is picked up by a microphone, amplified, and then processed through a 340-2700 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. The shape of the modulated envelope signal, however, is affected by the input signal level because this device does not attempt to reduce or limit the input dynamic range [4]. 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 Fig. 2). 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. 2, bursts of the 16 kHz carrier appear in synchrony with the period of voiced segments as well as with 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., [5]). Rosen, et al. [6], found that for four patients the average percent correct score on consonant identification was 37%. Only exceptional patients were able to obtain scores above zero on monosyllabic word (NU-6) identification. In a study by Danhauer, et al. [7], only four patients (out of 18) achieved a 2% correct score, and only one patient achieved a 4% correct score on monosyllabic word identification.



1. Block diagram of the House/3M single-channel implant. The signal is processed through a 340-2700 Hz filter, modulated with a 16 kHz carrier signal, and then transmitted (without any demodulation) to a single electrode implanted in the scala tympani.



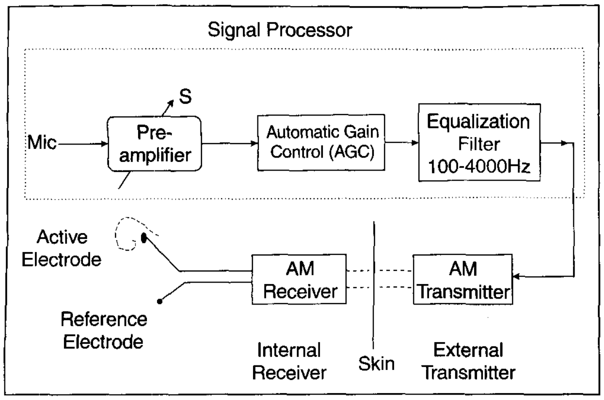
2. The time waveform (top) of the word "aka," and the amplitude modulated waveform (bottom) processed through the House/3M implant for input signal level exceeding 70 dB SPL.

Vienna/3M Device

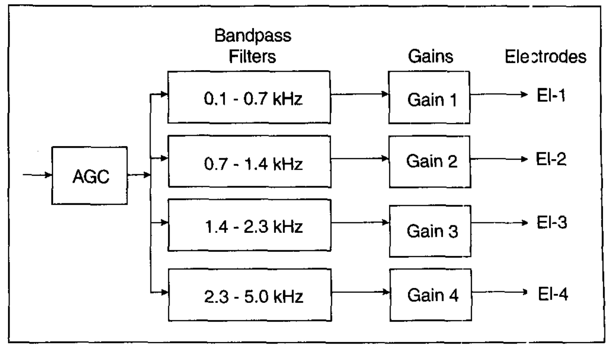
The Vienna single-channel implant was developed at the Technical University of Vienna, Austria, in the early 1980s [8] and is shown in the block diagram of Fig. 3. The signal is first preamplified and then compressed using a gain-controlled amplifier with a short attack time (0.5 msec). The amount of compression is adjusted according to the patient's dynamic range. The compressed signal is then fed through a frequency-equalization filter (Fig. 4), which attenuates frequencies outside the range of 100-4000 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 of 100-4000 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 at low frequencies and higher at high frequencies (< 300 Hz) [9, 10]. The frequency response of the equalization filter 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 [11], some of the exceptional patients were able to identify words in sentences with 86% accuracy. Word identification



3. 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 (100-4000 Hz), and is amplitude-modulated for transcutaneous transmission. The implanted receiver demodulates the radio-frequency signal and delivers it to the implanted electrode.



4. 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.

scores ranged from 15% to 86% correct across nine patients. Hochmair-Desoyer, et al. [12], 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. [5]) 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 [13]. 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 [9] showed that implant patients (as well as normal-hearing listeners [14]) cannot discriminate differences in pitch for frequencies above 300 Hz.

Single-channel stimulation restricts the amount of spectral information that an implant patient can receive to frequencies below 1 kHz. This is not sufficient, however, for speech perception, because there is important information in the speech signal up to 4000 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 [15].

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 by using an array of electrodes. Thus, 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 signal frequency. 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 first introduced in the 1980s, several questions were raised regarding multichannel stimulation:

1. 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?

2. Since more than one electrode will be stimulated, what kind of information should be transmitted to each electrode? Should it be some type of spectral features or attributes of the speech signal that are 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?

Single-channel stimulation does not exploit the place-code mechanism used by a normal cochlea for encoding frequencies

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 number of channels needed to obtain high levels of speech understanding is still the subject of debate (e.g., Shannon, et al. [16], Dorman, et al. [17]). 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 prosthesis can be divided into three main categories: waveform, feature-extraction and hybrid. These strategies differ in the way information is extracted from the speech signal and presented to the electrodes. Waveform strategies present some type of waveform (in analog or pulsatile form) derived by filtering the speech signal into different frequency bands. Feature-extraction strategies present some type of spectral features, such as formants, derived using feature extraction algorithms. Hybrid strategies present some of both. A review of these signal processing strategies is given below.

Waveform Strategies

Compressed-Analog (CA) Approach

The compressed-analog (CA) approach was originally used in the Ineraid device manufactured by Symbion, Inc., Utah [18] and is shown in the block diagram of Fig. 4. This approach was also used in the UCSF/Storz device [19], which is now discontinued. The signal is first compressed using 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 through a percutaneous connection to four intracochlear electrodes. The filtered waveforms are delivered simultaneously to four electrodes, spaced 4 mm apart, operating in monopolar configuration.

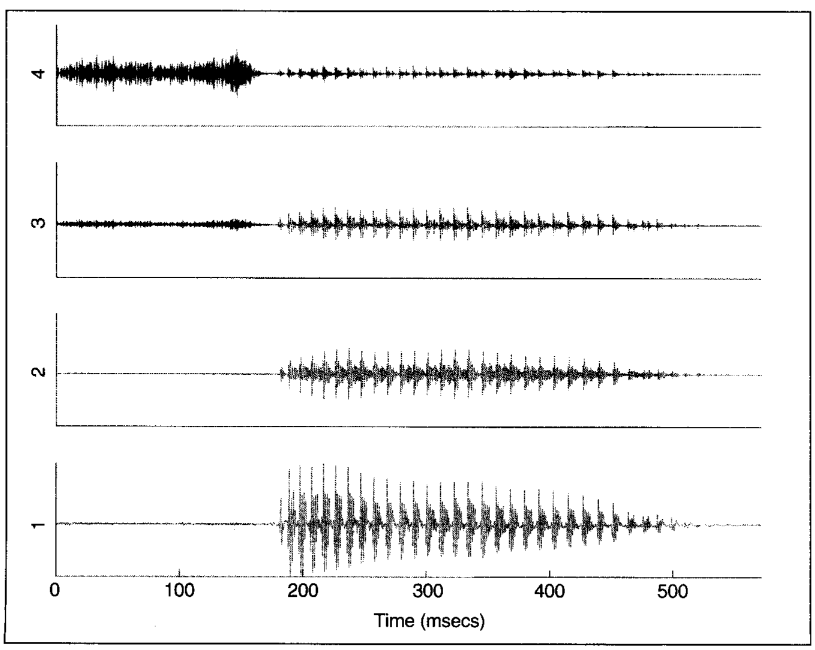
As an example, Figure 5 shows the four bandpassed waveforms produced for the syllable "sa" using a simplified implementation of the CA approach. As can be seen, spectral information is conveyed by the relative energy within each channel. During the "s" portion, most of the energy is contained in the high-frequency channel (channel 4) with almost no energy in the other, lower in frequency, channels. This distribution indicates that a high frequency sound (/s/) is present. On the other hand, during the "a" portion, most of the energy is contained in the low-frequency channels (channels 1 and 2), which indicates that a low-frequency sound (such as the vowel /a/) is present.

The CA approach was very successful because it enabled many patients to obtain open-set speech understanding. Dorman, et al. [20], report, for a sample of 50 Ineraid patients, a median score of 45% correct for word identification using the Central Institute for the Deaf (CID) sentences of everyday speech, a median score of 14% correct for monosyllabic word identification, and a median score of 14% correct for spondee (two-syllable) words. The CA approach clearly yielded superior speech-recognition performance over the single-channel approach [18]; not surprising given the increased frequency resolution provided by multiple-channel stimulation.

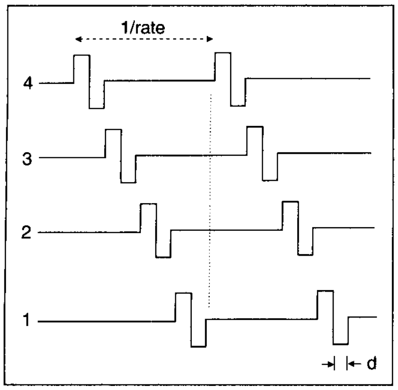
Continuous Interleaved Sampling (CIS) Approach

The CA approach uses analog stimulation that simultaneously delivers four continuous analog waveforms to four electrodes. A major concern associated with simultaneous stimulation is the interaction between channels caused by the summation of electrical fields from individual electrodes [21]. 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 [22], which addressed the channel interaction issue by using nonsimultaneous, interleaved pulses. Trains of biphasic pulses



5. 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).



6. Interleaved pulses used in the CIS strategy. The period between pulses on each channel (1/rate) and the pulse duration (d) per phase are indicated.

are delivered to the electrodes in a nonoverlapping (nonsimultaneous) fashion; i.e., such that only one electrode is stimulated at a time (Fig. 6). The amplitudes of the pulses are derived by extracting the envelopes of bandpassed waveforms. The CIS approach is shown in more detail in Fig. 7. The signal is first pre-emphasized and then passed through a bank of bandpass filters. The envelopes of the filtered waveforms are 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 logarithmic compression function is used to ensure that the envelope outputs fit the patient's dynamic range of electrically evoked hearing. Trains of balanced biphasic pulses, with amplitudes proportional to the envelopes, are delivered to the six electrodes at a constant rate in a nonoverlapping fashion (see Fig. 6). 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 higher performance than low-pulse rate. Figure 8 shows the pulsatile waveforms produced for the syllable "sa" using a simplified implementation of the CIS strategy. The pulse amplitudes were estimated by extracting the envelopes of the filtered waveforms (Fig. 6).

Several studies (e.g., [22-24]) were conducted by RTI and other institutions comparing the differences in performance between the CA and CIS strategies. The results [22] for seven patients tested on open-set recognition of 50 monosyllabic words (NU-6) and 100 keywords from the CID test showed that 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 [25, 26], Boex, et al. [24]). Several factors could be responsible for the success of the CIS 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: the Clarion, the Med-El, and the Nucleus CI24M.

CIS Parameters

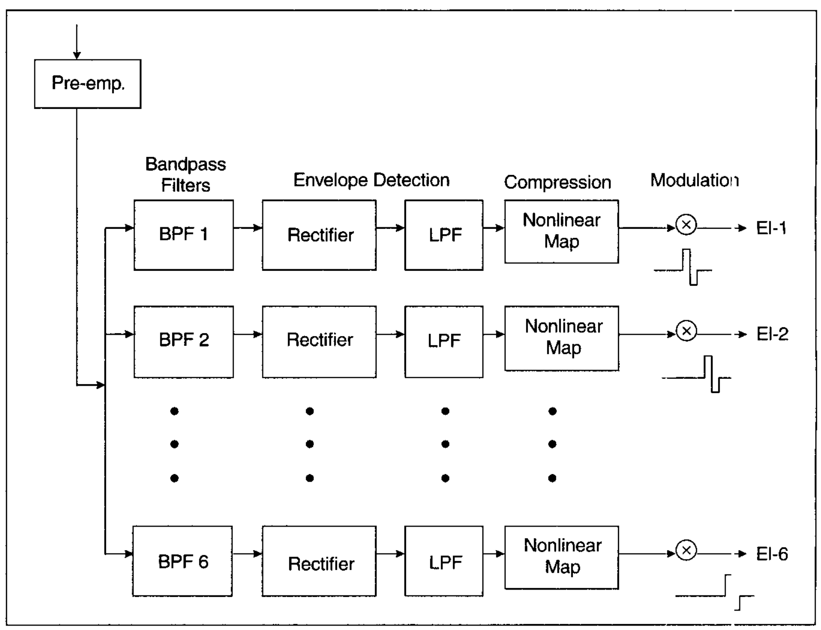
There are a number of parameters associated with the CIS approach that could be varied to optimize speech-recognition performance for each patient [23, 25]. These parameters include pulse rate and pulse duration, stimulation order, and compression function.

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 pps and as high as 2500 pps have been used. The "optimal" pulse rate for speech recognition varies from patient to patient. Wilson, et al. [23], reported that some patients obtain a maximum performance on the 16-consonant recognition task with 833 pps and a pulse duration of 33 µsecs/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 secs/phase pulses. One would expect that better performance would be obtained with very high pulse rates, since high-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 refers to the order with which the electrodes are stimulated. The stimulation order can be varied to minimize possible interaction between channels. One possibility is to stimulate the electrodes in an apex-to-base order; i.e., first stimulate electrode 1, then 2, etc., and lastly, 6. This way, signals in the low frequencies (apex) are stimulated first, and signals in the high frequencies (base) are stimulated last. This apex-to-base order, however, does not minimize the spatial



7. 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 lowpass 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. 6).

separation between sequentially stimulated electrodes. Alternatively, the electrodes can be stimulated in a so called "staggered" order (i.e., 6-3-5-2-4-1), which maximizes the spatial separation between stimulated electrodes. As with the pulse rate, preference for stimulation order varies from patient to patient. Some patients prefer the apex-to-base stimulation, because they say speech sounds more natural and intelligible, while other patients prefer the staggered-order stimulation.

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 of acoustic amplitudes in conversational speech is considerably larger than the implant patient's dynamic range. Dynamic range is 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. It has been shown [27, 28] 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 type 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 μAmps. For the power-law compression function, the constants A and B can be computed as follows:

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

$$B = THR - Ax_{min}^p \tag{2}$$

The values of threshold, THR, and most-comfortable levels, MCL, may vary from electrode to electrode.

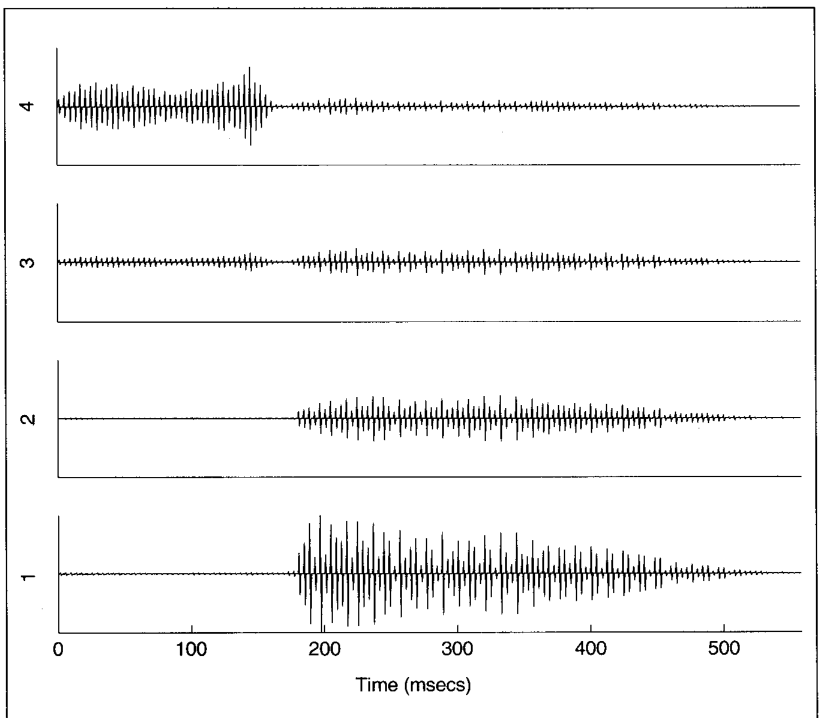
Feature-Extraction Strategies

Feature-extraction strategies were initially used in the multielectrode Nucleus implant, developed at the University of Melbourne, Australia, by Clark and his colleagues and manufactured by Nucleus Limited [29]. This device has gone through a number of improvements since it was first introduced in the early 1980s. Initially, it employed a feature-extraction approach based on a fundamentally different principle than the CA or CIS approach. Rather than presenting waveform information obtained by the filtering the speech signal into a few frequency bands, the processor presented spectral features, such as formants, obtained by formant-extraction algorithms. A review of the feature-extraction strategies used in the Nucleus implant is given below.

With multichannel implants different auditory nerve fibers can be stimulated at different places in the cochlea.

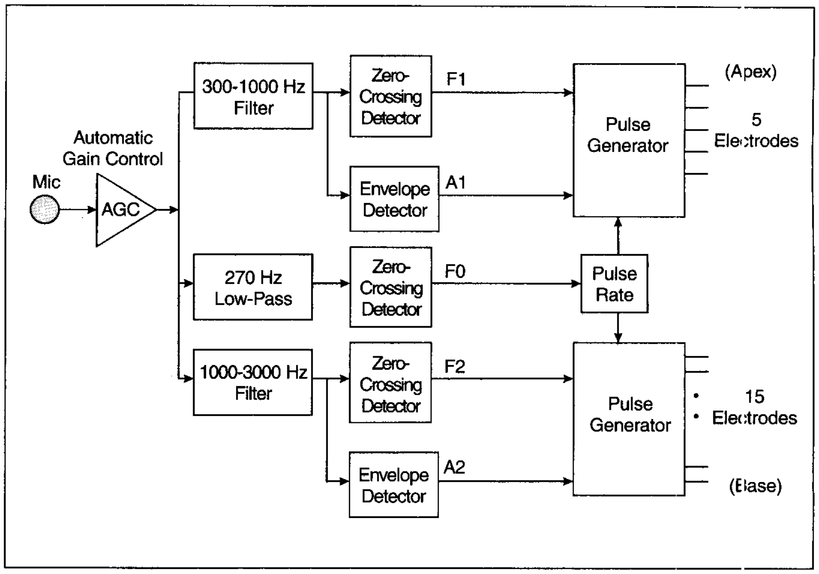
F0/F2 Strategy

The F0/F2 strategy was first developed for the Nucleus device in the early 1980s [29, 30]. In this strategy, the fundamental frequency (F0) and the second formant



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

(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 lowpass filter, and another zero-crossing detector is used to estimate F2 from the output of a 1000-4000 Hz bandpass filter. The amplitude of F2 is estimated with an envelope detector by rectifying and lowpass 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 are set in proportion to the amplitude of F2. During unvoiced segments (e.g., consonant /s/) the selected electrode is stimulated at quasi-random intervals at an average rate of 100 pulses/sec. Initial results [31] with the F0/F2 strategy were encouraging as it enabled some patients to obtain open-set speech understanding.



9. 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.

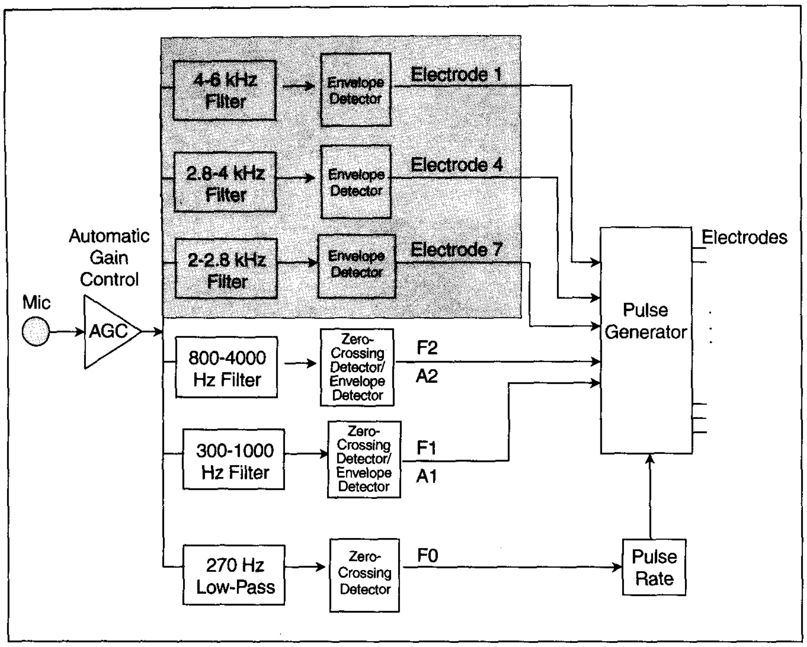
F0/F1/F2 Strategy

The F0/F2 strategy was later modified to include information about the first formant frequency (F1) [32] 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 280-1000 Hz bandpass filter. The block diagram of the F0/F1/F2 processor is shown in Fig. 9. 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 are dedicated to F1, since they can represent frequencies up to 1000 Hz, while the remaining 15 electrodes are dedicated to F2 since they can represent frequencies above 1000 Hz. For voiced segments, two electrical pulses are produced. One pulse is applied to an electrode pair chosen according to F2, and the second pulse is applied to an electrode pair chosen according to F1. The pulses are biphasic, with each phase lasting 200 µsec. A 800-µsec spacing between pulses is used to avoid any interaction that might occur due to temporal channel interactions. The amplitudes of the two pulses are set in proportion to the corresponding amplitudes of F1 and F2. As in the F0/F2 processor, the electrodes were stimulated at F0 pps for voiced segments and at an average rate of 100 pps for unvoiced segments.

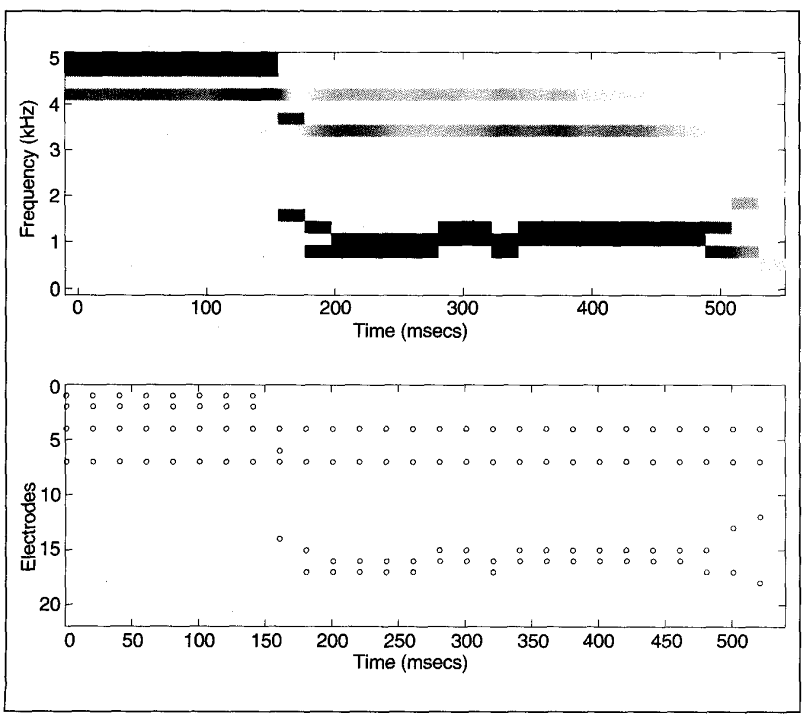
The addition of F1 information improved the speech-recognition performance of patients wearing the Nucleus cochlear implant. Within-patient comparisons between the F0/F2 and F0/F1/F2 strategies demonstrated improvements in speech understanding with the F0/F1/F2 strategy. Dowell, et al. [33], found that 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. [34], also reported that the mean scores on monosyllabic word identification (NU-6) correspondingly improved from 8% to 28%. No significant difference was found on tests of consonant recognition in the hearing-only condition. Significant improvements 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 this strategy emphasizes low-frequency information, which is required for vowel recognition. The majority of the consonants, however, contain high-frequency information, and this has motivated the refinement of the F0/F1/F2 strategy to the MPEAK strategy.

MPEAK Strategy

Further improvements were made in the late 1980s by Cochlear Pty. Limited (a subsidiary of Nucleus Limited) in collaboration with the University of Melbourne [35, 36]. 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 processing was used, which considerably reduced the size and weight of the new processor, now called a miniature speech processor (MSP). A new coding strategy, called MULTIPEAK (or MPEAK), extracted high-frequency information from the speech signal, in addition to formant information. The block diagram of the MPEAK strategy is shown in Fig. 10. 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 800-4000 Hz. Additional high-frequency information was extracted from the frequency bands 2000-2800 Hz, 2800-4000 Hz, and 4000-6000 Hz. The motivation for using the three additional bands was twofold: (1) to enhance the representation of the second formant (F2), and (2) to include high-frequency information, which is important for the perception of consonants. The estimated envelope amplitudes of the



10. Block diagram of the MPEAK strategy. Similar to the F0/F1/F2 strategy, the formant frequencies (F1,F2), and fundamental frequency (F0) 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 F0 pulses/sec for voiced sounds, and at a quasi-random rate for unvoiced sounds.



11. 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.

three bandpass filters were delivered to fixed electrodes 7, 4, and 1, which 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 F0 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 F1 and F2 electrodes and on the high-frequency electrodes 4 (2000-2800 Hz) and 7 (2800-4000 Hz). The high-frequency electrode 1 is not stimulated because there is generally little energy in the spectrum above 4 kHz for voiced sounds. For unvoiced sounds, stimulation occurs on the high-frequency electrodes 1, 4, and 7, as well as on the electrode corresponding to F2. The electrode corresponding to F1 is not stimulated because there is generally little energy below 1000 Hz for unvoiced sounds (e.g., /s/). Figure 11 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 [37], using a group of five patients, found that the MPEAK strategy yielded a mean improvement of 17% on consonant identification. A mean improvement of 28% on open-set sentence recognition was also found with the MPEAK strategy. Several other studies (e.g., [38, 39]) confirmed that the MPEAK strategy achieved significant improvements over the F0/F1/F2 strategy on open-set speech recognition.

Although the MPEAK strategy has proven 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, tends 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 speech processor —the SMSP processor.

"N-of-M" Strategies

In "n-of-m" strategies, the signal is filtered into m frequency bands, and the processor selects the n ($n < m$) envelope outputs with the largest energy. Only the

Waveform strategies present some type of waveform derived by filtering the speech signal into different frequency bands.

electrodes corresponding to the *n* selected outputs are stimulated at each cycle. This strategy can be considered to be a hybrid strategy, in that it combines feature and waveform representations.

Interleaved Processor (IP)

The interleaved processor (IP), developed at the Research Triangle Institute, was the first processor to use an "n-of-m" strategy [40]. The signal was filtered into six frequency bands, and the two envelope outputs with the highest energy were selected at each stimulation cycle. Pulses were then delivered in an interleaved manner to the corresponding two electrodes at a maximum rate of 313 pps per channel.

Spectral Maxima Sound Processor (SMSP)

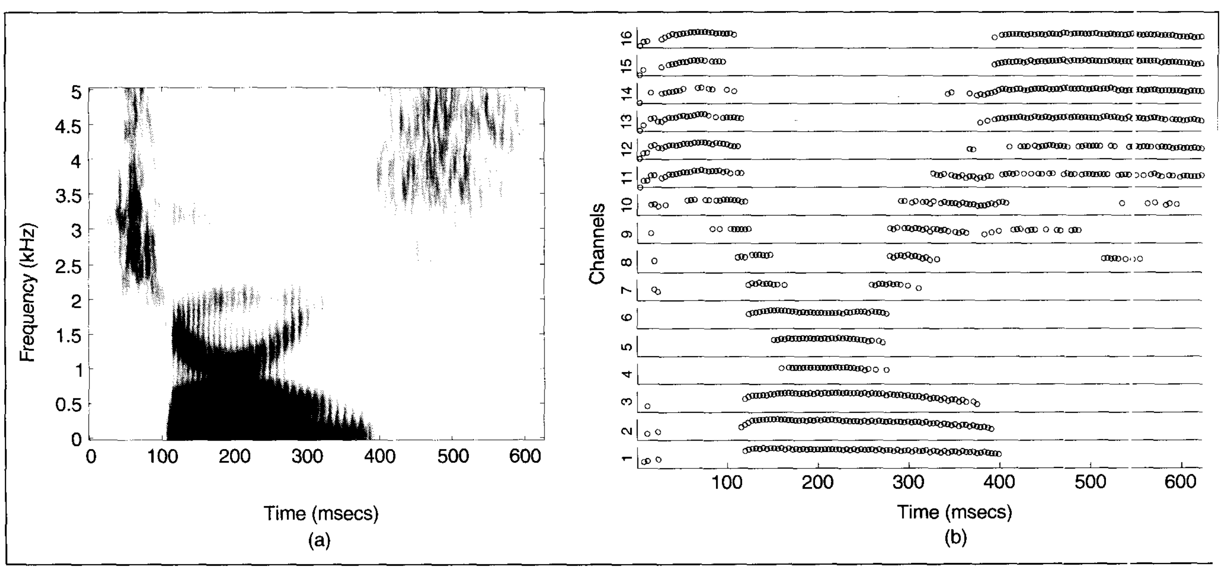
A new processor, called the spectral maxima sound processor (SMSP), was developed in the early 1990s [41] for the Nucleus implant. Unlike previous processors developed, the SMSP processor did not extract any features (e.g., F1, F2) from the speech waveform. Instead it analyzed the speech signal using a bank of 16 bandpass filters and a spectral maxima detector. The signal from the microphone is first pre-amplified and then sent through a bank of 16 bandpass filters with center frequencies ranging from 250 to 5400 Hz. The output of each filter is rectified and lowpass filtered with a cutoff frequency of 200 Hz. After computing all 16 filter outputs, the SMSP processor selects, at 4 msec intervals, the six largest filter outputs. The SMSP processor therefore uses a 6-of-16 strategy. The six amplitudes of the spectral maxima are then 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.

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 SMSP processor delivers biphasic pulses to the electrodes at a constant rate of 250 pps for both voiced and unvoiced sounds. Figure 12 illustrates the pattern of electrical stimulation for the word "choice." As can be seen, the electrodes selected for stimulation in each cycle vary, depending upon the spectral content of the signal.

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 [41]. The SMSP strategy was later refined and incorporated in the Nucleus Spectra 22 processor.

Current State-of-the-Art 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, manufac-



12. Example of the SMSP strategy using the word "choice." (a) shows the spectrogram of the word choice and (b) 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 "ɔ" portion of the word, low-frequency channels (1-6) are selected.

tured by Med-El Corporation, Austria, which is currently in clinical trials in the United States.

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 [42]. 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. 13). 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 [42] 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 6 maxima. Figure 14 shows examples of electrical stimulation patterns for four different sounds using the SPEAK strategy.

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 patient's individual parameters. For broadband spectra, more maxima are selected and the stimulation rate is slowed. 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.

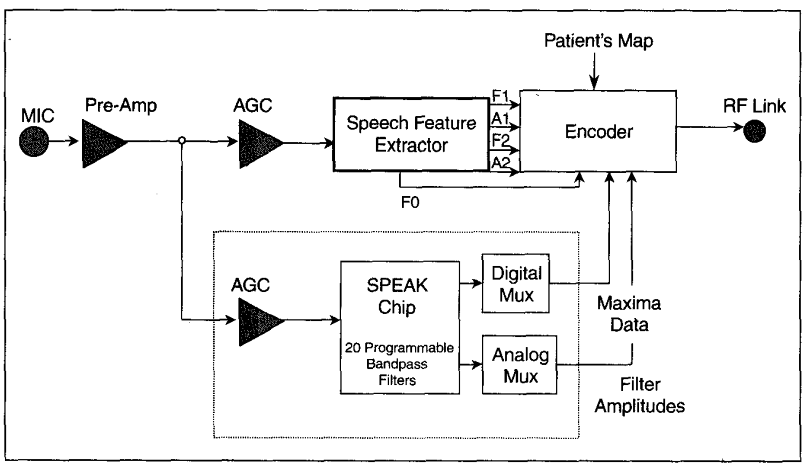
Comparison of the SPEAK strategy and the MPEAK strategy by Skinner, et al. [43], using 60 patients showed that the SPEAK strategy performed better than the MPEAK strategy on vowel, consonant, monosyllabic word and sentence recognition. Especially large improvements in performance were found with tests 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 which does not extract any features from the speech signal.

The SPEAK strategy is now incorporated in a new investigational device 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 worn on the waist, and the ear-level size worn behind the ear. The ear-level version is the size of a behind-the-ear hearing aid. Some of the features of the Nucleus 24 (CI24M) system include: (1) two additional electrodes to be placed outside the inner ear to allow different modes of stimulation, (2) a removable internal magnet for future MRI compatibility, and (3) high-rate stimulation strategies including a CIS strategy.

The "n-of-m" strategy combines feature and waveform representations.

Clarion Processor

The Clarion cochlear-implant system [44, 45] 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 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



13. 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.

There is still much to be learned about electrical stimulation of the auditory nerve.

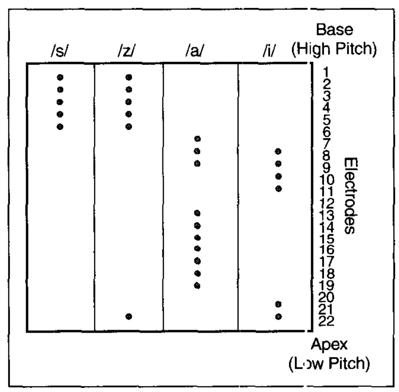
then delivered simultaneously to eight electrode pairs. Analog waveforms are delivered to each electrode at a rate of 13,000 samples/sec per channel. The CA strategy emphasizes detailed temporal information at the expense of reduced spatial selectivity, due to simultaneous stimulation. For some patients, 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 then passed through a bank of eight bandpass filters. The envelopes of the filtered waveforms are then extracted by full-wave rectification and lowpass 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 pps per channel in an interleaved fashion.

The Clarion processor (version 1.0) was recently approved by FDA, and the initial results on open-set speech recognition were very encouraging. In a recent study by Loeb and Kessler [46], 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. [47], showed that the pulsatile version (CIS) of the Clarion processor (ver. 1.0) obtained superior performance over the CA version. 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). 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 [47].

Several changes were recently made to the Clarion implant, resulting in Clarion ver. 1.2. Some of those changes include: (1) a smaller speech processor, (2) improved filter implementation using bandpass filters with 30 dB/octave rolloff, 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 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 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 rate per channel to 1666 pps.

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 [48] is based on the Motorola 56001 DSP, and can be programmed with either a high-rate CIS strategy or a high-rate "n-of-m" (SPEAK-type) strategy. The Med-El cochlear implant (also referred to as COMBI-40 [49]) uses a very soft electrode carrier specially designed to facilitate deep electrode insertion into the cochlea. Because of this deep insertion



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

Table 1. 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	Clarion 1.0
CIS	Envelope signals	Pulsatile	8	833 pps	Clarion 1.0
CIS	Envelope signals	Pulsatile	12	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

(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 implants. 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 pps 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 applied to a bank of eight (logarithmically spaced) bandpass Butterworth filters of sixth-order. The bandpass filter outputs are full-wave rectified and lowpass filtered with a cutoff of 400 Hz. The 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 1515 pps per channel. The pulses are transmitted transcutaneously through a radio-frequency link [50]. 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 pps 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). This "n-of-m" strategy is 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 rate.

The Med-El implant processor is widely used in Europe [49]. A percutaneous version 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 reported by Dorman and Loizou [25, 26, 51, 52]. Significant improvements were obtained on all test materials compared to the performance with the Ineraid device (CA strategy).

Conclusions and Future Directions

Much of the success of cochlear implants is due to the advancement of signal-processing techniques developed over the years (see Table 1). While this success is very encouraging, there is still much to be learned about electrical stimulation of the auditory nerve. Future research in cochlear prosthesis should:

1. Continue investigating the strengths and limitations of present signal-processing strategies including CIS and SPEAK. The findings of such investigations may lead to the development of signal-processing techniques capable of transmitting more information to the brain.
2. Develop noise-reduction algorithms that will help implant patients better communicate in noisy environments.
3. Identify factors that contribute to the variability in performance among patients [53]. Knowing these factors may help us develop signal-processing techniques that are patient-specific. Patients will then be optimally fitted 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 performance between "poorly performing" and "better-performing" patients.
4. Develop preoperative procedures that can predict how well a patient will perform with a particular type of cochlear implant.
5. 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.
6. 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.
7. Investigate the effect of high-rate pulsatile stimulation (> 3000 pps) on speech perception as well as on music appreciation using more than eight channels.

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

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. R55 DC03421 from NIH/NIDCD.

Philipos C. 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, Acoustical Society of America, Eta Kappa Nu, and Phi Kappa Phi.

Address for Correspondence: Philipos C. Loizou, Assistant Professor, Department of Applied Science, University of Arkansas at Little Rock, Little Rock, AR 72204-1099. Tel: (501) 569-8067. Fax: (501) 569-8020. E-mail: loizou@ualr.edu.

References

1. **House W:** A personal perspective on cochlear implants, in *Cochlear Implants* (R. Schindler and M. Merzenich, eds.), New York: Raven Press, pp. 13-16, 1985.

2. **House W, Urban J:** Long term results of electrode implantation and electronic stimulation of the cochlea in man. *Ann Otology, Rhinology and Laryngology,* 82: 504-517, 1973.

3. **House W, Berliner K:** Cochlear implants: Progress and perspectives. *Ann Otology, Rhinology and Laryngology*, Suppl. 91: 1-124, 1982.

4. **Edgerton B, Brimacombe J:** Effects of signal processing by the House-3M cochlear implant on consonant perception. *Acta Otolaryngologica*, Suppl. 411: 115-123, 1984.

5. **Gantz B, Tyler R, Knutson J, Woodworth G, Abbas P, et al.:** Evaluation of five different cochlear implant designs: Audiologic assessment and predictors of performance. *Laryngoscope*, 98: 1100-1106, 1988.

6. **Rosen S, Walliker J, Brimacombe J, and Edgerton B:** Prosodic and segmental aspects of speech perception with the House/3M single-channel implant. *J Speech and Hearing Res,* 32: 93-111, 1989.

7. **Danhauer J, Ghadialy F, Eskwitt D, and Mendel L:** Performance of 3M/House cochlear implant users on tests of speech perception. J. *American Academy of Audiology,* 1: 236-239, 1990.

8. **Hochmair-Desoyer I, Hochmair E:** Percepts elicited by different speech-coding strategies, *Ann New York Academy of Sciences.* 405: 268-279, 1983.

9. **Shannon R:** Multi-channel electrical stimulation of the auditory nerve in man: I. Basic psychophysics. *Hearing Res,* 11: 157-189, 1983.

10. **Shannon R:** Psychophysics, in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.).

Singular Publishing Group, Inc., pp. 357-388, 1993.

11. **Tyler R:** Open-set recognition with the 3M/Vienna single-channel cochlear implant. *Archives of Otolaryngology, Head and Neck Surgery*, 114: 1123-1126, 1988.

12. **Hochmair-Desoyer I, Hochmair E and Stiglbrunner H:** Psychoacoustic tempora processing and speech understanding in cochlear implant patients, in *Cochlear Implants* (R. Schindler and M. Merzenich, eds.). New York: Raven Press, pp. 291-304, 1985.

13. **Abbas P:** Electrophysiology, in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.), Singular Publishing Group, Inc, pp. 317-356, 1993.

14. **Burns E, Viemester N:** Nonspectral pitch. *J Acoust Soc Amer*, 60:. 863-869, 1976.

15. **Dorman M:** Speech perception by adults, in *Cochlear Implants: Audiological Foundations* (R. Tyler, ed.). Singular Publishing Group, Inc, pp. 145-190, 1993.

16. **Shannon R, Zeng F, Kamath V, Wygonski J , and Ekelid M:** Speech recognition with primarily temporal cues. *Science*, 270: 303-304, 1995.

17. **Dorman M, 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, *J Acoust Soc Amer*, 102: 2403-2411, 1997.

18. **Eddington D:** Speech discrimination in deaf subjects with cochlear implants, *J Acoust Soc Amer*, 68(3): 885-891, 1980.

19. **Merzenich M, Rebscher S, Loeb G, Byers C, Schindler R:** The UCSF cochlear implant project: State of development. *Advances in Audiology*, 2: 119-144, 1984.

20. **Dorman M, Hannley M, Dankowski K, Smith L, and McCandless G:** Word recognition by 50 patients fitted with the Symbion multi-channel cochlear implant. *Ear and Hearing*, 10: 44-49, 1989.

21. **White M, Merzenich M, Gardi J:** Multi-channel cochlear implants: Channel interactions and processor design. *Archives of Otolaryngology*, 110: 493-501, 1984.

22. **Wilson B, Finley C, Lawson D, Wolford R, Eddington D, W. Rabinowitz W:** Better speech recognition with cochlear implants. *Nature*, 352: 236-238, July 1991.

23. **Wilson B, Lawson D, Zerbi M:** Advances in coding strategies for cochlear implants. *Advances in Otolaryngology - Head and Neck Surgery*, 9: 105-129, 1995.

24. **Boex C, Pelizzone M, Montandon P:** Improvements in speech recognition with the CIS strategy for the Ineraid multi-channel intracochlear implant, in *Advances in Cochlear Implants* (I. Hochmair-Desoyer and E. Hochmair, eds.). Vienna: Manz, pp. 136-140, 1994.

25. **Dorman M, Loizou P:** 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*, 18: 147-155, 1997.

26. **Dorman M, Loizou P:** Mechanisms of vowel recognition for Ineraid patients fit with continuous interleaved sampling processors, *J Acoust Soc Amer*, 102: 581-587, 1997.

27. **Eddington D, Dobelle W, Brachman D, Mladevosky M, Parkin J:** Auditory prosthesis research using multiple intracochlear stimulation in man. *Ann Otology, Rhinology and Laryngology*, 87 (Suppl. 53): 1-39, 1978.

28. **Zeng F, Shannon R:** Loudness balance between acoustic and electric stimulation, *Hearing Res*, 60: 231-235, 1992.

29. **Clark G:** The University of Melbourne-Nucleus multi-electrode cochlear implant. *Advances in Oto-Rhino-Laryngology*, 38: 1-189, 1987.

30. **Seligman P, Patrick J, Tong Y, Clark G, Dowell R, Crosby P:** A signal processor for a multiple-electrode hearing prosthesis. *Acta Otolaryngologica*, Suppl. 411: 135-139, 1984.

31. **Dowell R, Mecklenburg D, and Clark G:** Speech recognition for 40 patients receiving multi-channel cochlear implants. *Archives of Otolaryngology, Head and Neck Surgery*, 112: 1054-1059, 1986.

32. **Blamey P, Dowell R, Clark G:** Acoustic parameters measured by a formant-estimating speech processor for a multiple-channel cochlear implant. *J Acoust Soc Amer*, 82: 38-47, 1987.

33. **Dowell R, Seligman P, Blamey P, Clark G:** Evaluation of a two-formant speech processing strategy for a multi-channel cochlear prosthesis, *Ann Otology, Rhinology and Laryngology*, 96 (Suppl. 128): 132-134, 1987.

34. **Tye-Murray N, Lowder M, and Tyler R:** Comparison of the F0/F2 and F0/F1/F2 processing strategies for the Cochlear Corporation cochlear implant. *Ear and Hearing*, 11: 195-200, 1990.

35. **Patrick J, Seligman P, Money D, Kuzma J:** Engineering, in *Cochlear Prostheses* (Clark G, Tong Y, and Patrick J, eds.), Edinburgh: Churchill Livingstone, pp. 99-124, 1990.

36. **Patrick J, Clark G:**, The Nucleus 22-channel cochlear implant system. *Ear and Hearing*, 12(Suppl. 1): 3-9, 1991.

37. **Wallenberger E, Battmer R:** Comparative speech recognition results in eight subjects using two different coding strategies with the Nucleus 22 channel cochlear implant. *Br J Audiology*, 25: 371-380, 1991.

38. **Dowell R, Dawson P, Dettman S, Shepherd R, Whitford L, Seligman P, Clark G:** Multi-channel cochlear implantation in children: A summary of current work at the University of Melbourne, *American Journal of Otology*, 12(Suppl. 1): 137-143, 1991.

39. **Skinner M, Holden L, Holden T, Dowell R, Seligman P, Brimacombe J, Beiter A:** 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*,12: 3-22, 1991.

40. **Wilson B, Finley C, et al.:** Comparative studies of speech processing strategies for cochlear implants, *Laryngoscope*, 98: 1069-1077, 1988.

41. **McDermott H, McKay C, Vandali A:** A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant. *J Acoust Soc Amer*, 91: 3367-3371, 1992.

42. **Seligman P, McDermott H:** Architecture of the Spectra 22 speech processor, *Ann Otology, Rhinology and Laryngology*, 104(Suppl. 166): 139-141, 1995.

43. **Skinner M, Clark G, Whitford L, Seligman P, Staller S, et al.:** Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system, *Amer J Otology*, 15(Suppl. 2): 15-27, 1994.

44. **Schindler R, Kessler D:** Preliminary results with the Clarion cochlear implant, *Laryngoscope*, 102: 1006-1013, 1992.

45. **Kessler D, Schindler R:** Progress with a multi-strategy cochlear implant system: The Clarion, *in Advances in Cochlear Implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, pp. 354-362, 1994.

46. **Loeb G, Kessler D:** Speech recognition performance over time with the Clarion cochlear prosthesis, *Ann Otology, Rhinology and Laryngology*, 104 (Suppl. 166): 290-292, 1995.

47. **Tyler R, Gantz B, Woodworth G, Parkinson A, Lowder M, Schum L:** Initial independent results with the Clarion cochlear implant, *Ear and Hearing*, 17: 528-536, 1996.

48. **Zierhofer C, Peter O, Bril S, Pohl P, Hochmair-Desoyer I, Hochmair E:** A multi-channel cochlear implant system for high-rate pulsatile stimulation strategies, in *Advances in Cochlear Implants* (I. Hochmair-Desoyer and E. Hochmair, eds.), Vienna: Manz, pp. 204-207, 1994.

49. **Helms J, Muller J, Schon F, et al.:** Evaluation of performance with the COMBI 40 cochlear implant in adults: A multicentric clinical study, *ORL*, vol. 59, pp. 23-35, 1997.

50. **Zierhofer C, Hochmair-Desoyer I, and Hochmair E:** Electronic design of a cochlear implant for multi-channel high-rate pulsatile stimulation strategies, *IEEE Trans Rehab Eng*, 3(1): 112-116, March 1995.

51. **Loizou P, M Dorman M, Powell V:** The recognition of vowels produced by men, women, boys and girls by cochlear implant patients using a 6 channel CIS processor, *J Acoust Soc Amer*, 103: 1141-1149, 1998.

52. **M Dorman and P Loizou:** Improving consonant intelligibility for Ineraid patients fit with Continuous Interleaved Sampling (CIS) processors by enhancing contrast among channel outputs. *Ear and Hearing*, 17: 308-313, 1996.

53. Cochlear implants in adults and children, *NIH Consensus Statement*, 13: 1-30, May 1995.

54. **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.