Interfacing Sensors With the Nervous System: Lessons From the Development and Success of the Cochlear Implant

Blake S. Wilson, Senior Member, IEEE, and Michael F. Dorman

Invited Paper

Abstract—The cochlear implant is the most successful neural prosthesis to date and may serve as a paradigm for the development or further development of other systems to interface sensors with the nervous system, e.g., visual or vestibular prostheses. This paper traces the history of cochlear implants and describes how the current levels of performance have been achieved. Lessons and insights from this experience are presented in concluding sections.

Index Terms—Auditory prosthesis, cochlear implant, deafness, hearing, nervous system, neural prosthesis, sensors, speech perception, vestibular prosthesis, visual prosthesis.

I. INTRODUCTION

THE COCHLEAR implant is one of the great success stories of modern medicine. Just 30 years ago, cochlear implants provided little more than a sensation of sound and sound cadences. They were useful as an aid to lipreading. Now, a majority of implant users enjoy high levels of speech recognition using hearing alone; indeed, many can use the telephone without difficulty. This is a long trip in a short time, and the restoration of function—from total or nearly total deafness to useful hearing—is truly remarkable.

In this paper, we trace this history and indicate how the present levels of performance have been achieved. The design of cochlear implants is described in some detail to provide an example of ways in which sensors can be successfully interfaced to the nervous system. Results from studies with implant patients are presented. In addition, we describe some of the limitations of present systems and possibilities for overcoming them. We conclude with a section on the lessons learned from cochlear implants and how those lessons might inform the designs of other types of sensory neural prostheses, such as prostheses for the restoration of vision or balance.

II. A BRIEF HISTORY

As recently as the early 1980s, many eminent and highly knowledgeable people believed that cochlear implants would provide only an awareness of environmental sounds and possibly speech cadences to their users. Many were skeptical of implants and thought that mimicking or reinstating the function of the exquisite machinery in the normal inner ear was a fool’s dream. Among these critics were world-renowned experts in otology and auditory physiology. Fortunately, pioneers persisted in the face of this intense criticism and provided the foundations for present devices. Detailed reviews of the early history of cochlear implants are presented in [1]–[3].

A timeline of assessments in the development of cochlear implants is presented in Table I. These range from frank skepticism at the beginning to high enthusiasm by 1995.

The first implant of a device for electrical stimulation of the auditory nerve was performed by Djourno and Eyrès in Paris in 1957. An induction coil was used, with one end placed on the stump of the auditory nerve or adjacent brainstem and the other end within the temporalis muscle (the patient had had bilateral cholesteatomas which had been removed in prior operations, taking the cochleas and peripheral parts of the auditory nerves with them). The patient used the device for several months before it failed, and was able to sense the presence of environmental sounds but could not understand speech or discriminate among speakers or many sounds. In 1961, Dr. William F. House implanted two patients in Los Angeles, each with single gold wires inserted a short distance into the (deaf) cochlea. By 1975, more patients had been implanted worldwide, most by Dr. House, and 13 had functioning, single-channel devices. The United States National Institutes of Health (NIH) commissioned a study at that point, to assess the performance of those devices and to determine whether support by the NIH for the further development of cochlear implants would be wise. The report from the study [4], the “Bilger report,” is a landmark in the field. Its key conclusion was that “although the subjects could not understand speech through their prostheses, they did score significantly higher on tests of lipreading and recognition of environmental sounds with their prostheses activated than without them.” This and earlier assessments are included in Table I.
Shortly after the Bilger report was published, the NIH did elect to support research and development efforts in the field. The rapid progress thereafter in the design and performance of implant systems was in very large part the direct result of this decision. In particular, work supported through the Neural Prosthesis Program at the NIH, first directed by Dr. F. Terry Hambrecht and later by Dr. William J. Heetderks, produced many important innovations in electrode and speech processor designs that remain in use to this day.

In 1988, NIH convened the first of two consensus development conferences on cochlear implants. Multichannel systems—those with multiple channels of processing and with multiple sites of stimulation in the cochlea—had come into use at that time. The consensus statement from the 1988 conference [5] suggested that multichannel implants were more likely to be effective than single-channel implants, and indicated that about 1 in 20 patients could carry out a normal conversation without lipreading. Approximately 3000 patients had received cochlear implants by 1988.

New and highly effective processing strategies for cochlear implants were developed in the late 1980s and early 1990s, principally through the Neural Prosthesis Program. Among these were the continuous interleaved sampling (CIS) [6], n-of-m [7], and spectral peak (SPEAK) [8] strategies. Large gains in speech reception performance were achieved with these strategies, two of which remain in widespread use today (CIS and n-of-m). A detailed review of processing strategies and their lines of development is presented in [9].

The second NIH consensus development conference was held in 1995. By then, approximately 12,000 patients had received implants. A major conclusion from the 1995 conference [10] was that “a majority of those individuals with the latest speech processors for their implants will score above 80 percent correct on high-context sentences even without visual cues.”

By the middle of 2006, the cumulative number of implants worldwide exceeded 110,000. This number is orders of magnitude higher than the numbers for all other types of neural protheses, including those for restoration of motor or other sensory functions.

Fig. 1 shows the number of cochlear implants over time, beginning in 1957 with the first implant operation by Djouorno and Eyriès. The growth in numbers since then is exponential.

III. DESIGN OF COCHLEAR IMPLANTS

A. Aspects of Normal Hearing

In normal hearing, sound waves traveling through air reach the tympanic membrane via the ear canal, causing vibrations that move the three small bones of the middle ear. This action produces a piston-like movement of the stapes, the third
bone in the chain. The “footplate” of the stapes is attached to a flexible membrane in the bony shell of the cochlea called the oval window. Inward and outward movements of this membrane induce pressure oscillations in the cochlear fluids, which in turn initiate a traveling wave of displacement along the basilar membrane (BM), a highly specialized structure that divides the cochlea along its length. This membrane has graded mechanical properties. At the base of the cochlea, near the stapes and oval window, it is narrow and stiff. At the other end of the cochlea, near the apex, the membrane is wide and flexible. These properties give rise to the traveling wave and to points of maximal response according to the frequency or frequencies of the pressure oscillations in the cochlear fluids. The traveling wave propagates from the base to the apex. For an oscillation with a single frequency, the magnitude of displacements increases up to a particular point along the membrane and then drops precipitously thereafter. High frequencies produce maxima near the base of the cochlea, whereas low frequencies produce maxima near the apex.

Motion of the BM is sensed by the inner hair cells (IHCs) in the cochlea, which are attached to the top of the BM in a matrix of cells called the organ of Corti. Each hair cell has fine rods of protein, called stereocilia, emerging from one end. When the BM moves at the location of a hair cell, the rods are deflected as if hinged at their bases. Such deflections in one direction increase the release of chemical transmitter substance at the base (other end) of the cells, and deflections in the other direction inhibit the release. The variations in the concentration of the chemical transmitter substance act at the terminal ends of auditory neurons, which are immediately adjacent to the bases of the IHCs. Increases in chemical transmitter substance increase discharge activity in the nearby neurons, whereas decrements in the substance inhibit activity. Changes in neural activity thus reflect events at the BM. These changes are transmitted to the brain via the auditory nerve, the collection of all neurons that innervate the cochlea.

The steps described above are illustrated in the top panel of Fig. 2. This shows a cartoon of the main anatomical structures, including the tympanic membrane, the three bones of the middle ear, the oval window, the BM, the IHCs, and the adjacent neurons of the auditory nerve.

B. Loss of Hearing

The principal cause of hearing loss is damage to or complete destruction of the sensory hair cells. Unfortunately, the hair cells are fragile structures and are subject to a wide variety of insults, including but not limited to genetic defects, infectious diseases (e.g., rubella and meningitis), overexposure to loud sounds, certain drugs (e.g., kanamycin, streptomycin, and cisplatin), and aging. In the deaf or deafened cochlea, the hair cells are largely or completely absent, severing the connection between the peripheral and central auditory systems. The function of a cochlear prosthesis is to bypass the (missing) hair cells by stimulating directly the surviving neurons in the auditory nerve.

The anatomical situation faced by designers of cochlear implants is illustrated in the bottom panel of Fig. 2. The panel shows a complete absence of hair cells. In general, a small number of cells may remain for some patients, usually in the apical (low frequency) part of the cochlea.

Without the normal stimulation provided by the hair cells, the peripheral part of the neurons—between the cell bodies in the spiral ganglion and the terminals within the organ of Corti—undergo “retrograde degeneration” and eventually die [11]. Fortunately, the cell bodies are far more robust. At least some usually survive, even for prolonged deafness or for virulent etiologies such as meningitis [11]–[13]. These cells, or more specifically the nodes of Ranvier just distal or proximal to them, are the putative sites of excitation for cochlear implants.

C. Electrical Stimulation of the Auditory Nerve

Direct stimulation of the nerve is produced by currents delivered through electrodes placed in the scala tympani (ST), one of three fluid-filled chambers along the length of the cochlea. (The boundary between the ST and the scala media is formed by the BM and organ of Corti.) A cutaway drawing of the implanted cochlea is presented in Fig. 3. The figure shows a partial insertion of an array of electrodes into the ST. The array is inserted through a drilled opening made by the surgeon in the bony shell of the cochlea overlying the ST (called a “cochleostomy”) and close to the base of the cochlea. Alternatively, the array may be
inserted through the second flexible membrane of the cochlea, the round window membrane, which also is close to the basal end of the cochlea and ST (see drawing).

The depth of insertion is limited by the decreasing lumen of the ST from base to apex, the curvature of the cochlear spiral, and an uneven and unsmooth lumen particularly in the apical region. No array has been inserted farther than about 30 mm, and typical insertions are much less than that, e.g., 18–26 mm. (The total length of the typical human cochlea is about 35 mm.) In some cases, only shallow insertions are possible, such as when bony obstructions in the lumen impede further insertion.

Different electrodes in the implanted array may stimulate different subpopulations of neurons. As described above, neurons at different positions along the length of the cochlea respond to different frequencies of acoustic stimulation in normal hearing. Implant systems attempt to mimic or reproduce this “tonotopic” encoding by stimulating basally situated electrodes (first turn of the cochlea and lower part of the drawing) to indicate the presence of high-frequency sounds, and by stimulating electrodes at more apical positions (deeper into the ST and ascending along the first and second turns in the drawing) to indicate the presence of sounds with lower frequencies. Closely spaced pairs of bipolar electrodes are illustrated here, but arrays of single electrodes that are each referenced to a remote electrode outside the cochlea also may be used. This latter arrangement is called a “monopolar coupling configuration” and is used in all present-day implant systems that are widely applied worldwide. (There are three such systems and they constitute more than 99% of the cochlear implant market.)

The spatial specificity of stimulation with a ST electrode most likely depends on a variety of factors, including the orientation and geometric arrangement of the electrodes, the proximity of the electrodes to the target neural structures, and the condition of the implanted cochlea in terms of nerve survival and ossification. An important goal of electrode design is to maximize the number of largely nonoverlapping populations of neurons that can be addressed with the electrode array. Present evidence suggests, however, that no more than 4–8 independent sites may be available using current designs, even for arrays with as many as 22 electrodes [14]–[19]. Most likely, the number of independent sites is limited by substantial overlaps in the electric fields from adjacent (and more distant) electrodes. The overlaps are unavoidable for electrode placements in the ST, as the electrodes are sitting in the highly conductive fluid of the perilymph and additionally are relatively far away from the target neural tissue in the spiral ganglion. A closer apposition of the electrodes next to the inner wall of the ST would move them a bit closer to the target cells (see Fig. 3), and such placements have been shown in some cases to produce an improvement in the spatial specificity of stimulation [20]. However, a large gain in the number of independent sites may well require a fundamentally new type of electrode, or a fundamentally different placement of electrodes. The many issues related to electrode design, along with prospects for the future, are discussed in [20]–[30].

D. Components of Cochlear Implant Systems

The essential components in a cochlear prosthesis include: 1) a microphone for sensing sound in the environment; 2) a speech processor to transform the microphone input into a set of stimuli for the implanted array of electrodes; 3) a transcutaneous link for the transmission of power and stimulus information across the skin; 4) an implanted receiver/stimulator to decode the information received from the radio-frequency signal produced by an external coil and then to generate stimuli using the instructions obtained from the decoded information; 5) a cable to connect the outputs of the receiver/stimulator to the electrodes; and 6) the array of electrodes. These components must work together as a system to support excellent performance and a weakness in a component can degrade performance significantly. For example, a limitation in the data bandwidth of the transcutaneous link can restrict the types and rates of stimuli that can be specified by the external speech processor and this, in turn, can limit performance. A thorough discussion of considerations for the design of cochlear prostheses and their constituent parts is presented in [27].

We note that an earlier implant system, the Ineraid® device, had a percutaneous connector rather than a transcutaneous link. In addition, several experimental implant systems included percutaneous connectors. Although use of these through-the-skin connectors increased the risk of infection, they also provided
direct electrical access to the implanted electrodes from an external speech processor or other stimulating or recording equipment. This access allowed full stimulus control and high-fidelity recordings of intracochlear evoked potentials. A wide variety of speech processing strategies was evaluated with subjects having percutaneous connectors. This was vital for the development of strategies now in widespread use and for the acquisition of knowledge about the stimulus-response properties of the electrically stimulated auditory nerve in humans.

E. Transformation of a Microphone Input Into Stimuli for the Implant

An important aspect of the design for any type of sensory neural prosthesis is how to transform an input from a sensor or array of sensors into a set of stimuli that can be interpreted by the nervous system. The stimuli can be electrical or tactile, for examples, and usually involve multiple sites of stimulation, corresponding to the spatial mapping of inputs and representations of those inputs in the nervous system. One approach to the transformation – and probably the most effective approach – is to mimic or replicate at least to some extent the damaged or missing physiological functions that are bypassed or replaced by the prosthesis.

Of course, limitations in other parts of the prosthesis system may restrict what can be done with the transformation. Effects of limitations in the bandwidth of the transcutaneous link for cochlear implant systems have been mentioned. Also, a lack of independence among stimulus sites can greatly reduce the number of channels of information that can be conveyed to the nervous system. In such cases, a high number of channels in processing the input(s) from the sensor(s) would not in general produce any benefit and might even degrade performance.

For cochlear implants, this part of the design is called the processing strategy. As noted previously, advances in processing strategies have produced quite large improvements in the speech reception performance of implant patients, from recognition of a tiny percentage of monosyllabic words with the first strategies and multisite stimulation, for example, to recognition of a high percentage of the words with current strategies and multisite stimulation.

One of the simpler approaches supporting the recent levels of performance with implants is illustrated in Fig. 4. This is the CIS strategy, which is used as the default strategy or as a processing option in all implant systems now in widespread clinical use.

The CIS strategy filters speech or other input sounds into bands of frequencies with a bank of bandpass filters. Envelope variations in the different bands are represented at corresponding electrodes in the cochlea with modulated trains of biphasic electrical pulses. The envelope signals extracted from the bandpass filters are compressed with a nonlinear mapping function prior to the modulation, in order to map the wide dynamic range of sound in the environment (around 90 dB) into the narrow dynamic range of electrically evoked hearing (about 10 dB or somewhat higher). The output of each bandpass channel is directed to a single electrode, with low-to-high channels assigned to apical-to-basal electrodes, to mimic at least the order, if not the precise locations, of frequency mapping in the normal cochlea. The pulse trains for the different channels and corresponding electrodes are interleaved in time, so that the pulses across channels and electrodes are non-simultaneous. This eliminates a principal component of electrode interaction, which otherwise would be produced by direct vector summation of the electric fields from different (simultaneously stimulated) electrodes. The corner frequency of the low-pass filter in each envelope detector typically is set at 200 Hz or higher, so that the fundamental frequencies of speech sounds are represented in the modulation waveforms. CIS gets its name from the continuous sampling of the (compressed) envelope signals by rapidly presented pulses that are interleaved across electrodes. Between 4 and 22 channels (and corresponding stimulus sites) have been used in CIS implementations to date.

Other strategies also have produced excellent results. Among these are the \( \eta \)-of-\( m \) strategy mentioned above, and the advanced combination encoder (ACE) strategy [31], which is similar in design and performance to the \( \eta \)-of-\( m \) strategy [9]. The principal difference between CIS and the \( \eta \)-of-\( m \) or ACE strategies is that the channel outputs are “scanned” in the latter two strategies to select the \( \eta \) channels with the highest envelope signals prior to each frame of stimulation across electrodes. Stimulus pulses are delivered only to the subset of \( m \) electrodes that correspond to the \( \eta \) selected channels. This spectral or channel “peak picking” scheme is designed, in part, to reduce the density of stimulation while still representing the most important aspects of the acoustic environment. The deletion of low-amplitude channels (and associated stimuli) for each frame of stimulation may reduce the overall level of masking or interference across electrode and stimulus regions in the cochlea. To the extent that the omitted channels do not contain significant information, such “unmasking” may improve the perception of the input signal by the patient. In addition, for positive signal-to-noise ratios (\( S/N \)s), selection of the highest peaks in the spectra may emphasize the primary speech signal with respect to the noise. Detailed descriptions of these and related processing strategies, along with detailed descriptions of prior strategies, are presented in [9].
Fig. 5. Percent correct scores for 55 users of the COMBI 40 implant and the CIS processing strategy. Scores for recognition of the Hochmair–Schultz–Moser (HSM) sentences are presented in the top panel, and scores for recognition of the Freiburger monosyllabic words are presented in the bottom panel. The solid line in each panel shows the median of the scores, and the dashed and dotted lines show the interquartile ranges. The data are an updated superset of those reported in [32], kindly provided by Patrick D’Haese of Med El GmbH, in Innsbruck, Austria. The experimental conditions and implantation criteria are described in [32]. All subjects took both tests at each of the indicated intervals following initial fitting of their speech processors. Identical scores at a single test interval are displaced horizontally for clarity. Thus, for example, the horizontal “line” of scores in the top right portion of the top panel all represent scores for the 24-month test interval. (Figure is from [9] and is used here with the permission of Whurr Publishing Ltd.)

IV. PERFORMANCE WITH PRESENT-DAY SYSTEMS

A. Average Performance and Range of Scores

Each of these strategies—CIS, ACE, and n-of-m—supports recognition of monosyllabic words on the order of 50% correct (using hearing alone), across populations of tested subjects (see [9, Table 2.4]). Variability in outcomes is high, however, with some patients achieving scores at or near 100% correct and with other patients scoring close to zero on this most difficult of standard audiological measures. Standard deviations of the scores range from about 10% to about 30% for the various studies conducted to date.

Results from a large and carefully controlled study are presented in Fig. 5. This figure shows scores for 55 users of the Med El COMBI 40 implant system (Med El GmbH, Innsbruck, Austria) and the CIS processing strategy. Scores for the Hochmair-Schultz-Moser (HSM) sentences are presented in the top panel, and scores for recognition of the Freiburger monosyllabic words are presented in the bottom panel. Results for five measurement intervals are shown, ranging from one month to two years following the initial fitting of the speech processor. The solid line in each panel shows the median of the individual scores and the dashed and dotted lines show the interquartile ranges. The data are a superset of those reported in [32], that include scores for additional subjects at various test intervals.

Most of the subjects used an 8-channel processor with a pulse rate of about 1500/s/electrode. Some of the subjects used fewer channels and a proportionately higher rate. (All processors used the maximum overall rate of 12,120 pulses/s across electrodes.)

As is evident from the figure, scores are broadly distributed at each test interval and for both tests. However, ceiling effects are encountered for the sentence test for many of the subjects, especially at the later test intervals. At 24 months postfitting, 47 of the 55 subjects score at 75% correct or higher, consistent with the 1995 NIH Consensus Statement. Scores for recognition of monosyllabic words are much more broadly distributed, with only a few subjects scoring 90% correct or higher.

An interesting aspect of the results presented in Fig. 5 is an apparent improvement in performance over time. This is easiest to see in the lower ranges of scores, e.g., in the steady increase in the lower interquartile lines (the dotted lines) across test intervals.

Improvements over time are even more evident in plots of mean scores for sentences and for words, as shown in Fig. 6 for these same data and for additional test intervals for the sentence test. The mean scores increase for both the sentence and word tests out to twelve months and then plateau thereafter. The mean scores for the sentence test asymptote at about 90% correct, and the mean scores for the word test asymptote at about 55% correct. Such results typify performance with the best of the modern cochlear implant systems and processing strategies, for electrical stimulation on one side with a unilateral implant.

These results are especially remarkable for the top scorers, given that only a maximum of eight broadly overlapping sectors
of the auditory nerve are stimulated with this device and the implementation of CIS used with it. This number is quite small in comparison to the normal complement of approximately 30,000 neurons in the human auditory nerve.

The results also show a learning or accommodation effect, with continuous improvements in scores over the first 12 months of use. This suggests the likely importance of brain function in determining outcomes, and the reorganization or “knitting” (brain plasticity) that must occur to utilize such sparse inputs to the maximum extent possible.

B. Top Performers

The top performers with present-day cochlear implants can achieve remarkably high scores in tests of speech recognition. Scores for one such subject, implant subject HR4, are shown in the black bars in Fig. 7 for a comprehensive and difficult set of tests. Mean scores for six undergraduate students with normal hearing and taking the same tests are shown in the gray bars, along with the standard error of the mean for each test. HR4 was totally deaf prior to receiving his implant. The tests included recognition of monosyllabic, consonant-nucleus-consonant (CNC) words; recognition of City University of New York (CUNY) sentences; recognition of Hearing in Noise Test (HINT) sentences; recognition of Arizona Biomedical Institute (AzBio) sentences; identification of consonants (Cons) in an /e/-consonant-/e/ context; identification of vowels (Vowels) in a /b/-vowel-/t/ context; and recognition of CUNY and AzBio (Az) sentences presented in competition with a four-talker babble, at the indicated speech-to-babble ratios (+5 or +10 dB). (Figure is from [33] and is used here with the permission of the IEEE.)

Fig. 7. Percent-correct scores for implant subject HR4 and for six subjects with normal hearing. Tests included recognition of monosyllabic, consonant-nucleus-consonant (CNC) words; recognition of City University of New York (CUNY) sentences; recognition of Hearing in Noise Test (HINT) sentences; recognition of Arizona Biomedical Institute (AzBio) sentences; identification of consonants (Cons) in an /e/-consonant-/e/ context; identification of vowels (Vowels) in a /b/-vowel-/t/ context; and recognition of CUNY and AzBio (Az) sentences presented in competition with a four-talker babble, at the indicated speech-to-babble ratios (+5 or +10 dB). (Figure is from [33] and is used here with the permission of the IEEE.)

Recognition of CUNY and AzBio sentences presented in competition with a four-talker babble, at the speech-to-babble ratio (S/B) of +10 dB for the CUNY sentences and that ratio and +5 dB for the AzBio sentences. Further details about the subjects, tests, and testing procedures are presented in [33].

Fig. 7 shows a spectacular restoration of function for a user of a sensory neural prosthesis. All of the scores for HR4 are high. His scores for speech material presented in quiet, including words, sentences, consonants, and vowels, match or closely approximate the scores for the control group. His score for the most difficult test used in standard clinical practice, recognition of the monosyllabic CNC words is 100% correct. In contrast, some of his scores for sentences presented in competition with speech babble are worse than normal. Although his score for the CUNY sentences at the S/B of +10 dB is 98% correct, his scores for the AzBio sentences at the S/Bs of +10 dB and +5 dB are below those of the normal-hearing subjects. In all, HR4 scored at or near the ceiling of 100% correct for seven of the nine tests, and he attained scores of 77% correct or better for the remaining two tests. (The subjects with normal hearing scored at or near the ceiling for all nine tests.) HR4 scored at the ceiling for all tests given in standard clinical practice to identify deficits in hearing. His results indicate a full restoration of clinically-normal function, at least for speech reception. He used a 16-channel CIS processor, as implemented in the Clarion® CII cochlear prosthesis (Advanced Bionics Corp., Sylmar, CA, USA) [34]. This prosthesis also includes a high-bandwidth transcutaneous link, current sources with short rise and fall times, an array of 16 intracochlear electrodes, and (in the version used) a positioning device to place the electrodes next to the inner wall of the ST.

Such high scores overall are consistent with HR4’s ability to communicate with ease in most listening situations. He has no difficulty at all in telephone communications. He can understand conversations not directed to him and can identify speakers by regional dialect. He can mimic voices and accents that he has heard only after receiving the implant. His speech reception abilities are truly remarkable, abilities that could not have been imagined 20 years ago, even by the most-optimistic proponents of cochlear implants.

Other patients, using this and other implant systems, and also other processing strategies (including the $n$-of-$m$ and ACE strategies), have achieved similarly high scores. For example, one of the subjects in Fig. 5 achieved a score of 98% correct in the Freiburger monosyllabic word test at the two-year interval. This subject used a COMBI 40 implant system, with its eight channels of CIS processing and eight sites of stimulation. This system also has a high-bandwidth transcutaneous link and current sources with short rise and fall times. It does not include a positioning device; nor do other versions of the Clarion prosthesis or other implant systems, that also support stellar scores for some patients.

Although more than a few patients have achieved scores like those shown in Fig. 7, most patients have lower scores, typically much lower scores for the difficult tests, as also indicated in the lower panel of Fig. 5. However, the results obtained with HR4 and his peers are an existence proof of what is possible with electrical stimulation of the auditory nerve in a totally deafened ear.
V. STRENGTHS AND LIMITATIONS OF PRESENT SYSTEMS

A. Efficacy of Sparse Representations

Some patients achieve spectacularly high scores with present-day cochlear implants. Indeed, their scores are in the normal ranges even for the most difficult of standard audiological tests. Such results are both encouraging and surprising in that the implants provide only a very crude mimicking of only some aspects of the normal physiology. In cases like that of patient HR4, 16 overlapping sectors of the auditory nerve are stimulated with 16 intracochlear electrodes. In other cases, other patients have achieved similarly high scores with 6–8 sites of stimulation in the cochlea, as noted above. The spatial specificity of stimulation with implants is much lower than that demonstrated in neural tuning curves for normal hearing [35], especially for monopolar stimulation, which is used in all present-day systems. Such broad and highly overlapping activation of the nerve most likely limits the number of perceptually separable channels to 4–8, even if more than eight electrodes are used, as also noted before. The information presented through the implant is limited to envelope variations in the 16 or fewer frequency bands for these patients. (Similar numbers apply for patients also achieving high scores but using processing strategies other than CIS.) For HR4 and others, the upper frequency of envelope variations has been set at 200–700 Hz [9], e.g., by using a cutoff frequency in the range of 200–700 Hz for the low-pass filters in the envelope detectors shown in Fig. 4. A substantial fraction of this information may be perceived by the better patients [36]–[38], and whatever is perceived is sufficient for high levels of speech recognition.

The performance achieved by HR4 and the others like him brings into question the significance for speech reception of the intricate processing, and the interplay between and among processing steps, that occur in the normal cochlea. The details of the traveling wave of mechanical displacements along the BM in response to acoustic stimuli [39], and the spatial sharpening of the membrane response by active processes at the outer hair cells (OHCs) [39], [40], are not necessary for effective representations of speech information. Also, the noninstantaneous compression function at the synapses between the IHCs and single fibers of the auditory nerve [41] is not necessary. Additional aspects of normal hearing that are not replicated with implants include multiple stages of compression (at the BM/OHC complex, at the IHCs, and at the IHC/neuron synapses); effects of efferent action on the OHCs and other structures in the cochlea [42]; the broad distributions of thresholds for the multiple afferent fibers innervating each IHC [43]; and effects of spontaneous activity in the nerve [44], which is absent or largely absent in the deafened ear [45]–[47]. Despite these many missing steps or severed connections, cochlear implants can restore clinically normal function in terms of speech reception for some patients. This is remarkable.

B. Variability in Outcomes

One of the major remaining problems with cochlear implants is the broad distribution of outcomes, especially for difficult tests and as exemplified in the bottom panel of Fig. 5. That is, patients using exactly the same implant system—with the same speech processor, transcutaneous link, implanted receiver/stimulator, and implanted electrode array—can have scores ranging from the floor to the ceiling for such tests. Indeed, only a small fraction of patients achieve the spectacularly high scores discussed above.

C. Likely Importance of Cortical Function

Accumulating and compelling evidence is pointing to differences in cortical or auditory pathway function as a likely contributor to the variability in outcomes with cochlear implants. On average, patients with short durations of deafness prior to their implants fare better than patients with long durations of deafness [48]. This may be the result of sensory deprivation for long periods, which adversely affects connections between and among neurons in the central auditory system [49] and may allow encroachment by other sensory inputs of cortical areas normally devoted to auditory processing (this encroachment is called “cross-modal plasticity,” see [50] and [51]). Although one might think that differences in nerve survival at the periphery could explain the variability, either a negative correlation or no relationship has been found between the number of surviving ganglion cells and prior word recognition scores, for deceased implant patients who in life had agreed to donate their temporal bones (containing the cochlea) for postmortem histological studies [52]–[55]. In some cases, survival of the ganglion cells was far shy of the normal complement, and yet these same patients achieved high scores in monosyllabic word tests. Conversely, in some other cases, survival of the ganglion cells was excellent, and yet these patients did not achieve high scores on the tests. Although some number of ganglion cells must be required for the function of a cochlear implant, this number appears to be small. Above that putative threshold, the brains of the better-performing patients apparently can utilize a sparse input from even a small number of surviving cells for high levels of speech reception.

Similarly, it seems likely that representation of speech sounds with the cochlear implant needs to be above some threshold in order for the brain to utilize the input for good speech reception. Single-channel implant systems did not rise above this second putative threshold; nor did prior processing strategies for multichannel implants. The combination of multiple sites of stimulation in the cochlea (at least 6–8), relatively new processing strategies such as the CIS, p-of-m, and ACE strategies, and some minimum survival of ganglion cells is sufficient for a high restoration of function in some patients. Those patients are likely to have intact “auditory brains” that can utilize these still sparse and distorted inputs, compared with the inputs the brain receives from the normal cochlea.

Other patients may not have the benefit of normal or nearly normal processing central to the auditory nerve. The effects of auditory deprivation for long periods have been mentioned. In addition, the brains of children become less “plastic” or adaptable to new inputs beyond their third or fourth birthdays. This may explain why deaf children implanted before then generally have much better outcomes than deaf children implanted at age five and older [50], [56], [57].

The brain may be the “tail that wags the dog” in determining outcomes with present-day cochlear implants. The brain “saves
us” in achieving high scores with those implants, in somehow utilizing a crude and sparse and distorted representation at the periphery. In addition, strong learning or accommodation effects—over long periods ranging from about three months to a year or more—indicate a principal role of the brain in reaching asymptotic performance with implants (see Fig. 6). Multiple lines of evidence further indicate or suggest that impairments in brain function—including damage to the auditory pathways in the brainstem, or compromised function in the areas of cortex normally devoted to auditory processing, or reduced cortical plasticity, or cross-modal plasticity—can produce highly deleterious effects on results obtained with cochlear implants.

D. Likely Importance of Electrode Designs

Present designs and placements of electrodes for cochlear implants do not support more than 4–8 effective sites of stimulation, or effective or functional channels, as described in Section III-C above. Contemporary cochlear implants use between 12 and 22 intracochlear electrodes, so the number of electrodes exceeds the number of effective channels (or sites of stimulation) for practically all patients and for all current devices. The number of effective channels depends on the patient and the speech reception measure to evaluate performance. For example, increases in scores with increases in the number of active electrodes generally plateau at a lower number for consonant identification than for vowel identification. (This makes sense from the perspective that consonants may be identified with combinations of temporal and spectral cues, whereas vowels are identified primarily or exclusively with spectral cues, that are conveyed through independent sites of stimulation.) Patients with low speech reception scores generally do not have more than four effective channels for any test, whereas patients with high scores may have as many as eight or slightly more channels depending on the test (e.g., [18] and [58]).

Results from studies using acoustic simulations of implant processors and subjects with normal hearing indicate that a higher number of effective channels or sites of stimulation for implants could be beneficial. Dorman et al. found, for example, that with the simulations and normal-hearing subjects, as many as ten channels are needed to reach asymptotic performance (for difficult tests) using a CIS-like processor [59]. Other investigators have found that even more channels are needed for asymptotic performance, especially for difficult tests such as identification of vowels or recognition of speech presented in competition with noise or a multi-talker babble [18], [60]. For example, Friesen et al. found that identification of vowels for listeners with normal hearing continued to improve with the addition of channels in the acoustic simulations up to the tested limit of 20 channels, for vowels presented in quiet and at progressively worse speech-to-noise ratios out to and including +5 dB [18].

Large improvements in the performance of cochlear implants might well be obtained with an increase in the number of effective sites of stimulation, which would help narrow the gap between implant patients and subjects with normal hearing. This gap is especially wide for the many patients who do not have more than four functional channels across wide ranges of speech reception measures. Just a few more channels for the top performers with implants would almost without doubt help them in listening to speech in demanding situations, such as speech presented in competition with noise or other talkers. An increase in the number of functional channels for patients presently at the low end of the performance spectrum could improve their outcomes substantially.

A highly plausible explanation for the limitation in effective channels with implants is that the electric fields from different intracochlear electrodes strongly overlap at the sites of neural excitation (e.g., [58] and [61]). Such overlaps (or electrode interactions) may well impose an upper bound on the number of electrodes that are sufficiently independent to convey perceptually separate channels of information. In addition, a central processing deficit may contribute to the limitation, perhaps especially for patients with low speech reception scores and (usually) a relatively low number of effective channels.

A problem with ST implants is that the electrodes are relatively far from the target tissue (the spiral ganglion), even for placements of electrodes next to the inner wall of the ST. Close apposition of the target and the electrode is necessary for a high spatial specificity of stimulation [62]. One possibility for providing a close apposition is to promote the growth of neurites from the ganglion cells toward the electrodes in the ST with controlled delivery of neurotrophic drugs into the perilymph [63]–[66]. Such growth of neurites would bring the target to the electrodes. Another possibility is to implant an array of electrodes directly within the auditory nerve (an intramodiolar implant), through an opening made in the basal part of the cochlea [24]–[26], [28]–[30]. In this case, the electrodes would be placed immediately adjacent to axons of the auditory nerve. Studies are underway to evaluate each of these possibilities, including safety and efficacy studies. Results from studies to evaluate the intramodiolar implant have demonstrated that it is feasible and that the number of independent sites of stimulation with that implant may be substantially higher than the number for ST implants [29], [30].

E. Recent Advances

Two recent advances in the design and performance of cochlear implants are: 1) electrical stimulation of both ears with bilateral cochlear implants and 2) combined electric and acoustic stimulation (EAS) of the auditory system for persons with residual hearing at low frequencies. Bilateral electrical stimulation may reinstate at least to some extent the interaural amplitude and timing difference cues that allow people with normal hearing to lateralize sounds in the horizontal plane and to selectively “hear out" a voice or other source of sound from among multiple sources at different locations. Additionally, stimulation on both sides may allow users to make use of the acoustic shadow cast by the head for sound sources off the midline. In such cases, the S/N may well be more favorable at one ear compared with the other for multiple sources of sound, and users may be able to attend to the ear with the better S/N. Combined EAS may preserve a relatively normal hearing ability at low frequencies, with excellent frequency resolution and other attributes of normal hearing, while providing a complementary representation of high-frequency sounds with the
cochlear implant and electrical stimulation. Various surgical techniques and drug therapies have been developed to preserve low-frequency hearing in an implanted cochlea, including deliberately shallow insertions of the electrode array (6, 10, or 20 mm) so as not to damage the apical part of the cochlea and remaining hair cells there; insertion of the electrode array through the round window membrane rather than through a cochleostomy to eliminate deleterious effects of drilling (loud and possibly damaging levels of noise, introduction of blood and bone dust into the perilymph, possible damage to delicate cochlear structures such as the BM); use of “soft surgery” techniques to minimize trauma; use of thin and highly flexible electrodes; use of a lubricant such as hyaluronic acid to facilitate insertion of the array; and use of corticosteroids and other drugs to help preserve cochlear structures in the face of surgical manipulations and the introduction of a foreign body into the inner ear. Moderate-to-excellent preservation of residual hearing has been reported using the shallow insertions and some or all of the additional procedures and techniques just mentioned [67]–[80]. Among the tested methods, insertion through the round window for placement of 20 mm arrays or use of shorter arrays have produced especially good results [77], [80]–[82]. The “soft surgery” methods also have been identified as important (e.g., [76] and [83]). Studies aimed at the further development of surgical techniques, adjunctive drug therapies, and special electrode arrays are in progress; both short- and long-term preservation of residual hearing in an implanted cochlea remain as major challenges and concerns.

Each of these approaches—bilateral electrical stimulation and combined EAS—has produced large improvements in speech reception performance compared with control conditions. In particular, bilateral stimulation can provide a substantial benefit in recognizing difficult speech materials such as monosyllabic words and in recognizing speech presented in competition with spatially distinct noise, in comparison to scores obtained with either unilateral implant alone [38], [84]–[100]. In addition, use of both implants supports an improved ability to lateralize sounds, again compared with either unilateral implant [86], [89], [90], [92], [93], [97], [98], [100]–[105]. (This ability is nonexistent or almost nil with a unilateral implant.) Combined EAS also provides a substantial benefit for listening to speech in noise or in competition with a multi-talker babble, compared with either electric stimulation only or acoustic stimulation only [38], [67], [68], [70], [71], [73], [75]–[79], [106]–[110]. Indeed, in some cases the score for combined EAS is greater than the sum of the scores for the electric-only and acoustic-only conditions. This has been described as a synergistic effect [38], [71], [75], [106], [111]. In addition, identification of melodies and reception of musical sounds is greatly improved with combined EAS compared with electric stimulation alone [73], [78], [109], [112], [113]. (Scores with acoustic stimulation alone closely approximate the scores with combined EAS, for melody and music reception.)

These gains from bilateral electrical stimulation most likely arise from a partial or full restoration of the binaural difference cues and to the head shadow effect, as suggested above. In addition, gains may result from a “binaural summation” effect that is produced in normal hearing by redundant stimulation on the two sides. Detailed descriptions of these various contributors to an overall binaural benefit for normal hearing and possible contributors for prosthetic hearing are presented in [38]. The evidence to date indicates that almost all recipients of bilateral cochlear implants benefit from the head shadow effect and that some benefit from: 1) the binaural squelch effect that is made possible with presentation and perception of the binaural timing-difference cue; 2) the binaural summation effect; or 3) both.

In addition to these binaural effects that occur in normal hearing and to a variable extent in prosthetic hearing, electric stimulation on both sides may help fill “gaps” in the representation of frequencies on one side due to uneven survival of spiral ganglion cells along the cochlear spiral—with complementary excitation of surviving neurons at the same frequency place on the contralateral side. For example, a lack of input to the central nervous system (CNS) at the 5 kHz position on one side may be at least partly bridged or compensated by stimulation of remaining neurons at the 5 kHz position in the other ear. This mechanism and the binaural summation effect may underlie the large improvements observed with bilateral implants for the recognition of difficult speech material presented from in front of the subjects and without any interfering noise, where the interaural difference cues and the head shadow effect do not come into play. The mechanism also may contribute to the good results observed for other conditions, in which the difference cues and the head shadow effect are also present.

A further possible mechanism contributing to the observed benefits of bilateral electric stimulation is a higher number of effective channels. Bilateral implants, in general, provide a doubling or near doubling of physical stimulus sites, compared with either unilateral implant alone. This may provide some gain in the number of effective channels, especially in cases of uneven nerve survival across the two sides, where stimulation of an area on one side that is “dead” on the other side may add an effective channel. As noted before, even a small gain in the number of effective channels could produce a large benefit, particularly for patients who otherwise would have low levels of performance and particularly for reception of difficult speech materials or for listening to speech in adverse S/N conditions.

An example of findings from studies with recipients of bilateral implants is presented in Fig. 8. These results are from studies conducted by Müller and coworkers at the Julius–Maximilians Universität in Würzburg, Germany [87]. Nine subjects participated. The left and middle columns show individual and average scores for the recognition of sentences presented in competition with speech-spectrum noise at the S/N of +10 dB and with the sentences presented through a loudspeaker in front of the subject and the noise presented through a loudspeaker to the right of the subject (left column) or to the left of the subject (middle column). The right column shows results for the recognition of monosyllabic words in quiet, presented from the loudspeaker in front of the subject. For the sentence tests, the difference in scores for the left
implant only versus the right implant only shows the magnitude of the head shadow benefit, which is large (see lower-left and lower-middle panels). For these same tests, the difference in scores for the bilateral condition versus the score for the single implant at the side opposite to the noise source shows the magnitude of a “binaural processing benefit,” which is a combination of binaural squelch, binaural summation, and possibly other effects. This binaural processing benefit is smaller than the head shadow benefit but still significant. For the word test (right column), the difference in scores between the bilateral condition and either of the unilateral conditions may be attributable to a binaural summation effect, or a filling of gaps in nerve survival across the two sides, or a principal contribution from the better of the two ears, or a higher number of effective channels, or some combination of these, for the bilateral condition. The improvement obtained with stimulation on both sides is large, comparable to the head shadow benefits demonstrated by the results from the sentence tests. This improvement is larger than what would be expected from binaural summation effects alone.

The gains from combined EAS may arise from a normal or nearly normal input to the CNS for low-frequency sounds from the acoustic stimulation, in conjunction with a crude representation of high-frequency sounds from the electric stimulation with a partially inserted cochlear implant. The CNS apparently is able to integrate these seemingly disparate inputs into a single auditory percept, that is judged as sounding natural and intelligible. The likely ability to separate different “auditory streams” on the basis of different fundamental frequencies (and trajectories of fundamental frequencies) for different sounds may at least in part underlie the large advantages produced with combined EAS compared with electric stimulation only [77], [108], [109], [114], [115]. In particular, these fundamental frequencies (and one or more of their first several harmonics) occur at low frequencies and are within the range of residual hearing for most if not all users of combined EAS, i.e., below 500–1000 Hz. Perception and utilization of fine frequency differences in this range may allow an effective separation of a signal from interfering sounds. Also, the likely ability to “track” low frequencies almost certainly underlies the large improvements in melody recognition and music reception that have been reported (e.g., [113]).

Each of these relatively new approaches utilizes or reinstates a part of the natural system. Two ears are better than one, and
use of even a part of normal or nearly normal hearing at low frequencies can provide a highly significant advantage.

F. Possibilities for Further Improvements

Tremendous progress has been made in the design and performance of cochlear prostheses. However, much room remains for improvements. Patients with the best results still do not hear as well as listeners with normal hearing, particularly in demanding situations such as speech presented in competition with noise or other talkers. Users of standard unilateral implants do not have much access to music and other sounds that are more complex than speech. Most importantly, speech reception scores still vary widely across patients for relatively difficult tests, such as recognition of monosyllabic words, with any of the implant systems now in widespread use.

Fortunately, major steps forward have been made recently and many other possibilities for further improvements in implant design and function are on the horizon. Electrical stimulation on both sides with bilateral cochlear implants, and combined EAS for persons with some residual hearing, have been mentioned. These are new approaches, which may well be refined or optimized for still higher levels of performance. Some of the possibilities for such improvements are just now being explored, including development and evaluation of surgical techniques and adjunctive therapies aimed at the preservation of residual hearing in an implanted cochlea. In addition, other approaches—such as reinstatement of spontaneous-like activity in the auditory nerve [116], representation of “fine structure” or “fine frequency” information with novel patterns of electric stimuli [117]–[119], or a closer mimicking of the processing that occurs in the normal cochlea [118], [120]—may also produce improvements in performance, especially for patients with good or relatively good function in the central auditory pathways and in the cortical areas that process auditory information.

Further improvements for all patients might be produced by somehow increasing the number of effective channels supported by cochlear implants. Several possibilities for this have been mentioned, including intramodiolar implants and drug-induced growth of neurites toward the electrodes of ST implants. An additional possibility is to regard bilateral implants as a collection of many stimulus sites and to choose for activation the perceptually separable sites among them. Alternatively, one might “interlace” stimulus sites across the two sides, where the most basal region of one cochlea is stimulated on one side, the next most basal region on the other side, the next most basal region on the first side, and so forth until the full tonotopic map is spanned. In this way, all the frequencies would be represented but the distance between active electrodes in each implant would be doubled, which would in turn reduce the interactions among them, compared with stimulation of adjacent electrodes. These different ways of using bilateral implants have the potential to increase the number of effective channels [38], [121], but almost certainly at the cost of diminishing or eliminating a useful representation of the binaural difference cues. This may be a good tradeoff for some patients.

Each of the approaches described above is aimed at improving the representation at the periphery. A fundamentally new approach may be needed to help those patients presently at the low end of the performance spectrum, however. They may have compromised “auditory brains” as suggested above and by many recent findings. For them, a “top-down” or “cognitive neuroscience” approach to implant design may be more effective than the traditional “bottom-up” approach. In particular, the new (top-down) approach would ask what the compromised brain needs as an input in order to perform optimally, in contrast to the traditional approach of replicating insofar as possible the normal patterns of activity at the auditory nerve. The patterns of stimulation specified by the new approach are quite likely to be different from the patterns specified by the traditional approach.

A related possibility that may help all patients at least to some extent is directed training to encourage and facilitate desired plastic changes in brain function (or, to put it another way, to help the brain in its task to learn how to utilize the inputs from the periphery provided by a cochlear implant). Such training if well designed may shorten the time needed to reach asymptotic performance and may produce higher levels of auditory function at that point and beyond. The ideal training procedure for an infant or young child may be quite different from the ideal procedure for older children or adults due to differences in brain plasticity. For example, the “step size” for increments in the difficulty of a training task may need to be much smaller for adults than for infants and young children [122]. However, all patients may benefit from appropriately designed procedures, that respect the differences in brain plasticity according to age.

The brain is a critical part of a prosthesis system. For patients with a fully intact brain, the “bottom-up” approach to implant design probably is appropriate, i.e., an ever-closer approximation to the normal patterns of neural discharge at the periphery is likely to provide the inputs that the brain “expects” and is configured to receive and process. For patients with a compromised brain, such inputs may not be optimal. In those cases, a “top-down” approach to implant design, or a combination of “top-down” and “bottom-up” approaches, may produce the best results. For example, a “top-down” approach combined with techniques to minimize electrode interactions at the periphery may be especially effective for patients presently shackled with poor outcomes.

VI. INTERFACING SENSORS WITH THE NERVOUS SYSTEM

The full restoration of clinically normal function with a cochlear implant, as demonstrated by the findings for subject HR4 (Fig. 7) and others like him, bodes well for the development of other types of sensory neural prostheses. In particular, a sparse and distorted representation at the periphery may be sufficient for restoration of high levels of function for other sensory inputs as well, e.g., visual or vestibular inputs. As with cochlear implants, a putative threshold of the amount and quality of information in the peripheral representation may need to be exceeded before good outcomes can be achieved. However, this threshold may be quite low and a full replication of the exquisite and complex machinery at the periphery is certainly not necessary for the restoration of useful hearing and may not be necessary for the restoration of other senses either.
That said, reproduction of some aspects of the normal physiology is likely to be important. In cochlear implants, for example, a crude replication of the normal tonotopic representation of frequencies—with multichannel processing strategies and with multiple (and perceptually separable) sites of stimulation in the cochlea—was necessary to achieve high levels of performance. Perhaps a topographic representation would work well for a visual prosthesis, as has been suggested (e.g., [123]–[133]). As with cochlear implants, we expect some threshold of resolution in the stimulation will need to be exceeded for good function, and that the difficult problems of electrode interactions will need to be addressed for useful restoration of vision and other senses. However, the threshold may be surprisingly low. (A low threshold may be essential for a successful visual prosthesis, as the optic nerve has 1.2 million ganglion cells and associated axons, that receive inputs from 125 million photoreceptors in the retina. These numbers are substantially higher than the corresponding numbers for the cochlea, e.g., 1.2 million neurons in the optic nerve versus 30,000 neurons in the auditory nerve. The complexity of the retina and strategies for electrical stimulation using epiretinal or subretinal arrays, or electrical stimulation at more central sites in the visual pathway, are discussed in [123]–[127], [131], and [134]–[137].)

In addition, an intact or largely intact brain may well be a prerequisite for a topographic representation to work, at least initially and without training. Further, effects of cross-modal plasticity may preclude a good outcome with any type of sensory neural prosthesis, although a training approach has been proposed to mitigate or even possibly reverse these effects [125], [126].

An important consideration in the design of sensory neural prostheses is to regard the brain as a key part of the overall system. The brain of the user should be respected for what it does, and the design should foster a partnership between the brain and the prosthesis, perhaps with communications in both directions, i.e., from the brain to the prosthesis as well as from the prosthesis to the brain. Indeed, this was a principal theme of the Smart Prosthetics conference held at the Beckman Center, University of California, Irvine, in November 2006 and sponsored by the Keck Foundation and National Academies Futures Initiative, see http://www.keckfutures.org/ and the daughter pages. We expect this more holistic approach will be embraced in future designs.

The path between a sensor or an array of sensors and useful perception involves many steps and considerations. The path can be traversed, though, as demonstrated by cochlear implants.

VII. SUMMARY

In summary, the experience with cochlear implants either indicates or suggests the following.

- Experts can be stunningly wrong in assessments of a new approach or technology; perseverance in the face of intense criticism was essential for the successful development of cochlear implants and this may prove to be the case for other types of neural prostheses as well.
- The above is not an argument for wayward or uninformed efforts, of course, and the NIH vetted cochlear implants with the Bilger study [4] before investing many millions of dollars for the further development of implant systems. Still, though, the courage to take informed risks on the part of the NIH and the investigators (including the investigators worldwide who were supported by agencies other than the NIH) was as important as anything else in moving this marvelous technology forward; in addition, some of the earlier efforts that appeared to many to be wayward at the time later proved to be prescient.
- Multidisciplinary efforts of multiple teams were required to make the cochlear implant a success, and NIH support of a large number of these efforts was critically important. Development of other types of neural prostheses without these elements in place seems unimaginable.
- A decidedly sparse and crude and distorted representation at the periphery supports a remarkable restoration of function for some users of present-day cochlear implants. This bodes well for the development of vestibular, visual, or other types of sensory neural prostheses.
- However, this representation must exceed some putative threshold of quality and quantity of information. Most likely, this means that aspects of the normal physiology need to be mimicked or reinstated to some minimal extent.
- The experience with cochlear implants indicates that 1) not all aspects of the normal physiology need to be reproduced and 2) those aspects that are reinstated do not have to be perfectly reproduced by any means. Present-day implants—with multiple channels of processing, multiple sites of stimulation in the cochlea, and the CIS, n-of-m, ACE, or other modern processing strategies—have exceeded the putative threshold for the great majority of patients, in that most patients score at 80% correct or higher in sentence tests using hearing alone and many patients can use the telephone without difficulty. Prior implant systems did not exceed the threshold.
- Not surprisingly, the interface to the tissue is important. Present electrode arrays for cochlear implants do not support more than 4–8 functional channels even though the number of stimulating electrodes is higher than that. Overlapping excitation fields from different electrodes almost certainly degrade their independence; this is a general problem with neural prostheses that map outputs to thousands of neurons in very close proximity to each other, as in the retina.
- Interlacing of stimulus pulses across electrodes—such that only one electrode is active at any one time—has proved to be highly effective for cochlear implants in achieving the present levels of electrode and channel independence. Such interlacing of stimuli may be effective for other types of neural prostheses. In addition, novel electrode designs, placements of electrodes in close proximity to the target neurons, drug treatments to encourage the growth of neural tissue toward electrodes, or interlacing of stimuli across bilateral implants (e.g., across implants for each retina), or combinations of these, may well increase the number of functional sites of stimulation for cochlear, as well as other types of sensory neural prostheses.
- Any residual function should be preserved and utilized to the maximum extent possible, in conjunction with the
prosthesis, as in combined electric and acoustic stimulation of the auditory system for persons with some residual (low-frequency) hearing.

• For sensory systems with bilateral inputs—audition, vision, and balance—reinstatement of inputs on both sides may confer large benefits to users of prosthetic systems, as demonstrated by the experience with bilateral cochlear implants.

• Percutaneous access to the implanted electrodes may provide important advantages in the initial development of any sensory neural prosthesis, as certainly was the case with cochlear implants.

• Good results take time. Asymptotic performance is not achieved with cochlear implants until at least three months of daily use and in many cases longer or much longer than that. This and other findings indicate a principal role of the brain in determining outcomes with implants. It also indicates that results from acute studies may be misleading in that they may grossly underestimate the potential of an approach. The brain is likely to be vitally important in determining outcomes with other types of neural prostheses as well, and effects of cross-modal plasticity may preclude good outcomes for persons who have been deprived of a sensory input for all or most of their lives, in that the “cortical target” for the input has been encroached or recruited by other sensory modalities and such effects may not be reversible. (Such effects have not been reversed thus far. However, this does not mean categorically that the task cannot be achieved. Merabet and coworkers have, for example, proposed a yet-to-be-tested training procedure to mitigate or reverse effects of cross-modal plasticity in the context of a visual prosthesis, see [126].)

• The power of the intact or largely intact brain to utilize sparse and distorted inputs is impressive; and this most likely underpins in large part the success of cochlear implants.

• A sensory prosthesis and the brain are “partners” in an overall system, and simply focusing on the periphery in the design of a prosthesis may provide good results for persons with fully intact brains and sensory pathways, but probably will limit results for persons with impaired pathways or impaired or altered cortical processing.

• The amount of information from the periphery that can be utilized may be increased through plastic changes in the brain, especially for infants and very young children but also for older patients, albeit at a likely slower pace of adaptation and perhaps to a lesser extent than with young children.

• Desired plastic changes may be facilitated and augmented through directed training; the optimal training procedure is likely to vary according to the age of the patient, to the duration of sensory deprivation prior to the restoration of (some) function with a prosthesis, and whether or not the sense was first lost prior to the “critical period” for the normal development of that sensory pathway and processing in the midbrain and cortex. Training may or may not be effective for patients who lost a sense prior to or during the critical period and had it reinstated (at least to some extent) after the critical period had expired. Training may be most effective for persons who lost the sense following the critical period, and after the sensory pathways and associated cortical processing had been established.

• The highly deleterious effects of cross-modal plasticity or missing the critical period for maturation of the central auditory pathways and cortex are “moral imperatives” to screen infants for deafness or blindness or possibly other sensory losses and to provide at least some input to the appropriate part of the CNS if feasible and as soon as practicable for cases in which severe deficits are found.

• Cochlear implants are among the great success stories of modern medicine, and this has surprised many. Another surprise, with the development of another highly effective sensory neural prosthesis, is certainly possible.

ACKNOWLEDGMENT

The authors are grateful to the three anonymous reviewers of the present paper for their exceptionally thoughtful and constructive comments. Limited material also was drawn or adapted from several recent publications, [9], [33], [57].

DEDICATION: This paper is dedicated to F. T. Hambrecht, M.D., and W. J. Heetderks, M.D., Ph.D., whose vision, leadership, and scientific acumen made present-day neural prostheses possible.

REFERENCES


Blake S. Wilson (M’80–SM’06) recently retired from RTI International following 33 years of continuous service and has become its first Emeritus Senior Fellow. He also is an Adjunct Professor at the Duke University Medical Center; the Chief Strategy Advisor for Med El GmbH of Innsbruck, Austria; and the Overseas Expert for a large project at the International Center of Hearing and Speech in Kajetany (near Warsaw), Poland, to improve treatments of hearing loss. These are all ongoing positions. He is the inventor of most of the speech processing strategies used with present-day cochlear implants, including the continuous interleaved sampling (CIS), spectral peak picking (e.g., ”n-of-m”), and virtual channel strategies, among others. The CIS and n-of-m strategies, or direct descendants of them, are used as the default strategies for all three implant systems now in widespread use. One of his papers, in the journal Nature, alternates with one other paper as the most highly cited publication in the field of cochlear implants. He has served as the Principal Investigator for 24 projects, including 13 projects for the National Institutes of Health. He also served as the Director of the Center for Auditory Prosthesis Research at RTI from its inception and for more years thereafter, until he was appointed as one of RTI’s first four Senior Fellows in 2002.

Dr. Wilson and his coworkers have been recognized by many awards and honors; most notably the 1996 Discover Award for Technological Innovation and the American Otolaryngology Society’s President’s Citation for “Major contributions to the restoration of hearing in profoundly deaf persons.” He has been the Guest of Honor at ten international conferences, and has been a keynote or invited speaker at more than 130 others. He has served as the Chair for two large international conferences and as the Co-Chair for two others. Most recently, he received the 2007 Distinguished Alumnus Award from the Pratt School of Engineering at Duke.

Michael F. Dorman received the Ph.D. degree in experimental child and developmental psychology (Linguistics minor) from the University of Connecticut, Storrs, in 1971. A Fellow of the Acoustical Society of America, he is currently a Professor in the Department of Speech and Hearing Science and the Program in Linguistics at Arizona State University. He is the author of over 100 publications in areas including speech perception by infants, adults, hearing-impaired listeners and listeners fit with cochlear implants, and also cortical lateralization of function and neural plasticity. His research has been supported by the National Institutes of Health since 1973.
Interfacing Sensors With the Nervous System: Lessons From the Development and Success of the Cochlear Implant

Blake S. Wilson, Senior Member, IEEE, and Michael F. Dorman

Invited Paper

Abstract—The cochlear implant is the most successful neural prosthesis to date and may serve as a paradigm for the development or further development of other systems to interface sensors with the nervous system, e.g., visual or vestibular prostheses. This paper traces the history of cochlear implants and describes how the current levels of performance have been achieved. Lessons and insights from this experience are presented in concluding sections.

Index Terms—Auditory prosthesis, cochlear implant, deafness, hearing, nervous system, neural prosthesis, sensors, speech perception, vestibular prosthesis, visual prosthesis.

I. INTRODUCTION

The COCHLEAR implant is one of the great success stories of modern medicine. Just 30 years ago, cochlear implants provided little more than a sensation of sound and sound cadences. They were useful as an aid to lipreading. Now, a majority of implant users enjoy high levels of speech recognition using hearing alone; indeed, many can use the telephone without difficulty. This is a long trip in a short time, and the restoration of function—from total or nearly total deafness to useful hearing—is truly remarkable.

In this paper, we trace this history and indicate how the present levels of performance have been achieved. The design of cochlear implants is described in some detail to provide an example of ways in which sensors can be successfully interfaced to the nervous system. Results from studies with implant patients are presented. In addition, we describe some of the limitations of present systems and possibilities for overcoming them. We conclude with a section on the lessons learned from cochlear implants and how those lessons might inform the designs of other types of sensory neural prostheses, such as prostheses for the restoration of vision or balance.

II. A BRIEF HISTORY

As recently as the early 1980s, many eminent and highly knowledgeable people believed that cochlear implants would provide only an awareness of environmental sounds and possibly speech cadences to their users. Many were skeptical of implants and thought that mimicking or reinstating the function of the exquisite machinery in the normal inner ear was a fool’s dream. Among these critics were world-renowned experts in otology and auditory physiology. Fortunately, pioneers persisted in the face of this intense criticism and provided the foundations for present devices. Detailed reviews of the early history of cochlear implants are presented in [1]–[3].

A timeline of assessments in the development of cochlear implants is presented in Table I. These range from frank skepticism at the beginning to high enthusiasm by 1995.

The first implant of a device for electrical stimulation of the auditory nerve was performed by Djourno and Eyrès in Paris in 1957. An induction coil was used, with one end placed on the stump of the auditory nerve or adjacent brainstem and the other end within the temporalis muscle (the patient had had bilateral cholesteatomas which had been removed in prior operations, taking the cochleas and peripheral parts of the auditory nerves with them). The patient used the device for several months before it failed, and was able to sense the presence of environmental sounds but could not understand speech or discriminate among speakers or many sounds. In 1961, Dr. William F. House implanted two patients in Los Angeles, each with single gold wires inserted a short distance into the (deaf) cochlea. By 1975, more patients had been implanted worldwide, most by Dr. House, and 13 had functioning, single-channel devices. The United States National Institutes of Health (NIH) commissioned a study at that point, to assess the performance of those devices and to determine whether support by the NIH for the further development of cochlear implants would be wise. The report from the study [4], the “Bilger report,” is a landmark in the field. Its key conclusion was that “although the subjects could not understand speech through their prostheses, they did score significantly higher on tests of lipreading and recognition of environmental sounds with their prostheses activated than without them.” This and earlier assessments are included in Table I.
Shortly after the Bilger report was published, the NIH did elect to support research and development efforts in the field. The rapid progress thereafter in the design and performance of implant systems was in very large part the direct result of this decision. In particular, work supported through the Neural Prosthesis Program at the NIH, first directed by Dr. F. Terry Hambrecht and later by Dr. William J. Heetderks, produced many important innovations in electrode and speech processor designs that remain in use to this day.

In 1988, NIH convened the first of two consensus development conferences on cochlear implants. Multichannel systems—with multiple channels of processing and with multiple sites of stimulation in the cochlea—had come into use at that time. The consensus statement from the 1988 conference [5] suggested that multichannel implants were more likely to be effective than single-channel implants, and indicated that about 1 in 20 patients could carry out a normal conversation without lipreading. Approximately 3000 patients had received cochlear implants by 1988.

New and highly effective processing strategies for cochlear implants were developed in the late 1980s and early 1990s, principally through the Neural Prosthesis Program. Among these were the continuous interleaved sampling (CIS) [6], n-of-m [7], and spectral peak (SPEAK) [8] strategies. Large gains in speech reception performance were achieved with these strategies, two of which remain in widespread use today (CIS and n-of-m). A detailed review of processing strategies and their lines of development is presented in [9].

The second NIH consensus development conference was held in 1995. By then, approximately 12,000 patients had received implants. A major conclusion from the 1995 conference [10] was that “a majority of those individuals with the latest speech processors for their implants will score above 80% correct on high-context sentences even without visual cues.” By the middle of 2006, the cumulative number of implants worldwide exceeded 110,000. This number is orders of magnitude higher than the numbers for all other types of neural pros- theses, including those for restoration of motor or other sensory functions.

Fig. 1 shows the number of cochlear implants over time, beginning in 1957 with the first implant operation by Djourno and Eyriès. The growth in numbers since then is exponential.

III. DESIGN OF COCHLEAR IMPLANTS

A. Aspects of Normal Hearing

In normal hearing, sound waves traveling through air reach the tympanic membrane via the ear canal, causing vibrations that move the three small bones of the middle ear. This action produces a piston-like movement of the stapes, the third
bone in the chain. The “footplate” of the stapes is attached to a flexible membrane in the bony shell of the cochlea called the oval window. Inward and outward movements of this membrane induce pressure oscillations in the cochlear fluids, which in turn initiate a traveling wave of displacement along the basilar membrane (BM), a highly specialized structure that divides the cochlea along its length. This membrane has graded mechanical properties. At the base of the cochlea, near the stapes and oval window, it is narrow and stiff. At the other end of the cochlea, near the apex, the membrane is wide and flexible. These properties give rise to the traveling wave and to points of maximal response according to the frequency or frequencies of the pressure oscillations in the cochlear fluids. The traveling wave propagates from the base to the apex. For an oscillation with a single frequency, the magnitude of displacements increases up to a particular point along the membrane and then drops precipitously thereafter. High frequencies produce maxima near the base of the cochlea, whereas low frequencies produce maxima near the apex.

Motion of the BM is sensed by the inner hair cells (IHCs) in the cochlea, which are attached to the top of the BM in a matrix of cells called the organ of Corti. Each hair cell has fine rods of protein, called stereocilia, emerging from one end. When the BM moves at the location of a hair cell, the rods are deflected as if hinged at their bases. Such deflections in one direction increase the release of chemical transmitter substance at the base (other end) of the cells, and deflections in the other direction inhibit the release. The variations in the concentration of the chemical transmitter substance act at the terminal ends of auditory neurons, which are immediately adjacent to the bases of the IHCs. Increases in chemical transmitter substance increase discharge activity in the nearby neurons, whereas decrements in the substance inhibit activity. Changes in neural activity thus reflect events at the BM. These changes are transmitted to the brain via the auditory nerve, the collection of all neurons that innervate the cochlea.

The steps described above are illustrated in the top panel of Fig. 2. This shows a cartoon of the main anatomical structures, including the tympanic membrane, the three bones of the middle ear, the oval window, the BM, the IHCs, and the adjacent neurons of the auditory nerve.

B. Loss of Hearing

The principal cause of hearing loss is damage to or complete destruction of the sensory hair cells. Unfortunately, the hair cells are fragile structures and are subject to a wide variety of insults, including but not limited to genetic defects, infectious diseases (e.g., rubella and meningitis), overexposure to loud sounds, certain drugs (e.g., kanamycin, streptomycin, and cisplatin), and aging. In the deaf or deafened cochlea, the hair cells are largely or completely absent, severing the connection between the peripheral and central auditory systems. The function of a cochlear prosthesis is to bypass the (missing) hair cells by stimulating directly the surviving neurons in the auditory nerve.

The anatomical situation faced by designers of cochlear implants is illustrated in the bottom panel of Fig. 2. The panel shows a complete absence of hair cells. In general, a small number of cells may remain for some patients, usually in the apical (low frequency) part of the cochlea.

Without the normal stimulation provided by the hair cells, the peripheral part of the neurons—between the cell bodies in the spiral ganglion and the terminals within the organ of Corti—undergo “retrograde degeneration” and eventually die [11]. Fortunately, the cell bodies are far more robust. At least some usually survive, even for prolonged deafness or for virulent etiologies such as meningitis [11]–[13]. These cells, or more specifically the nodes of Ranvier just distal or proximal to them, are the putative sites of excitation for cochlear implants.

C. Electrical Stimulation of the Auditory Nerve

Direct stimulation of the nerve is produced by currents delivered through electrodes placed in the scala tympani (ST), one of three fluid-filled chambers along the length of the cochlea. (The boundary between the ST and the scala media is formed by the BM and organ of Corti.) A cutaway drawing of the implanted cochlea is presented in Fig. 3. The figure shows a partial insertion of an array of electrodes into the ST. The array is inserted through a drilled opening made by the surgeon in the bony shell of the cochlea overlying the ST (called a “cochleostomy”) and close to the base of the cochlea. Alternatively, the array may be...
inserted through the second flexible membrane of the cochlea, the round window membrane, which also is close to the basal end of the cochlea and ST (see drawing).

The depth of insertion is limited by the decreasing lumen of the ST from base to apex, the curvature of the cochlear spiral, and an uneven and unsmooth lumen particularly in the apical region. No array has been inserted farther than about 30 mm, and typical insertions are much less than that, e.g., 18–26 mm. (The total length of the typical human cochlea is about 35 mm.) In some cases, only shallow insertions are possible, such as when bony obstructions in the lumen impede further insertion.

Different electrodes in the implanted array may stimulate different subpopulations of neurons. As described above, neurons at different positions along the length of the cochlea respond to different frequencies of acoustic stimulation in normal hearing. Implant systems attempt to mimic or reproduce this "tonotopic" encoding by stimulating basally situated electrodes (first turn of the cochlea and lower part of the drawing) to indicate the presence of high-frequency sounds, and by stimulating electrodes at more apical positions (deeper into the ST and ascending along the first and second turns in the drawing) to indicate the presence of sounds with lower frequencies. Closely spaced pairs of bipolar electrodes are illustrated here, but arrays of single electrodes that are each referenced to a remote electrode outside the cochlea also may be used. This latter arrangement is called a "monopolar coupling configuration" and is used in all present-day implant systems that are widely applied worldwide. (There are three such systems and they constitute more than 99% of the cochlear implant market.)

The spatial specificity of stimulation with a ST electrode most likely depends on a variety of factors, including the orientation and geometric arrangement of the electrodes, the proximity of the electrodes to the target neural structures, and the condition of the implanted cochlea in terms of nerve survival and ossification. An important goal of electrode design is to maximize the number of largely nonoverlapping populations of neurons that can be addressed with the electrode array. Present evidence suggests, however, that no more than 4–8 independent sites may be available using current designs, even for arrays with as many as 22 electrodes [14–19]. Most likely, the number of independent sites is limited by substantial overlaps in the electric fields from adjacent (and more distant) electrodes. The overlaps are unavoidable for electrode placements in the ST, as the electrodes are sitting in the highly conductive fluid of the perilymph and additionally are relatively far away from the target neural tissue in the spiral ganglion. A closer apposition of the electrodes next to the inner wall of the ST would move them a bit closer to the target cells (see Fig. 3), and such placements have been shown in some cases to produce an improvement in the spatial specificity of stimulation [20]. However, a large gain in the number of independent sites may well require a fundamentally new type of electrode, or a fundamentally different placement of electrodes. The many issues related to electrode design, along with prospects for the future, are discussed in [20]–[30].

Fig. 3 shows a complete presence of hair cells (in the labeled organ of Corti) and a pristine survival of cochlear neurons. However, the number of hair cells is zero or close to it in cases of total deafness. In addition, survival of neural processes peripheral to the ganglion cells (the "dendrites") is rare in the deafened cochlea, as noted before. Survival of the ganglion cells and central processes (the axons) ranges from sparse to substantial. The pattern of survival is in general not uniform, with reduced or sharply reduced counts of cells in certain regions of the cochlea. In all, the neural substrate or target for a cochlear implant can be quite different from one patient to the next. A detailed review of these observations and issues is presented in [13].

D. Components of Cochlear Implant Systems

The essential components in a cochlear prosthesis include: 1) a microphone for sensing sound in the environment; 2) a speech processor to transform the microphone input into a set of stimuli for the implanted array of electrodes; 3) a transcutaneous link for the transmission of power and stimulus information across the skin; 4) an implanted receiver/stimulator to decode the information received from the radio-frequency signal produced by an external coil and then to generate stimuli using the instructions obtained from the decoded information; 5) a cable to connect the outputs of the receiver/stimulator to the electrodes; and 6) the array of electrodes. These components must work together as a system to support excellent performance and a weakness in a component can degrade performance significantly. For example, a limitation in the data bandwidth of the transcutaneous link can restrict the types and rates of stimuli that can be specified by the external speech processor and this, in turn, can limit performance. A thorough discussion of considerations for the design of cochlear prostheses and their constituent parts is presented in [27].

We note that an earlier implant system, the Ineraid® device, had a percutaneous connector rather than a transcutaneous link. In addition, several experimental implant systems included percutaneous connectors. Although use of these through-the-skin connectors increased the risk of infection, they also provided...
direct electrical access to the implanted electrodes from an external speech processor or other stimulating or recording equipment. This access allowed full stimulus control and high-fidelity recordings of intracochlear evoked potentials. A wide variety of speech processing strategies was evaluated with subjects having percutaneous connectors. This was vital for the development of strategies now in widespread use and for the acquisition of knowledge about the stimulus-response properties of the electrically stimulated auditory nerve in humans.

E. Transformation of a Microphone Input Into Stimuli for the Implant

An important aspect of the design for any type of sensory neural prosthesis is how to transform an input from a sensor or array of sensors into a set of stimuli that can be interpreted by the nervous system. The stimuli can be electrical or tactile, for examples, and usually involve multiple sites of stimulation, corresponding to the spatial mapping of inputs and representations of those inputs in the nervous system. One approach to the transformation – and probably the most effective approach – is to mimic or replicate at least to some extent the damaged or missing physiological functions that are bypassed or replaced by the prosthesis.

Of course, limitations in other parts of the prosthesis system may restrict what can be done with the transformation. Effects of limitations in the bandwidth of the transcutaneous link for cochlear implant systems have been mentioned. Also, a lack of independence among stimulus sites can greatly reduce the number of channels of information that can be conveyed to the nervous system. In such cases, a high number of channels in processing the input(s) from the sensor(s) would not in general produce any benefit and might even degrade performance.

For cochlear implants, this part of the design is called the processing strategy. As noted previously, advances in processing strategies have produced quite large improvements in the speech reception performance of implant patients, from recognition of a tiny percentage of monosyllabic words with the first strategies and multisite stimulation, for example, to recognition of a high percentage of the words with current strategies and multisite stimulation.

One of the simpler approaches supporting the recent levels of performance with implants is illustrated in Fig. 4. This is the CIS strategy, which is used as the default strategy or as a processing option in all implant systems now in widespread clinical use.

The CIS strategy filters speech or other input sounds into bands of frequencies with a bank of bandpass filters. Envelope variations in the different bands are represented as corresponding electrodes in the cochlea with modulated trains of biphasic electrical pulses. The envelope detector typically is set at 200 Hz or higher, so that the fundamental frequencies of speech sounds are represented. The output of the envelope detector generally uses a full-wave or half-wave rectifier (Rect.) followed by a low-pass filter (LPF). A Hilbert Transform or a half-wave rectifier without the low-pass filter also may be used. The outputs of the multipliers are directed to intracochlear electrodes (EL-1 to EL-n), via a transcutaneous link or a percutaneous connector. (Diagram adapted from [6] and used here with the permission of the Nature Publishing Group.)

Fig. 4. Block diagram of the CIS strategy. The strategy uses a pre-emphasis filter (Pre-emp.) to attenuate strong components in speech below 1.2 kHz. The pre-emphasis filter is followed by multiple channels of processing. Each channel includes stages of bandpass filtering (BPF), envelope detection, compression, and modulation. The envelope detectors generally use a full-wave or half-wave rectifier (Rect.) followed by a low-pass filter (LPF). A Hilbert Transform or a half-wave rectifier without the low-pass filter also may be used. Carrier waveforms for two of the multipliers are shown immediately below the two corresponding multiplier blocks (circles with a “x” mark within them). The outputs of the multipliers are directed to intracochlear electrodes (EL-1 to EL-n), via a transcutaneous link or a percutaneous connector. (Diagram adapted from [6] and used here with the permission of the Nature Publishing Group.)

The CIS strategy is similar in design and performance to the n-of-m strategy [9]. The principal difference between CIS and the n-of-m or ACE strategies is that the channel outputs are “scanned” in the latter two strategies to select the n channels with the highest envelope signals to each frame of stimulation across electrodes. Stimulus pulses are delivered only to the subset of m electrodes that correspond to the n selected channels. This spectral or channel “peak picking” scheme is designed, in part, to reduce the density of stimulation while still representing the most important aspects of the acoustic environment. The deletion of low-amplitude channels (and associated stimuli) for each frame of stimulation may reduce the overall level of masking or interference across electrode and stimulus regions in the cochlea. To the extent that the omitted channels do not contain significant information, such “unmasking” may improve the perception of the input signal by the patient. In addition, for positive signal-to-noise ratios (S/Ns), selection of the highest peaks in the spectra may emphasize the primary speech signal with respect to the noise. Detailed descriptions of these and related processing strategies, along with detailed descriptions of prior strategies, are presented in [9].
Fig. 5. Percent correct scores for 55 users of the COMBI 40 implant and the CIS processing strategy. Scores for recognition of the Hochmair–Schultz–Moser (HSM) sentences are presented in the top panel, and scores for recognition of the Freiburger monosyllabic words are presented in the bottom panel. The solid line in each panel shows the median of the scores, and the dashed and dotted lines show the interquartile ranges. The data are an updated superset of those reported in [32], kindly provided by Patrick D’Haese of Med El GmbH, in Innsbruck, Austria. The experimental conditions and implantation criteria are described in [32]. All subjects took both tests at each of the indicated intervals following initial fitting of their speech processors. Identical scores at a single test interval are displaced horizontally for clarity. Thus, for example, the horizontal “line” of scores in the top right portion of the top panel all represent scores for the 24-month test interval. (Figure is from [9] and is used here with the permission of Whurr Publishing Ltd.)

IV. PERFORMANCE WITH PRESENT-DAY SYSTEMS

A. Average Performance and Range of Scores

Each of these strategies—CIS, ACE, and \( \eta \)-of-\( m \)—supports recognition of monosyllabic words on the order of 50% correct (using hearing alone), across populations of tested subjects (see [9, Table 2.4]). Variability in outcomes is high, however, with some patients achieving scores at or near 100% correct and with other patients scoring close to zero on this most difficult of standard audiological measures. Standard deviations of the scores range from about 10% to about 30% for the various studies conducted to date.

Results from a large and carefully controlled study are presented in Fig. 5. This figure shows scores for 55 users of the Med El COMBI 40 implant system (Med El GmbH, Innsbruck, Austria) and the CIS processing strategy. Scores for the Hochmair-Schultz-Moser (HSM) sentences are presented in the top panel, and scores for recognition of the Freiburger monosyllabic words are presented in the bottom panel. Results for five measurement intervals are shown, ranging from one month to two years following the initial fitting of the speech processor. The solid line in each panel shows the median of the individual scores and the dashed and dotted lines show the interquartile ranges. The data are a superset of those reported in [32], that include scores for additional subjects at various test intervals.

Most of the subjects used an 8-channel processor with a pulse rate of about 1500/s/ electrode. Some of the subjects used fewer channels and a proportionately higher rate. (All processors used the maximum overall rate of 12 120 pulses/s across electrodes.)

As is evident from the figure, scores are broadly distributed at each test interval and for both tests. However, ceiling effects are encountered for the sentence test for many of the subjects, especially at the later test intervals. At 24 months postfitting, 47 of the 55 subjects score at 75% correct or higher, consistent with the 1995 NIH Consensus Statement. Scores for recognition of monosyllabic words are much more broadly distributed, with only a few subjects scoring 90% correct or higher.

An interesting aspect of the results presented in Fig. 5 is an apparent improvement in performance over time. This is easiest to see in the lower ranges of scores, e.g., in the steady increase in the lower interquartile lines (the dotted lines) across test intervals.

Improvements over time are even more evident in plots of mean scores for sentences and for words, as shown in Fig. 6 for these same data and for additional test intervals for the sentence test. The mean scores increase for both the sentence and word tests out to twelve months and then plateau thereafter. The mean scores for the sentence test asymptote at about 90% correct, and the mean scores for the word test asymptote at about 55% correct. Such results typify performance with the best of the modern cochlear implant systems and processing strategies, for electrical stimulation on one side with a unilateral implant.

These results are especially remarkable for the top scorers, given that only a maximum of eight broadly overlapping sectors...
of the auditory nerve are stimulated with this device and the implementation of CIS used with it. This number is quite small in comparison to the normal complement of approximately 30,000 neurons in the human auditory nerve.

The results also show a learning or accommodation effect, with continuous improvements in scores over the first 12 months of use. This suggests the likely importance of brain function in determining outcomes, and the reorganization or “knitting” (brain plasticity) that must occur to utilize such sparse inputs to the maximum extent possible.

B. Top Performers

The top performers with present-day cochlear implants can achieve remarkably high scores in tests of speech recognition. Scores for one such subject, implant subject HR4, are shown in the black bars in Fig. 7 for a comprehensive and difficult set of tests. Mean scores for six undergraduate students with normal hearing and taking the same tests are shown in the gray bars, along with the standard error of the mean for each test. HR4 was totally deaf prior to receiving his implant. The tests included recognition of monosyllabic, consonant-nucleus-consonant (CNC) words; recognition of City University of New York (CUNY) sentences; recognition of Hearing in Noise Test (HINT) sentences; recognition of Arizona Biomedical Institute (AzBio) sentences; identification of consonants (Cons) in an /l/-consonant-/l/ context; identification of vowels (Vowels) in a /b/-vowel-/t/ context; and recognition of CUNY and AzBio (Az) sentences presented in competition with a four-talker babble, at the indicated speech-to-babble ratios (+5 or +10 dB). (Figure is from [33] and is used here with the permission of the IEEE.)

Fig. 7. Percent-correct scores for implant subject HR4 and for six subjects with normal hearing. Tests included recognition of monosyllabic, consonant-nucleus-consonant (CNC) words; recognition of City University of New York (CUNY) sentences; recognition of Hearing in Noise Test (HINT) sentences; recognition of Arizona Biomedical Institute (AzBio) sentences; identification of consonants (Cons) in an /l/-consonant-/l/ context; identification of vowels (Vowels) in a /b/-vowel-/t/ context; and recognition of CUNY and AzBio (Az) sentences presented in competition with a four-talker babble, at the indicated speech-to-babble ratios (+5 or +10 dB). (Figure is from [33] and is used here with the permission of the IEEE.)

Fig. 7 shows a spectacular restoration of function for a user of a sensory neural prosthesis. All of the scores for HR4 are high. His scores for speech material presented in quiet, including words, sentences, consonants, and vowels, match or closely approximate the scores for the control group. His score for the most difficult test used in standard clinical practice, recognition of the monosyllabic CNC words is 100% correct. In contrast, some of his scores for sentences presented in competition with speech babble are worse than normal. Although his score for the CUNY sentences at the S/B of +10 dB is 98% correct, his scores for the AzBio sentences at the S/Bs of +10 dB and +5 dB are below those of the normal-hearing subjects. In all, HR4 scored at or near the ceiling of 100% correct for seven of the nine tests, and he attained scores of 77% correct or better for the remaining two tests. (The subjects with normal hearing scored at or near the ceiling for all nine tests.) HR4 scored at the ceiling for all tests given in standard clinical practice to identify deficits in hearing. His results indicate a full restoration of clinically-normal function, at least for speech reception. He used a 16-channel CIS processor, as implemented in the Clarion® CII cochlear prosthesis (Advanced Bionics Corp., Sylmar, CA, USA) [34]. This prosthesis also includes a high-bandwidth transcutaneous link and current sources with short rise and fall times, an array of 16 intracochlear electrodes, and (in the version used) a positioning device to place the electrodes next to the inner wall of the ST.

Such high scores overall are consistent with HR4’s ability to communicate with ease in most listening situations. He has no difficulty at all in telephone communications. He can understand conversations not directed to him and can identify speakers by regional dialect. He can mimic voices and accents that he has heard only after receiving the implant. His speech reception abilities are truly remarkable, abilities that could not have been imagined 20 years ago, even by the most-optimistic proponents of cochlear implants.

Other patients, using this and other implant systems, and also other processing strategies (including the t-of-m and ACE strategies), have achieved similarly high scores. For example, one of the subjects in Fig. 5 achieved a score of 98% correct in the Freiburger monosyllabic word test at the two-year interval. This subject used a COMBI 40 implant system, with its eight channels of CIS processing and eight sites of stimulation. This system also has a high-bandwidth transcutaneous link and current sources with short rise and fall times. It does not include a positioning device; nor do other versions of the Clarion prosthesis or other implant systems, that also support stellar scores for some patients.

Although more than a few patients have achieved scores like those shown in Fig. 7, most patients have lower scores, typically much lower scores for the difficult tests, as also indicated in the lower panel of Fig. 5. However, the results obtained with HR4 and his peers are an existence proof of what is possible with electrical stimulation of the auditory nerve in a totally deafened ear.
V. STRENGTHS AND LIMITATIONS OF PRESENT SYSTEMS

A. Efficacy of Sparse Representations

Some patients achieve spectacularly high scores with present-day cochlear implants. Indeed, their scores are in the normal ranges even for the most difficult of standard audiological tests. Such results are both encouraging and surprising in that the implants provide only a very crude mimicking of only some aspects of the normal physiology. In cases like that of patient HR4, 16 overlapping sectors of the auditory nerve are stimulated with 16 intracochlear electrodes. In other cases, other patients have achieved similarly high scores with 6–8 sites of stimulation in the cochlea, as noted above. The spatial specificity of stimulation with implants is much lower than that demonstrated in neural tuning curves for normal hearing [35], especially for monopolar stimulation, which is used in all present-day systems. Such broad and highly overlapping activation of the nerve most likely limits the number of perceptually separable channels to 4–8, even if more than eight electrodes are used, as also noted before. The information presented through the implant is limited to envelope variations in the 16 or fewer frequency bands for these patients. (Similar numbers apply for patients also achieving high scores but using processing strategies other than CIS.) For HR4 and others, the upper frequency of envelope variations has been set at 200–700 Hz [9], e.g., by using a cutoff frequency in the range of 200–700 Hz for the low-pass filters in the envelope detectors shown in Fig. 4. A substantial fraction of this information may be perceived by the better patients [36]–[38], and whatever is perceived is sufficient for high levels of speech recognition.

The performance achieved by HR4 and the others like him brings into question the significance for speech reception of the intricate processing, and the interplay between and among processing steps, that occur in the normal cochlea. The details of the traveling wave of mechanical displacements along the BM in response to acoustic stimuli [39], and the spatial sharpening of the membrane response by active processes at the outer hair cells (OHCs) [39], [40], are not necessary for effective representations of speech information. Also, the noninstantaneous compression function at the synapses between the IHCs and single fibers of the auditory nerve [41] is not necessary. Additional aspects of normal hearing that are not replicated with implants include multiple stages of compression (at the BM/OHC complex, at the IHCs, and at the IHC/neuron synapses); effects of efferent action on the OHCs and other structures in the cochlea [42]; the broad distributions of thresholds for the multiple afferent fibers innervating each IHC [43]; and effects of spontaneous activity in the nerve [44], which is absent or largely absent in the deafened ear [45]–[47]. Despite these many missing steps or severed connections, cochlear implants can restore clinically normal function in terms of speech reception for some patients. This is remarkable.

B. Variability in Outcomes

One of the major remaining problems with cochlear implants is the broad distribution of outcomes, especially for difficult tests and as exemplified in the bottom panel of Fig. 5. That is, patients using exactly the same implant system—with the same speech processor, transcutaneous link, implanted receiver/stimulator, and implanted electrode array—can have scores ranging from the floor to the ceiling for such tests. Indeed, only a small fraction of patients achieve the spectacularly high scores discussed above.

C. Likely Importance of Cortical Function

Accumulating and compelling evidence is pointing to differences in cortical or auditory pathway function as a likely contributor to the variability in outcomes with cochlear implants. On average, patients with short durations of deafness prior to their implants fare better than patients with long durations of deafness [48]. This may be the result of sensory deprivation for long periods, which adversely affects connections between and among neurons in the central auditory system [49] and may allow encroachment by other sensory inputs of cortical areas normally devoted to auditory processing (this encroachment is called “cross-modal plasticity,” see [50] and [51]). Although one might think that differences in nerve survival at the periphery could explain the variability, either a negative correlation or no relationship has been found between the number of surviving ganglion cells and prior word recognition scores, for deceased implant patients who in life had agreed to donate their temporal bones (containing the cochlea) for postmortem histological studies [52]–[55]. In some cases, survival of the ganglion cells was far shy of the normal complement, and yet these same patients achieved high scores in monosyllabic word tests. Conversely, in some other cases, survival of the ganglion cells was excellent, and yet these patients did not achieve high scores on the tests. Although some number of ganglion cells must be required for the function of a cochlear implant, this number appears to be small. Above that putative threshold, the brains of the better-performing patients apparently can utilize a sparse input from even a small number of surviving cells for high levels of speech reception.

Similarly, it seems likely that representation of speech sounds with the cochlear implant needs to be above some threshold in order for the brain to utilize the input for good speech reception. Single-channel implant systems did not rise above this second putative threshold; nor did prior processing strategies for multichannel implants. The combination of multiple sites of stimulation in the cochlea (at least 6–8), relatively new processing strategies such as the CIS, n-of-m, and ACE strategies, and some minimum survival of ganglion cells is sufficient for a high restoration of function in some patients. Those patients are likely to have intact “auditory brains” that can utilize these still sparse and distorted inputs, compared with the inputs the brain receives from the normal cochlea.

Other patients may not have the benefit of normal or nearly normal processing central to the auditory nerve. The effects of auditory deprivation for long periods have been mentioned. In addition, the brains of children become less “plastic” or adaptable to new inputs beyond their third or fourth birthdays. This may explain why deaf children implanted before then generally have much better outcomes than deaf children implanted at age five and older [50], [56], [57].

The brain may be the “tail that wags the dog” in determining outcomes with present-day cochlear implants. The brain “saves
“us” in achieving high scores with those implants, in somehow utilizing a crude and sparse and distorted representation at the periphery. In addition, strong learning or accommodation effects—over long periods ranging from about three months to a year or more—indicate a principal role of the brain in reaching asymptotic performance with implants (see Fig. 6). Multiple lines of evidence further indicate or suggest that impairments in brain function—including damage to the auditory pathways in the brainstem, or compromised function in the areas of cortex normally devoted to auditory processing, or reduced cortical plasticity, or cross-modal plasticity—can produce highly deleterious effects on results obtained with cochlear implants.

**D. Likely Importance of Electrode Designs**

Present designs and placements of electrodes for cochlear implants do not support more than 4–8 effective sites of stimulation, or effective or functional channels, as described in Section III-C above. Contemporary cochlear implants use between 12 and 22 intracochlear electrodes, so the number of electrodes exceeds the number of effective channels (or sites of stimulation) for practically all patients and for all current devices. The number of effective channels depends on the patient and the speech reception measure to evaluate performance. For example, increases in scores with increases in the number of active electrodes generally plateau at a lower number for consonant identification than for vowel identification. (This makes sense from the perspective that consonants may be identified with combinations of temporal and spectral cues, whereas vowels are identified primarily or exclusively with spectral cues, that are conveyed through independent sites of stimulation.) Patients with low speech reception scores generally do not have more than four effective channels for any test, whereas patients with high scores may have as many as eight or slightly more channels depending on the test (e.g., [18] and [58]).

Results from studies using acoustic simulations of implant processors and subjects with normal hearing indicate that a higher number of effective channels or sites of stimulation for implants could be beneficial. Dorman et al. found, for example, that with the simulations and normal-hearing subjects, as many as ten channels are needed to reach asymptotic performance (for difficult tests) using a CIS-like processor [59]. Other investigators have found that even more channels are needed for asymptotic performance, especially for difficult tests such as identification of vowels or recognition of speech presented in competition with noise or a multi-talker babble [18], [60]. For example, Friesen et al. found that identification of vowels for listeners with normal hearing continued to improve with the addition of channels in the acoustic simulations up to the tested limit of 20 channels, for vowels presented in quiet and at progressively worse speech-to-noise ratios out to and including +5 dB [18].

Large improvements in the performance of cochlear implants might well be obtained with an increase in the number of effective sites of stimulation, which would help narrow the gap between implant patients and subjects with normal hearing. This gap is especially wide for the many patients who do not have more than four functional channels across wide ranges of speech reception measures. Just a few more channels for the top performers with implants would almost without doubt help them in listening to speech in demanding situations, such as speech presented in competition with noise or other talkers. An increase in the number of functional channels for patients presently at the low end of the performance spectrum could improve their outcomes substantially.

A highly plausible explanation for the limitation in effective channels with implants is that the electric fields from different intracochlear electrodes strongly overlap at the sites of neural excitation (e.g., [58] and [61]). Such overlaps (or electrode interactions) may well impose an upper bound on the number of electrodes that are sufficiently independent to convey perceptually separate channels of information. In addition, a central processing deficit may contribute to the limitation, perhaps especially for patients with low speech reception scores and (usually) a relatively low number of effective channels.

A problem with ST implants is that the electrodes are relatively far from the target tissue (the spiral ganglion), even for placements of electrodes next to the inner wall of the ST. Close apposition of the target and the electrode is necessary for a high spatial specificity of stimulation [62]. One possibility for providing a close apposition is to promote the growth of neurites from the ganglion cells toward the electrodes in the ST with controlled delivery of neurotrophic drugs into the perilymph [63]–[66]. Such growth of neurites would bring the target to the electrodes. Another possibility is to implant an array of electrodes directly within the auditory nerve (an intramodiolar implant), through an opening made in the basal part of the cochlea [24]–[26], [28]–[30]. In this case, the electrodes would be placed immediately adjacent to axons of the auditory nerve. Studies are underway to evaluate each of these possibilities, including safety and efficacy studies. Results from studies to evaluate the intramodiolar implant have demonstrated that it is feasible and that the number of independent sites of stimulation with that implant may be substantially higher than the number for ST implants [29], [30].

**E. Recent Advances**

Two recent advances in the design and performance of cochlear implants are: 1) electrical stimulation of both ears with bilateral cochlear implants and 2) combined electric and acoustic stimulation (EAS) of the auditory system for persons with residual hearing at low frequencies. Bilateral electrical stimulation may reinstate at least to some extent the interaural amplitude and timing difference cues that allow people with normal hearing to lateralize sounds in the horizontal plane and to selectively “hear out” a voice or other source of sound from among multiple sources at different locations. Additionally, stimulation on both sides may allow users to make use of the acoustic shadow cast by the head for sound sources off the midline. In such cases, the S/N may well be more favorable at one ear compared with the other for multiple sources of sound, and users may be able to attend to the ear with the better S/N. Combined EAS may preserve a relatively normal hearing ability at low frequencies, with excellent frequency resolution and other attributes of normal hearing, while providing a complementary representation of high-frequency sounds with the
cochlear implant and electrical stimulation. Various surgical techniques and drug therapies have been developed to preserve low-frequency hearing in an implanted cochlea, including deliberately shallow insertions of the electrode array (6, 10, or 20 mm) so as not to damage the apical part of the cochlea and remaining hair cells there; insertion of the electrode array through the round window membrane rather than through a cochleostomy to eliminate deleterious effects of drilling (loud and possibly damaging levels of noise, introduction of blood and bone dust into the perilymph, possible damage to delicate cochlear structures such as the BM); use of “soft surgery” techniques to minimize trauma; use of thin and highly flexible electrodes; use of a lubricant such as hyaluronic acid to facilitate insertion of the array; and use of corticosteroids and other drugs to help preserve cochlear structures in the face of surgical manipulations and the introduction of a foreign body into the inner ear. Moderate-to-excellent preservation of residual hearing has been reported using the shallow insertions and some or all of the additional procedures and techniques just mentioned [67]–[80]. Among the tested methods, insertion through the round window for placement of 20 mm arrays or use of shorter arrays have produced especially good results [77], [80]–[82]. The “soft surgery” methods also have been identified as important (e.g., [76] and [83]). Studies aimed at the further development of surgical techniques, adjunctive drug therapies, and special electrode arrays are in progress; both short- and long-term preservation of residual hearing in an implanted cochlea remain as major challenges and concerns.

Each of these approaches—bilateral electrical stimulation and combined EAS—has produced large improvements in speech reception performance compared with control conditions. In particular, bilateral stimulation can provide a substantial benefit in recognizing difficult speech materials such as monosyllabic words and in recognizing speech presented in competition with spatially distinct noise, in comparison to scores obtained with either unilateral implant alone [38], [84]–[100]. In addition, use of both implants supports an improved ability to lateralize sounds, again compared with either unilateral implant [86], [89], [90], [92], [93], [97], [98], [100]–[105]. (This ability is nonexistent or almost nil with a unilateral implant.) Combined EAS also provides a substantial benefit for listening to speech in noise or in competition with a multi-talker babble, compared with either electric stimulation only or acoustic stimulation only [38], [67], [68], [70], [71], [73], [75]–[79], [106]–[110]. Indeed, in some cases the score for combined EAS is greater than the sum of the scores for the electric-only and acoustic-only conditions. This has been described as a synergistic effect [38], [71], [75], [106], [111]. In addition, identification of melodies and perception of musical sounds is greatly improved with combined EAS compared with electric stimulation alone [73], [78], [109], [112], [113]. (Scores with acoustic stimulation alone closely approximate the scores with combined EAS, for melody and music reception.)

These gains from bilateral electrical stimulation most likely arise from a partial or full restoration of the binaural difference cues and to the head shadow effect, as suggested above. In addition, gains may result from a “binaural summation” effect that is produced in normal hearing by redundant stimulation on the two sides. Detailed descriptions of these various contributors to an overall binaural benefit for normal hearing and possible contributors for prosthetic hearing are presented in [38]. The evidence to date indicates that almost all recipients of bilateral cochlear implants benefit from the head shadow effect and that some benefit from: 1) the binaural squelch effect that is made possible with presentation and perception of the binaural timing-difference cue; 2) the binaural summation effect; or 3) both. The largest contributor to improvements in listening to speech presented in competition with spatially distinct noise is the head shadow effect, which is a physical effect that is present and can be utilized whether or not the binaural processing mechanism in the brainstem is intact.

In addition to these binaural effects that occur in normal hearing and to a variable extent in prosthetic hearing, electric stimulation on both sides may help fill “gaps” in the representation of frequencies on one side—due to uneven survival of spiral ganglion cells along the cochlear spiral—with complementary excitation of surviving neurons at the same frequency place on the contralateral side. For example, a lack of input to the central nervous system (CNS) at the 5 kHz position on one side may be at least partly bridged or compensated by stimulation of remaining neurons at the 5 kHz position in the other ear. This mechanism and the binaural summation effect may underlie the large improvements observed with bilateral implants for the recognition of difficult speech material presented from in front of the subjects and without any interfering noise, where the interaural difference cues and the head shadow effect do not come into play. The mechanism also may contribute to the good results observed for other conditions, in which the difference cues and the head shadow effect are also present.

A further possible mechanism contributing to the observed benefits of bilateral electric stimulation is a higher number of effective channels. Bilateral implants, in general, provide a doubling or near doubling of physical stimulus sites, compared with either unilateral implant alone. This may provide some gain in the number of effective channels, especially in cases of uneven nerve survival across the two sides, where stimulation of an area on one side that is “dead” on the other side may add an effective channel. As noted before, even a small gain in the number of effective channels could produce a large benefit, particularly for patients who otherwise would have low levels of performance and particularly for reception of difficult speech materials or for listening to speech in adverse S/N conditions.

An example of findings from studies with recipients of bilateral implants is presented in Fig. 8. These results are from studies conducted by Müller and coworkers at the Julius–Maximilians Universität in Würzburg, Germany [87]. Nine subjects participated. The left and middle columns show individual and average scores for the recognition of sentences presented in competition with speech-spectrum noise at the S/N of +10 dB and with the sentences presented through a loudspeaker in front of the subject and the noise presented through a loudspeaker to the right of the subject (left column) or to the left of the subject (middle column). The right column shows results for the recognition of monosyllabic words in quiet, presented from the loudspeaker in front of the subject. For the sentence tests, the difference in scores for the left
implant only versus the right implant only shows the magnitude of the head shadow benefit, which is large (see lower-left and lower-middle panels). For these same tests, the difference in scores for the bilateral condition versus the score for the single implant at the side opposite to the noise source shows the magnitude of a “binaural processing benefit,” which is a combination of binaural squelch, binaural summation, and possibly other effects. This binaural processing benefit is smaller than the head shadow benefit but still significant. For the word test (right column), the difference in scores between the bilateral condition and either of the unilateral conditions may be attributable to a binaural summation effect, or a filling of gaps in nerve survival across the two sides, or a principal contribution from the better of the two ears, or a higher number of effective channels, or some combination of these, for the bilateral condition. The improvement obtained with stimulation on both sides is large, comparable to the head shadow benefits demonstrated by the results from the sentence tests. This improvement is larger than what would be expected from binaural summation effects alone.

The gains from combined EAS may arise from a normal or nearly normal input to the CNS for low-frequency sounds from the acoustic stimulation, in conjunction with a crude representation of high-frequency sounds from the electric stimulation with a partially inserted cochlear implant. The CNS apparently is able to integrate these seemingly disparate inputs into a single auditory percept, that is judged as sounding natural and intelligible. The likely ability to separate different “auditory streams” on the basis of different fundamental frequencies (and trajectories of fundamental frequencies) for different sounds may at least in part underlie the large advantages produced with combined EAS compared with electric stimulation only [77], [108], [109], [114], [115]. In particular, these fundamental frequencies (and one or more of their first several harmonics) occur at low frequencies and are within the range of residual hearing for most if not all users of combined EAS, i.e., below 500–1000 Hz. Perception and utilization of fine frequency differences in this range may allow an effective separation of a signal from interfering sounds. Also, the likely ability to “track” low frequencies almost certainly underlies the large improvements in melody recognition and music reception that have been reported (e.g., [113]).

Each of these relatively new approaches utilizes or reinstates a part of the natural system. Two ears are better than one, and
use of even a part of normal or nearly normal hearing at low frequencies can provide a highly significant advantage.

F. Possibilities for Further Improvements

Tremendous progress has been made in the design and performance of cochlear prostheses. However, much room remains for improvements. Patients with the best results still do not hear as well as listeners with normal hearing, particularly in demanding situations such as speech presented in competition with noise or other talkers. Users of standard unilateral implants do not have much access to music and other sounds that are more complex than speech. Most importantly, speech reception scores still vary widely across patients for relatively difficult tests, such as recognition of monosyllabic words, with any of the implant systems now in widespread use.

Fortunately, major steps forward have been made recently and many other possibilities for further improvements in implant design and function are on the horizon. Electrical stimulation on both sides with bilateral cochlear implants, and combined EAS for persons with some residual hearing, have been mentioned. These are new approaches, which may well be refined or optimized for still higher levels of performance. Some of the possibilities for such improvements are just now being explored, including development and evaluation of surgical techniques and adjunctive therapies aimed at the preservation of residual hearing in an implanted cochlea. In addition, other approaches—such as reinstatement of spontaneous-like activity in the auditory nerve [116], representation of “fine structure” or “fine frequency” information with novel patterns of electric stimuli [117]–[119], or a closer mimicking of the processing that occurs in the normal cochlea [118], [120]—may also produce improvements in performance, especially for patients with good or relatively good function in the central auditory pathways and in the cortical areas that process auditory information.

Further improvements for all patients might be produced by somehow increasing the number of effective channels supported by cochlear implants. Several possibilities for this have been mentioned, including intramodiolar implants and drug-induced growth of neurites toward the electrodes of ST implants. An additional possibility is to regard bilateral implants as a collection of many stimulus sites and to choose for activation the perceptually separable sites among them. Alternatively, one might “interlace” stimulus sites across the two sides, where the most basal region of one cochlea is stimulated on one side, the next most basal region on the other side, the next most basal region on the first side, and so forth until the full tonotopic map is spanned. In this way, all the frequencies would be represented but the distance between active electrodes in each implant would be doubled, which would in turn reduce the interactions among them, compared with stimulation of adjacent electrodes. These different ways of using bilateral implants have the potential to increase the number of effective channels [38], [121], but almost certainly at the cost of diminishing or eliminating a useful representation of the binaural difference cues. This may be a good tradeoff for some patients.

Each of the approaches described above is aimed at improving the representation at the periphery. A fundamentally new approach may be needed to help those patients presently at the low end of the performance spectrum, however. They may have compromised “auditory brains” as suggested above and by many recent findings. For them, a “top-down” or “cognitive neuroscience” approach to implant design may be more effective than the traditional “bottom-up” approach. In particular, the new (top-down) approach would ask what the compromised brain needs as an input in order to perform optimally, in contrast to the traditional approach of replicating insofar as possible the normal patterns of activity at the auditory nerve. The patterns of stimulation specified by the new approach are quite likely to be different from the patterns specified by the traditional approach.

A related possibility that may help all patients at least to some extent is directed training to encourage and facilitate desired plastic changes in brain function (or, to put it another way, to help the brain in its task to learn how to utilize the inputs from the periphery provided by a cochlear implant). Such training if well designed may shorten the time needed to reach asymptotic performance and may produce higher levels of auditory function at that point and beyond. The ideal training procedure for an infant or young child may be quite different from the ideal procedure for older children or adults due to differences in brain plasticity. For example, the “step size” for increments in the difficulty of a training task may need to be much smaller for adults than for infants and young children [122]. However, all patients may benefit from appropriately designed procedures, that respect the differences in brain plasticity according to age.

The brain is a critical part of a prosthesis system. For patients with a fully intact brain, the “bottom-up” approach to implant design probably is appropriate, i.e., an ever-closer approximation to the normal patterns of neural discharge at the periphery is likely to provide the inputs that the brain “expects” and is configured to receive and process. For patients with a compromised brain, such inputs may not be optimal. In those cases, a “top-down” approach to implant design, or a combination of “top-down” and “bottom-up” approaches, may produce the best results. For example, a “top-down” approach combined with techniques to minimize electrode interactions at the periphery may be especially effective for patients presently shackled with poor outcomes.

VI. INTERFACING SENSORS WITH THE NERVOUS SYSTEM

The full restoration of clinically normal function with a cochlear implant, as demonstrated by the findings for subject HR4 (Fig. 7) and others like him, bodes well for the development of other types of sensory neural prostheses. In particular, a sparse and distorted representation at the periphery may be sufficient for restoration of high levels of function for other sensory inputs as well, e.g., visual or vestibular inputs. As with cochlear implants, a putative threshold of the amount and quality of information in the peripheral representation may need to be exceeded before good outcomes can be achieved. However, this threshold may be quite low and a full replication of the exquisite and complex machinery at the periphery is certainly not necessary for the restoration of useful hearing and may not be necessary for the restoration of other senses either.
That said, reproduction of some aspects of the normal physiology is likely to be important. In cochlear implants, for example, a crude replication of the normal tonotopic representation of frequencies—with multichannel processing strategies and with multiple (and perceptually separable) sites of stimulation in the cochlea—was necessary to achieve high levels of performance. Perhaps a topographic representation would work well for a visual prosthesis, as has been suggested (e.g., [123]–[133]). As with cochlear implants, we expect some threshold of resolution in the stimulation will need to be exceeded for good function, and that the difficult problems of electrode interactions will need to be addressed for useful restoration of vision and other senses. However, the threshold may be surprisingly low. (A low threshold may be essential for a successful visual prosthesis, as the optic nerve has 1.2 million ganglion cells and associated axons, that receive inputs from 125 million photoreceptors in the retina. These numbers are substantially higher than the corresponding numbers for the cochlea, e.g., 1.2 million neurons in the optic nerve versus 30,000 neurons in the auditory nerve. The complexity of the retina and strategies for electrical stimulation using epiretinal or subretinal arrays, or electrical stimulation at more central sites in the visual pathway, are discussed in [123]–[127], [131], and [134]–[137].)

In addition, an intact or largely intact brain may well be a prerequisite for a topographic representation to work, at least initially and without training. Further, effects of cross-modal plasticity may preclude a good outcome with any type of sensory neural prosthesis, although a training approach has been proposed to mitigate or even possibly reverse these effects [125], [126].

An important consideration in the design of sensory neural prostheses is to regard the brain as a key part of the overall system. The brain of the user should be respected for what it does, and the design should foster a partnership between the brain and the prosthesis, perhaps with communications in both directions, i.e., from the brain to the prosthesis as well as from the prosthesis to the brain. Indeed, this was a principal theme of the Smart Prosthetics conference held at the Beckman Center, University of California, Irvine, in November 2006 and sponsored by the Keck Foundation and National Academies Futures Initiative, see http://www.keckfutures.org/ and the daughter pages. We expect this more holistic approach will be embraced in future designs.

The path between a sensor or an array of sensors and useful perception involves many steps and considerations. The path can be traversed, though, as demonstrated by cochlear implants.

VII. Summary

In summary, the experience with cochlear implants either indicates or suggests the following.

• Experts can be stunningly wrong in assessments of a new approach or technology; perseverance in the face of intense criticism was essential for the successful development of cochlear implants and this may prove to be the case for other types of neural prostheses as well.
• The above is not an argument for wayward or uninformed efforts, of course, and the NIH vetted cochlear implants with the Bilger study [4] before investing many millions of dollars for the further development of implant systems. Still, though, the courage to take informed risks on the part of the NIH and the investigators (including the investigators worldwide who were supported by agencies other than the NIH) was as important as anything else in moving this marvelous technology forward; in addition, some of the earlier efforts that appeared to many to be wayward at the time later proved to be prescient.
• Multidisciplinary efforts of multiple teams were required to make the cochlear implant a success, and NIH support of a large number of these efforts was critically important. Development of other types of neural prostheses without these elements in place seems unimaginable.
• A decidedly sparse and crude and distorted representation at the periphery supports a remarkable restoration of function for some users of present-day cochlear implants. This bodes well for the development of vestibular, visual, or other types of sensory neural prostheses.
• However, this representation must exceed some putative threshold of quality and quantity of information. Most likely, this means that aspects of the normal physiology need to be mimicked or reinstated to some minimal extent. The experience with cochlear implants indicates that 1) not all aspects of the normal physiology need to be reproduced and 2) those aspects that are reinstated do not have to be perfectly reproduced by any means. Present-day implants—with multiple channels of processing, multiple sites of stimulation in the cochlea, and the CIS, n-of-m, ACE, or other modern processing strategies—have exceeded the putative threshold for the great majority of patients, in that most patients score at 80% correct or higher in sentence tests using hearing alone and many patients can use the telephone without difficulty. Prior implant systems did not exceed the threshold.
• Not surprisingly, the interface to the tissue is important. Present electrode arrays for cochlear implants do not support more than 4–8 functional channels even though the number of stimulating electrodes is higher than that. Overlapping excitation fields from different electrodes almost certainly degrade their independence; this is a general problem with neural prostheses that map outputs to thousands of neurons in very close proximity to each other, as in the retina.
• Interlacing of stimulus pulses across electrodes—such that only one electrode is active at any one time—has proved to be highly effective for cochlear implants in achieving the present levels of electrode and channel independence. Such interlacing of stimuli may be effective for other types of neural prostheses. In addition, novel electrode designs, placements of electrodes in close proximity to the target neurons, drug treatments to encourage the growth of neural tissue toward electrodes, or interlacing of stimuli across bilateral implants (e.g., across implants for each retina), or combinations of these, may well increase the number of functional sites of stimulation for cochlear, as well as other types of sensory neural prostheses.
• Any residual function should be preserved and utilized to the maximum extent possible, in conjunction with the

...
prosthesis, as in combined electric and acoustic stimulation of the auditory system for persons with some residual (low-frequency) hearing.

• For sensory systems with bilateral inputs—audition, vision, and balance—reinstatement of inputs on both sides may confer large benefits to users of prosthetic systems, as demonstrated by the experience with bilateral cochlear implants.

• Percutaneous access to the implanted electrodes may provide important advantages in the initial development of any sensory neural prosthesis, as certainly was the case with cochlear implants.

• Good results take time. Asymptotic performance is not achieved with cochlear implants until at least three months of daily use and in many cases longer or much longer than that. This and other findings indicate a principal role of the brain in determining outcomes with implants. It also indicates that results from acute studies may be misleading in that they may grossly underestimate the potential of an approach. The brain is likely to be vitally important in determining outcomes with other types of neural prostheses as well, and effects of cross-modal plasticity may preclude good outcomes for persons who have been deprived of a sensory input for all or most of their lives, in that the “cortical target” for the input has been encroached or recruited by other sensory modalities and such effects may not be reversible. (Such effects have not been reversed thus far. However, this does not mean categorically that the task cannot be achieved. Merabet and coworkers have, for example, proposed a yet-to-be-tested training procedure to mitigate or reverse effects of cross-modal plasticity in the context of a visual prosthesis, see [126].)

• The power of the intact or largely intact brain to utilize sparse and distorted inputs is impressive; and this most likely underpins in large part the success of cochlear implants.

• A sensory prosthesis and the brain are “partners” in an overall system, and simply focusing on the periphery in the design of a prosthesis may provide good results for persons with fully intact brains and sensory pathways, but probably will limit results for persons with impaired pathways or impaired or altered cortical processing.

• The amount of information from the periphery that can be utilized may be increased through plastic changes in the brain, especially for infants and very young children but also for older patients, albeit at a likely slower pace of adaptation and perhaps to a lesser extent than with young children.

• Desired plastic changes may be facilitated and augmented through directed training; the optimal training procedure is likely to vary according to the age of the patient, to the duration of sensory deprivation prior to the restoration of (some) function with a prosthesis, and whether or not the sense was first lost prior to the “critical period” for the normal development of that sensory pathway and processing in the midbrain and cortex. Training may or may not be effective for patients who lost a sense prior to or during the critical period and had it reinstated (at least to some extent) after the critical period had expired. Training may be most effective for persons who lost the sense following the critical period, and after the sensory pathways and associated cortical processing had been established.

• The highly deleterious effects of cross-modal plasticity or missing the critical period for maturation of the central auditory pathways and cortex are “moral imperatives” to screen infants for deafness or blindness or possibly other sensory losses and to provide at least some input to the appropriate part of the CNS if feasible and as soon as practicable for cases in which severe deficits are found.

• Cochlear implants are among the great success stories of modern medicine, and this has surprised many. Another surprise, with the development of another highly effective sensory neural prosthesis, is certainly possible.

ACKNOWLEDGMENT

The authors are grateful to the three anonymous reviewers of the present paper for their exceptionally thoughtful and constructive comments. Limited material also was drawn or adapted from several recent publications, [9], [33], [57].

DEDICATION: This paper is dedicated to F. T. Hambrecht, M.D., and W. J. Heetdersk, M.D., Ph.D., whose vision, leadership, and scientific acumen made present-day neural prostheses possible.

REFERENCES


Wilson and Dorman: Interfacing Sensors with the Nervous System


