Two New Directions in Speech Processor Design for Cochlear Implants

Blake S. Wilson, Reinhold Schatzer, Enrique A. Lopez-Poveda, Xiaoan Sun, Dewey T. Lawson, and Robert D. Wolford

Two new approaches to the design of speech processors for cochlear implants are described. The first aims to represent "fine structure" or "fine frequency" information in a way that it can be perceived and used by patients, and the second aims to provide a closer mimicking than was previously possible of the signal processing that occurs in the normal cochlea.

(Ear & Hearing 2005;26;73S–81S)

Although great progress has been achieved in the design and performance of cochlear implant systems, much remains to be done. 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 other talkers or noise at typical signal-to-competitors ratios, for example, +5 dB. Users of standard unilateral implants do not have much access to music and other sounds that are more complex than speech. Even the "star" performers report a need for great concentration in attaining their high scores in speech-reception tests. Such a cognitive load must separate implant listeners from their normal-hearing peers. Perhaps most important, though, is the fact that scores still vary widely across patients for difficult tests, such as recognition of monosyllabic words, with any of the implant systems now in widespread use.

Quite recently, large steps forward have been made with use of bilateral cochlear implants (e.g., Müller et al., 2002) and with combined electric and acoustic stimulation (EAS) of the auditory system (e.g., von Ilberg et al., 1999), the latter for patients with some residual (low frequency) hearing either ipsilateral or contralateral to a cochlear implant. The gains are mostly seen for speech reception in noise. In addition, use of relatively high rates of

RTI International, Research Triangle Park, North Carolina (B.S.W.); the Division of Otolaryngology, Head and Neck Surgery, Department of Surgery, Duke University Medical Center, Durham, North Carolina (B.S.W., D.T.L.); the Center for Auditory Prosthesis Research, RTI International, Research Triangle Park, North Carolina (R.S., X.S., D.T.L., R.D.W.); and Laboratory of Neurobiology of Hearing, Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Avda. Alfonso X "El Sabio" s/n, 37007 Salamanca, Spain (E.A.L.-P.).

pulsatile stimulation, in conjunction with at least eight electrode sites and good current sources, has produced significant improvements in performance over prior approaches, again most notably for speech reception in noise (e.g., Frijns et al., 2003; Koch et al., 2004).

Additional approaches are being developed. Work is underway in our laboratories and elsewhere to (a) represent "fine structure" or "fine frequency" information in a way that it can be perceived and used by patients, and (b) provide a closer mimicking than was previously possible of the signal processing that occurs in the normal cochlea. These new approaches may support further gains in performance, either alone or in combination with other approaches.

The primary purpose of this report is to provide an overview of design considerations for these two new approaches. In addition, some preliminary results from studies in progress to evaluate one of the approaches are mentioned.

IMPORTANCE OF FINE STRUCTURE INFORMATION

The mathematician David Hilbert showed that signals can be decomposed into slowly varying envelopes modulating high-frequency carriers (Hilbert, 1912). An example of such a decomposition is presented in Figure 1. The instantaneous phase, or frequency (first derivative of the phase signal), of the carrier varies continuously. Hilbert described the carrier as the "fine structure" (FS) portion of the original signal.

Smith and coworkers (2002) have investigated the relative importance of envelope and FS information for speech reception, melody recognition, and sound localization. They created "auditory chimeras" by first processing two separate inputs with identical banks of bandpass filters and then multiplying the FS carriers derived from one bank of filters with the envelope signals derived from the other bank of filters. The modulated carriers were then summed to form the output. Thus, the chimeras presented conflicting cues—the envelope variations in a given number of bands for one sound, versus the FS variations in the same bands for another sound. Pairings of inputs included sen-
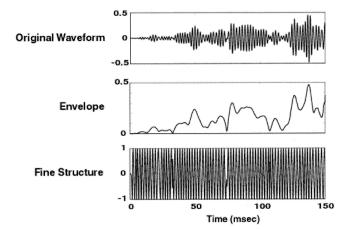
Fig. 1. Decomposition of a signal using the Hilbert transform. (Illustration from the website of Smith and colleagues, http://epl.meei.harvard.edu/~bard/chimera/; also see Smith et al., 2002.)

tences versus noise, sentences versus different sentences, melodies versus different melodies, and sentences with an interaural time delay (ITD) corresponding to a sound image at the left versus the same or different sentences with an ITD corresponding to a sound image at the right.

The sound heard or correctly identified by subjects with normal hearing depended on the type(s) of sounds in each pairing and on the number of processing channels (bands). Speech was identified by its envelope information for eight or more channels, whereas the FS information was more important for one or two channels. Both envelope and FS information contributed to sentence recognition for intermediate numbers of channels. Melodies were recognized almost exclusively by their FS information up to 32 channels. Envelope cues became dominant at 48 and 64 channels. Lateralization of sentences was difficult with a small number of channels but improved with increasing numbers up to the tested limit of 32. Lateralization was cued by the FS information in all cases.

These findings indicate the importance of the FS information for speech reception using fewer than about 8 processing channels and for music reception using fewer than about 40 channels. In addition, they indicate that ITD cues may be represented by the FS information but not the envelope information, for any number of channels up to (at least) 32.

Present-day electrode arrays for cochlear implants appear to support no more than four to eight independent channels, as indicated by a lack of increases in speech reception scores when the number of processing channels and associated sites of stimulation is increased beyond that number (Fishman et al., 1997; Friesen et al., 2001; Garnham et al., 2002; Lawson et al., Reference Note 1; Wilson,

1997). In this 4 to 8 range, both envelope and FS information contribute to speech recognition. Music is conveyed almost solely by FS cues.

The importance of the FS information seems indisputable, given the maximum number of effective channels with current implant devices. The question is, how can this information be presented in a way that it can be perceived and utilized by the patient?

PRESENT PROCESSING STRATEGIES FOR IMPLANTS

The continuous interleaved sampling (CIS), advanced combination encoder (ACE), spectral peak (SPEAK), n-of-m, and other processing strategies now in use for cochlear implants extract envelope signals from contiguous bandpass filters, which span the overall frequency range of speech and other inputs [see Wilson (2004), for detailed descriptions of these various strategies]. The envelope signals are used to determine patterns of stimulation in the implant. Thus, only the processed envelope information is presented to the user, and most or all FS information is discarded at the envelope-extraction stage.

In CIS processors, for instance, envelope signals are derived at the outputs of the bandpass filters and those derived signals are compressed into the narrow dynamic range of electrically evoked hearing (Fig. 2). The compressed signals modulate trains of pulses with a constant pulse rate. Frequency variations of signals within each bandpass are not represented in the stimuli unless (a) the cutoff frequency of the low-pass filter in the envelope detector is comparable to or higher than frequencies in the bandpass, (b) a half-wave rectifier is used instead of a full-wave rectifier in the envelope detector to avoid frequency-doubling effects, and (c) the pulse rate of the stimuli for the channel is sufficiently high to represent the relatively rapid variations in the modulation without significant distortions. Typical implementations of the CIS and the other processors use cutoff frequencies for the envelope detectors in the range of 200 to 400 Hz. This allows representations of the fundamental frequency for voiced speech sounds, voiced versus unvoiced distinctions, and rapid transient events in speech such as those associated with stop consonants. However, little or no information about frequencies within channels, even the channel with the lowest center frequency, is represented.

A variation of CIS, called "HiRes," uses high pulse rates and an effective cutoff for the envelope detectors at half the rates (Koch et al., 2004; Wilson, 2004). In addition, the detectors use a half-wave
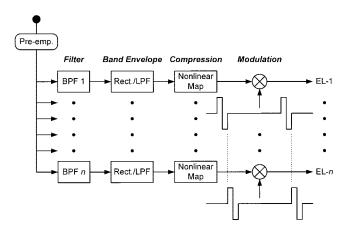
Fig. 2. Block diagram of the continuous interleaved sampling (CIS) strategy. The strategy uses a preemphasis filter (Preemp.) to attenuate strong components in speech below 1.2 kHz. The preemphasis filter is followed by multiple channels of processing. Each channel includes stages of bandpass filtering (BPF), envelope detection, compression, and modulation. The envelope detectors generally use a rectifier (Rect.) and 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 modulators are shown immediately below the two corresponding multiplier blocks (circles with a "x" mark within them). 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 Wilson et al., 1991.)

rectifier. The rates may be as high as about 2800 pulses/sec for each channel and associated electrode, for a 16-channel implementation that uses nonsimultaneous stimulation across electrodes. The rate can be almost doubled by using an implementation that presents pulses for two of the channels to their respective electrodes simultaneously. FS information may be presented for frequencies up to about 1400 Hz for the nonsimultaneous mode and up to about 2800 Hz for the "paired pulses" mode.

Another way in which FS information might be presented is to represent directly the "analog" outputs of the bandpass filters at the electrodes, as in the compressed analog (CA) or simultaneous analog stimulation (SAS) strategies (see full descriptions in Wilson, 2004). FS information is not discarded in the processing with these strategies.

Although FS information may be presented with the HiRes, CA, or SAS strategies, implant patients may not be able to perceive much if any of it. In particular, most patients do not perceive differences in the frequency of stimulation at individual electrodes as differences in pitch above a "pitch saturation limit" of about 300 Hz (e.g., Zeng, 2002). Thus, frequency variations may be presented by these strategies, but they cannot be utilized by the patients for any but the lowest frequencies. For loud-

ness-balanced stimuli, a sinusoid at 500 Hz does not sound any different to most patients than a 300 Hz sinusoid. Similarly, modulation of a pulse train at 500 Hz is not discriminable from modulation at 300 Hz.

A further concern with strategies that use simultaneous stimulation across electrodes is that such stimulation may exacerbate interactions among electrodes and thereby reduce the salience of channel-related cues (Favre and Pelizzone, 1993; Middlebrooks, 2004; White et al., 1984; Wilson, 2004). Thus, any gain produced through presentation of FS information may be counteracted, or more than counteracted, by an increase in electrode interactions.

This possible tradeoff between a representation of FS information on the one hand, and increased electrode interactions on the other hand, may vary across patients. Some patients have relatively high pitch saturation limits and therefore may have greater access than others to the presented FS information. In addition, some patients have relatively low electrode interactions, perhaps due to excellent survival of neural elements in the implanted cochlea or close placements of electrodes next to excitable tissue or both. Patients with low interactions and high pitch saturation limits may achieve especially good results with the CA, SAS, or "paired pulses" HiRes strategies. Patients with high pitch saturation limits only may receive the greatest benefit from the nonsimultaneous pulses version of HiRes and other implementations of CIS that use high cutoff frequencies for the envelope detectors.

In general, though, representation of FS information as variations in frequencies or rates of stimulation at the electrodes seems limited. Most patients will not have any access to the information above about 300 Hz, and no patient will have access to it above about 1000 Hz. In addition, the difference limens for frequencies below the limit are very much poorer for implant patients than for listeners with normal hearing, usually more than 10 times worse (e.g., Baumann & Nobbe, 2004; Zeng, 2002). Thus, even for frequencies below the pitch saturation limit, patients may be able to perceive only gross features in the presented FS information.

The typical pitch saturation limit of 300 Hz just reaches or barely exceeds the lower cutoff frequency for the bandpass filter with the lowest center frequency in the various processors mentioned above. With the possible exception of this lowest band, no information about frequency components within the bands of the filter bank will be available to most implant users with "temporal" representations of FS information.
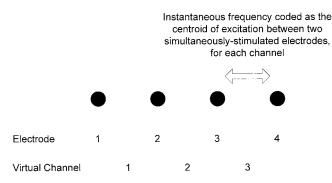
Fig. 3. Representation of fine-structure or fine-frequency information using current steering or "virtual channels." The centroid of excitation between electrodes along the length of the cochlea is shifted continuously, as instructed by an instantaneous frequency signal derived using a Hilbert transform or other means (e.g., a simple peak or zero-crossing detector) for each bandpass channel. Only four electrodes are illustrated. Present-day implants in widespread use include 12 to 22 electrodes or electrode positions. Adjacent electrodes are stimulated simultaneously, with an amplitude selected for each of the electrodes to place the centroid in the desired spot. Stimulation across such pairs of electrodes, and across channels, is nonsimultaneous, as in standard continuous interleaved sampling (CIS) processors, to avoid or minimize interactions.

SOME ALTERNATIVES FOR REPRESENTING FS INFORMATION

Possibilities for better representations of FS information include (a) the acoustic stimulation part of combined EAS and (b) fine adjustments in the sites of stimulation along the electrode array of an implant, as instructed by an instantaneous frequency signal for each channel. The acoustic stimulation part of combined EAS may be perceived in a way similar to that of low-frequency sounds in normal hearing. In such a case, FS information would be fully or largely available to the user up to the frequency limit of the residual hearing, typically 500 to 1000 Hz for EAS patients. Fine resolution of frequencies in the low-frequency range can support a high level of music reception (e.g., Gantz et al., 2004; Gfeller et al., 2002). In addition, the FS information in the low-frequency band, when combined with electric stimuli for a coarse representation of higher frequencies, may support especially high levels of speech reception in noise (Gantz et al., 2004; Kiefer et al., 2005; Turner et al., 2004; Wilson et al., Reference Note 2).

Fine adjustments in sites of stimulation might be made using virtual channels (e.g., Wilson et al., 1994), as illustrated in Figure 3, or through selection of a particular electrode among many. Coding by place of stimulation might be far more effective than coding by frequency or rate of stimulation, as de-scribed above. Coding by place may allow representation of frequencies within bands for all channels of the implant processor, not just the lowest band (at best). Such coding may be beneficial for EAS patients as well, in that FS information might be conveyed for all parts of the spectrum, using acoustic stimulation and residual hearing for the low-frequency part of the spectrum and place coding with electric stimuli for higher-frequency parts of the spectrum. This might be (even) more effective than combined EAS as presently applied, without the place coding for the higher frequencies.

A further alternative for representing FS information has been described by Stickney et al. (2002) and Zeng and coworkers (Reference Note 3). It involves frequency modulation of the carrier pulses for a CIS-like processor to reflect the instantaneous frequency for each bandpass channel. The mean rate is quite low compared with the rates of (fixed rate) carriers used in standard CIS processors. This ad-justment is made so that differences in rate pro-duced by the frequency modulation can be perceived by implant patients, that is, the maximum rate for any one channel cannot exceed the pitch saturation limit if it is to be perceived as distinct from lower rates. This approach also might present the FS information in a way that it can be perceived, although the information is transposed to much lower frequencies and differentially so across channels. Such transformations may or may not degrade the representation (and perception) of the FS information.

A general concern with the approach is that use of relatively low carrier rates in standard CIS proces-sors produces reductions in performance (e.g., Loizou et al., 2000). Thus, a tradeoff may exist between representation of FS information on the one hand versus deleterious effects of low carrier rates on the other hand.

Studies are underway in our laboratories to eval-uate the possibilities mentioned above for place coding of FS information. The subjects include users of the Ineraid device and users of an experimental version of the Nucleus device, which includes a Contour electrode array and a percutaneous connec-tor. The percutaneous access available with both of these devices allows current steering between simul-taneously stimulated electrodes as required for the construction of virtual channels. The high number of intracochlear electrodes in the Nucleus device (22 versus the six of the Ineraid device) also allows evaluation of coding through selection of one elec-trode among at least two for each processing chan-nel, at each update for the channel. Thus, the "virtual channel" approach is being evaluated in tests with both groups of subjects, and the "electrode

selection" approach is being evaluated in tests with the Nucleus subjects. Results from these various studies should be available in the near future.

Studies also are underway at the University of California at Irvine, under the able direction of Professor Zeng (Zeng, 2004a, 2004b; Zeng et al., Reference Note 3). Results from those studies should be available soon as well, and should shed light on the likely tradeoff mentioned above.

Additional approaches for representing FS information have been suggested. For example, Rubinstein et al. (1999) and Litvak et al. (2003) have suggested that improving the neural representation of relatively rapid temporal variations using "conditioner pulses" might in turn improve perception of frequency changes within channels. To the extent that perception is improved, this also might be an effective approach. To date, however, and to our knowledge, no such improvements have been demonstrated. (Dynamic range is increased with the use of conditioner pulses, but improvements in frequency discrimination, or extensions in the pitch saturation limit, have yet to be demonstrated.)

In our view, an effective representation of FS information could lead to a breakthrough in implant design and performance. Several possibilities are being pursued. Among these, we believe representations based on place coding, or place coding in conjunction with the acoustic-stimulation part of combined EAS, have the greatest promise at this time.

CLOSER MIMICKING OF PROCESSING IN THE NORMAL COCHLEA

Recent advances in electrode and stimulus design have increased the level of control that implants can exert over spatial and temporal patterns of responses in the auditory nerve. The advances include perimodiolar electrode arrays, use of high-rate carriers or high-rate conditioner pulses, and current steering to produce virtual channels or sites of stimulation between adjacent electrodes. All but the last of these advances are reviewed in Wilson et al. (2003). Virtual channels and their construction are described in Wilson et al. (1994).

The higher levels of control may be exploited to produce a closer mimicking with implants of the signal processing that occurs in the normal cochlea. The target for such an approach is illustrated in Figure 4, which shows a simplified block diagram of the normal auditory periphery. The processing includes (1) highly nonlinear filtering of the mechanical input by the basilar membrane (BM) and associated structures, including level-dependent tuning and compression, which is produced by a local feed-
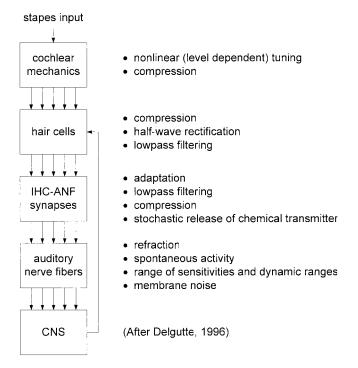

Fig. 4. Simplified block diagram of the normal auditory periphery. IHC-ANF, inner hair cell–auditory nerve fiber. (Diagram adapted from Delgutte, 1996.)

back loop involving electromotile contractions of the outer hair cells; (2) rectification, low-pass filtering, and a further compression in the transduction of BM movements to membrane potentials at the inner hair cells (IHCs); (3) a further noninstantaneous compression and adaptation at the synapses between IHCs and adjacent type I fibers of the auditory nerve; (4) random release of chemical transmitter substance at the base of the IHCs into the synaptic cleft even in the absence of stimulation, which gives rise to spontaneous activity in auditory neurons and statistical independence in discharge patterns among neurons; (5) the inability of single neurons to respond immediately after prior stimulation due to refractory effects; (6) a wide distribution of spontaneous rates among the 10 to 20 fibers that innervate each IHC; (7) a wide distribution of thresholds and dynamic ranges of those fibers, which is related to the distribution of spontaneous activities among the fibers (e.g., fibers with low rates have high thresholds and relatively wide dynamic ranges, and fibers with high rates have low thresholds and relatively narrow dynamic ranges); (8) feedback control from the central nervous system that can alter the response properties of the hair cells.

References and additional details about the processing in the normal auditory periphery are presented in Wilson et al. (2003).
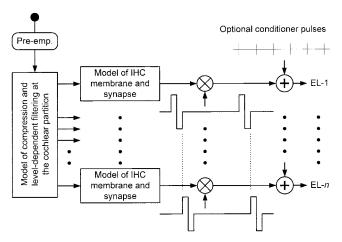
Fig. 5. Suggested processor structure for a closer mimicking of normal auditory functions. Possible models to be incorporated into a speech processor design are listed beneath the corresponding blocks. Pre-emp, Preemphasis filter; IHC, inner hair cell; EL, electrode. This diagram is reproduced from Wilson et al. (2003), with the permission of the Annual Reviews.

Present processing strategies for cochlear implants, such as the CIS strategy shown in Figure 2, provide only a very crude approximation to processing in the normal cochlea. For example, a bank of linear bandpass filters is used instead of the nonlinear and coupled filters that would model normal auditory function. Also, a single nonlinear map is used in the CIS and other strategies to produce the overall compression that the normal system achieves in multiple steps. The compression in CIS and other processors is instantaneous, whereas compression at the IHC/neuron synapse in normal hearing is noninstantaneous, with large adaptation effects.

Deng & Geisler (1987), among others, have shown that nonlinearities in BM filtering greatly enhance the neural representation of speech sounds presented in competition with noise. Similarly, findings of Tchorz & Kollmeier (1999) have indicated the importance of adaptation at the IHC/neuron synapse in representing temporal events or markers in speech, especially for speech presented in noise. Aspects of the normal processing are responsible for the sharp tuning, high sensitivity, wide dynamic range, and high resolution of normal hearing. Those aspects, and indeed entire steps and feedback loops, are missing in the processing used today for cochlear implants.

An approach for providing a much closer approximation to normal processing is suggested in Figure 5. The idea is to use better models of the normal processing, whose outputs may be fully or largely conveyed through the higher levels of neural control now available with implants.

Comparison of Figures 2 and 5 shows that in the new structure, a model of nonlinear filtering is used instead of the bank of linear filters, and a model of the IHC membrane and synapse is used instead of an envelope detector and nonlinear mapping function. Note that the mapping function is not needed in the new structure, because the multiple stages of compression implemented in the models should provide the overall compression required for mapping the wide dynamic range of processor inputs onto stimulus levels appropriate for neural activation. (Some scaling may be needed, but the compression functions should be at least approximately correct.) The compression achieved in this way would be much more analogous to the way it is achieved in normal hearing.

Conditioner pulses or high carrier rates may be used if desired, to impart spontaneous-like activity in auditory neurons and stochastic independence among neurons (Rubinstein et al., 1999; Wilson et al., 1997). This can increase the dynamic range of auditory neuron responses to electrical stimuli, bringing it closer to that observed for normal hearing using acoustic stimuli. Stochastic independence among neurons also may be helpful in representing rapid temporal variations in the stimuli at each electrode, in the collected (ensemble) responses of all neurons in the excitation field (e.g., Parnas, 1996; Wilson et al., 1997). (This does not necessarily mean that the represented variations can be perceived, as noted before in the sections on "fine structure" processors.)

The approach shown in Figure 5 is intended as a move in the direction of closer mimicking. It does not include feedback control from the CNS, and it does not include a way to stimulate fibers close to an electrode differentially, to mimic the distributions of thresholds and dynamic ranges of the multiple neurons innervating each IHC in the normal cochlea. However, it does have the potential to reinstate other important aspects of the normal processing, including details of filtering at the BM and associated structures, and including non-instantaneous compression and adaptation at the IHCs and their synapses.

IMPLEMENTATIONS OF "CLOSER MIMICKING" PROCESSORS

Studies are underway in our laboratories to evaluate various implementations of processors designed to provide a closer approximation than before to normal cochlear functions. We are proceeding in steps, including (a) substitution of a bank of dual-resonance, nonlinear (DRNL) filters (Lopez-Poveda & Meddis, 2001; Meddis et al., 2001) for the bank of

linear filters used in a standard CIS processor; (b) substitution of the Meddis IHC model (Meddis, 1986, 1988) for the envelope detector and for some of the compression ordinarily provided by the nonlinear mapping table in a standard CIS processor; and (c) combinations of (a) and (b) and fine tuning of the interstage gains and amounts of compression at various stages. Work thus far has focused on implementation and evaluation of processors using DRNL filters [step (a)]. For those processors, the envelope detectors and nonlinear mapping tables are retained, but the amount of compression provided by the tables is greatly reduced as substantial compression is provided by the DRNL filters. The DRNL filters have many parameters whose adjustment may affect performance. We have started with a set designed to provide a close approximation to filtering along the human BM (Lopez-Poveda & Meddis, 2001) but also are exploring effects produced by manipulations in the parameters, that is, to broaden tuning properties of the filters so that their responses overlap at least to some extent across channels.

In general, the frequency responses of the DRNL filters are much sharper than those of the Butterworth filters used in standard CIS processors, at least for 6 to 12 channels of processing and stimulation and at least for low-to-moderate input levels. Thus, if one simply substitutes DRNL filters for the Butterworth filters without alteration, then substantial gaps will be introduced in the represented spectra of lower-level inputs to the filter bank. Such a "picket fence" effect might degrade performance, even though other aspects of DRNL processing may be beneficial.

Studies to date have included evaluation of DRNL-based processors with broadened filters, as noted above. In addition, we have tested n-to-m constructs, in which more than one channel of DRNL processing is assigned to each stimulus site. In one variation, the average of outputs from the multiple channels is calculated and then that average is used to determine the amplitude of a stimulus pulse for a particular electrode. Each DRNL channel includes a DRNL filter, an envelope detector, and a lookup table for compressive mapping of envelope levels onto pulse amplitudes. Thus, the average is the average of mapped amplitudes for the number of DRNL channels assigned to the electrode. We call this the "avg n-to-m approach," in which m is the maximum number of electrodes available in the implant and in which n is the total number of DRNL channels, an integer multiple of m. In another variation, the maximum among outputs from the channels for each electrode is identified and then that maximum is used to determine the amplitude of the

stimulus pulse. We call this the "max n-to-m approach." Both approaches are designed to retain the sharp tuning of DRNL filters using the standard parameters while minimizing or eliminating the "picket fence" effect.

A final approach tested with one subject to date combines DRNL filters with virtual-channel stimulation. The high number of discriminable stimulus sites made available with current steering (and virtual channels) allows a high number of processing channels without having to resort to the n-to-m approaches described above.

Results from these preliminary studies are presented in Schatzer et al. (Reference Note 4) and in Wilson et al. (Reference Note 5). The studies described in Schatzer et al. included seven subjects and evaluations of various implementations of DRNL-based processors. The studies described in Wilson et al. included the one subject tested to date with processors that combined DRNL filtering with virtual-channel stimulation. Details about the processor implementations are presented in these references.

In broad terms, the results have been encouraging. Processors using n-to-m approaches have in general supported speech reception performance that is immediately on a par with that of the standard CIS processors used in everyday life by the subjects. For the one tested subject, a processor using DRNL filters in combination with virtual-channel stimulation supported significantly better performance than the standard CIS processor, especially for speech reception in noise.

SUMMARY AND CONCLUDING REMARKS

Developments of the "fine structure" and "better mimicking" strategies are in their nascent stages. The importance of the FS information seems indisputable. The question now is, how can the information be presented in a way that it can be perceived? Several promising lines of investigation are in progress. Work also is underway to develop new approaches for providing a much closer replication than was previously possible of signal processing steps in the normal cochlea. Recent advances in electrode and stimulus designs have greatly increased the control implants can exert on patterns of neural responses. This higher level of neural control might be exploited to convey the subtleties of the normal processing. Accurate models of normal processing may be utilized in future speech processor designs, in place of the very crude approximations used in present designs.

As noted in the Introduction, combinations of effective approaches may support better perfor-

mance than the best single approach. Once those most-effective single approaches have been identified, then combinations should of course also be tested.

The future is bright for cochlear implants. Some quite large steps forward have been made in the past few years. We have every prospect for continuing on that path.

REFERENCES

Baumann, U., & Nobbe, A. (2004). Pulse rate discrimination with deeply inserted electrode arrays. *Hearing Research*, *196*, 49–57.

Carney, L. H. (1993). A model for the responses of low-frequency auditory-nerve fibers in cat. *Journal of the Acoustical Society of America*, *93*, 401–417.

Delgutte, B. (1996). Physiological models for basic auditory percepts. In H. L. Hawkins, T. A. McMullen, A. N. Popper, & R. R. Fay (Eds.), *Auditory Computation* (pp. 157–220). New York: Springer.

Deng, L., & Geisler, C. D. (1987). A composite auditory model for processing speech sounds. *Journal of the Acoustical Society of America*, *82*, 2001–2012.

Favre, E., & Pelizzone, M. (1993). Channel interactions in patients using the Ineraid multichannel cochlear implant. *Hearing Research*, *66*, 150–156.

Fishman, K. E., Shannon, R. V., & Slattery, W. H. (1997). Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *Journal of Speech, Language, & Hearing Research*, *40*, 1201–1215.

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *Journal of the Acoustical Society of America*, *110*, 1150–1163.

Frijns, J. H., Klop, W. M., Bonnet, R. M., & Briaire, J. J. (2003). Optimizing the number of electrodes with high-rate stimulation of the Clarion CII cochlear implant. *Acta Oto-Laryngologica*, *123*, 138–142.

Gantz, B. J., Turner, C., & Gfeller, K. (2004). Expanding cochlear implant technology: Combined electrical and acoustical speech processing. *Cochlear Implants International, 5(Suppl. 1),* 8–14.

Garnham, C., O'Driscoll, M., Ramsden, R., & Saeed, S. (2002). Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets. *Ear & Hearing*, *23*, 540–552.

Gfeller, K. E., Turner, C., Woodworth, G., Mehr, M., Fearn, R., Knutson, J., Witt, S., & Stordahl, J. (2002). Recognition of familiar melodies by adult cochlear implant recipients and normal hearing adults. *Cochlear Implants International*, *3*, 29–53.

Hilbert, D. (1912). Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. Leipzig: Teubner.

Kiefer, J., Pok, M., Adunka, O., Stürzebecher, E., Baumgartner, W., Schmidt, M., Tillein, J., Ye, Q., & Gstoettner, W. (2005). Combined electric and acoustic stimulation (EAS) of the auditory system – Results of a clinical study. *Audiology & Neuro-Otology*, *10*, 134–144.

Koch, D. B., Osberger, M. J., Segel, P., & Kessler, D. (2004). HiResolution and conventional sound processing in the HiResolution bionic ear: Using appropriate outcome measures to assess speech recognition ability. *Audiology & Neuro-Otology*, *9*, 214–223.

Litvak, L. M., Delgutte, B., & Eddington, D. K. (2003). Improved neural representation of vowels in electric stimulation using desynchronizing pulse trains. *Journal of the Acoustical Society of America*, *114*, 2099–2111.

Loizou, P. C., Poroy, O., & Dorman, M. (2000). The effect of parametric variations of cochlear implant processors on speech understanding. *Journal of the Acoustical Society of America*, *108*, 790–802.

Lopez-Poveda, E. A., & Meddis, R. (2001). A human nonlinear cochlear filterbank. *Journal of the Acoustical Society of America*, *110*, 3107–3118.

Meddis, R. (1986). Simulation of mechanical to neural transduction in the auditory receptor. *Journal of the Acoustical Society of America*, *79*, 702–711.

Meddis, R. (1988). Simulation of auditory-neural transduction: Further studies. *Journal of the Acoustical Society of America*, *83*, 1056–1063.

Meddis, R., O'Mard, L. P., & Lopez-Poveda, E. A. (2001). A computational algorithm for computing nonlinear auditory frequency selectivity. *Journal of the Acoustical Society of America*, *109*, 2852–2861.

Middlebrooks, J. C. (2004). Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *Journal of the Acoustical Society of America*, *116*, 452–468.

Müller, J., Schön, F., & Helms, J. (2002). Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system. *Ear & Hearing*, *23*, 198–206.

Parnas, B. R. (1996). Noise and neuronal populations conspire to encode simple waveforms reliably. *IEEE Transactions on Biomedical Engineering*, *43*, 313–318.

Robert, A., & Eriksson, J. L. (1999). A composite model of the auditory periphery for simulating responses to complex tones. *Journal of the Acoustical Society of America*, *106*, 1852–1864.

Rubinstein, J. T., Wilson, B. S., Finley, C. C., & Abbas, P. J. (1999). Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation. *Hearing Research*, *127*, 108–118.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*, 87–90.

Stickney, G. S., Nie, K., & Zeng, F.-G. (2002). Realistic listening improved by adding fine structure. *Journal of the Acoustical Society of America, 112(Pt. 2)*, 2355.

Tchorz, J., & Kollmeier, B. (1999). A model of auditory perception as a front end for automatic speech recognition. *Journal of the Acoustical Society of America*, *106*, 2040–2050.

Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., & Henry, B. A. (2004). Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing. *Journal of the Acoustical Society of America*, *115*, 1729–1735.

von Ilberg, C., Kiefer, J., Tillein, J., Pfennigdorff, T., & Hartmann, R., Sturzebecher, E., & Klinke, R. (1999). Electric-acoustic stimulation of the auditory system. *ORL Journal of Otorhinolaryngology & Related Specialties*, *61*, 334–340.

White, M. W., Merzenich, M. M., & Gardi, J. N. (1984). Multichannel cochlear implants: Channel interactions and processor design. *Archives of Otolaryngology – Head & Neck Surgery*, *110*, 493–501.

Wilson, B. S. (1997). The future of cochlear implants. *British Journal of Audiology*, *31*, 205–225.

Wilson, B. S. (2004). Engineering design of cochlear implants. In F.-G. Zeng, A. N. Popper, & R. R. Fay (Eds.), *Cochlear Implants: Auditory Prostheses and Electric Hearing* (pp. 14–52). New York: Springer-Verlag.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, *352*, 236–238.

Wilson, B. S., Finley, C. C., Lawson, D. T., & Zerbi, M. (1997). Temporal representations with cochlear implants. *American Journal of Otology*, *18*, S30–S34.

Wilson, B. S., Lawson, D. T., Müller, J. M., Tyler, R. S., & Kiefer, J., et al. (2003). Cochlear implants: Some likely next steps. *Annual Review of Biomedical Engineering*, *5*, 207–249.

Wilson, B. S., Lawson, D. T., Zerbi, M., & Finley, C. C. (1994). Recent developments with the CIS strategies. In I. J. Hochmair-Desoyer, & E. S. Hochmair (Eds.), *Advances in Cochlear Implants* (pp. 103–112). Vienna: Manz.

Wilson, B. S., Sun, X., Schatzer, R., & Wolford, R. D. (2004). Representation of fine-structure or fine-frequency information with cochlear implants. *International Congress Series*, *1273*, 3–6.

Zeng, F.-G. (2002). Temporal pitch in electric hearing. *Hearing Research*, *174*, 101–106.

Zeng, F.-G. (2004a). Trends in cochlear implants. *Trends in Amplification*, *8*, 1–34.

Zeng, F.-G. (2004b). Auditory prostheses: Past, present, and future. In F.-G. Zeng, A. N. Popper, & R. R. Fay (Eds.), *Cochlear Implants: Auditory Prostheses and Electric Hearing* (pp. 1–13). New York: Springer-Verlag.

Zhang, X., Heinz, M. G., Bruce, I. C., & Carney, L. H. (2001). A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *Journal of the Acoustical Society of America*, *109*, 648–670.

REFERENCE NOTES

1 Lawson, D. T., Wilson, B. S., Zerbi, M., & Finley, C. C. (1996). Speech processors for auditory prostheses: 22 electrode percutaneous study: results for the first five subjects. Third Quarterly Progress Report, NIH project NO1-DC-5–2103, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available at http://www.rti.org/capr/caprqprs.htm.)

2 Wilson, B. S., Wolford, R. D., Lawson, D. T., & Schatzer, R. (2002). Speech processors for auditory prostheses: additional perspectives on speech reception with combined electric and acoustic stimulation. Third Quarterly Progress Report, NIH project NO1-DC-2–1002, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available at http://www.rti.org/capr/caprqprs.htm.)

3 Zeng, F.-G., Nie, K., Stickney, G., Kong, Y.-Y., Chen, H., & Vongphoe, M. (2003). Using fine structure to improve cochlear implant performance. Presented at the *2003 Conference on Implantable Auditory Prostheses*, Pacific Grove, CA, p. 39 in the book of Abstracts.

4 Schatzer, R., Wilson, B. S., Wolford, R. D., & Lawson, D. T. (2003). Speech processors for auditory prostheses: Signal processing strategy for a closer mimicking of normal auditory functions. Sixth Quarterly Progress Report, NIH project N01-DC-2–1002. Neural Prosthesis Program, National Institutes of Health, Bethesda, MD. (This report is available at http://www.rti.org/capr/caprqprs.htm).

5 Wilson, B. S., Schatzer, R., & Lopez-Poveda, E. A. (in press). Possibilities for a closer mimicking of normal auditory functions with cochlear implants. In S.B. Waltzman & J. T. Roland, Jr. (Eds.), *Cochlear Implants.* 2nd Edition. New York: Thieme Medical Publishers.

6 Wilson, B. S. (in press). Sound processing strategies. In H. Cooper & L. Craddock (Eds.), *Cochlear Implants: A Practical Guide.* 2nd Edition. London: Whurr.