Annals of Biomedical Engineering, Vol. 30, pp. 157–168, 2002
Printed in the USA. All rights reserved.

0090-6964/2002/30(2)/157/12/$15.00

Biological Basis of Hearing-Aid Design

MURRAY B. SACHS,[1] IAN C. BRUCE,[1] ROGER L. MILLER,[2] and ERIC D. YOUNG[1]

[1]Center for Hearing Sciences and Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD and [2]Hearing Research Laboratories, Division of Otolaryngology—Head and Neck Surgery, Department of Surgery, Duke University Medical Center, Durham, NC

(Received 15 October 2001; accepted 20 December 2001)

Abstract—We show that we can accurately model the auditory-nerve discharge patterns in response to sounds as complex as speech and ask how we may exploit this knowledge to test new strategies for hearing-aid signal processing. We describe the auditory-nerve representations of vowels in normal and noise-damaged ears. The normal representations are predicted well by a cochlear signal processing model originally developed by Carney (Carney, L. H. *J. Acoust. Soc. Am.* 93:401–417, 1993). Basilar-membrane tuning is represented by a time-varying narrow-band filter. Outer hair cell control of tuning is exerted by a nonlinear feedback path. We show that the effects of noise-induced outer hair cell damage can be modeled by scaling the feedback signal appropriately and use the model to test one strategy for hearing-aid speech processing. We conclude by discussing some aspects of future trends in biomedical engineering approaches to problems of hearing impairment. © *2002 Biomedical Engineering Society.*
[DOI: 10.1114/1.1458592]

Keywords—Cochlear models, Hearing impairment, Basilar membrane, Auditory nerve, Hearing aids.

INTRODUCTION

Hearing science has a history in common with biomedical engineering that dates at least to the middle of the nineteenth century. Although biomedical engineering did not emerge as a recognized discipline until the second half of the twentieth century, we would argue that the work of Helmholtz in the nineteenth century epitomized the interdisciplinary approach to problem solving typical of biomedical engineering. His monumental achievements include the 1863 publication of *On the Sensations of Tone*, a book which is to this day essential reading for anyone interested in auditory perception. Helmholtz's early biomedical contributions included measurement of the speed of conduction of nerve impulses along frog nerve and invention of the ophthalmoscope. This combination of fundamental and applied research activity is what many biomedical engineers strive for today. In 1855 he became full professor of physiology and anatomy in Bonn, where his teaching was considered unusual ''because he mixed fundamental considerations, and indeed a little mathematics, with his anatomy.'' [11] One of the questions that Helmholtz considered is how acoustic stimuli are represented in the auditory system. In this paper, we will review some aspects of the answer to that question and consider how the representation changes with hearing impairment.

The chain of events leading to auditory perception begins in the inner ear, the locus of a remarkable frequency analysis and transduction process. The sequence of steps is reviewed in Fig. 1(A). Sound impinges on the outer ear and causes the eardrum to vibrate; these vibrations are transmitted to the inner ear via the middle ear bones. The motion of the middle ear leads to fluid motion in the cochlea and ultimately to vibratory displacement of the basilar membrane (BM). Understanding the nonlinear motion of the BM through modeling has been the focus of numerous biomedical engineers since the time of Helmholtz and continues to be the source of intense research.[6,29] The important aspect of BM motion is that it is tuned by frequency, so that different frequency components of a sound cause BM vibration at different places along its length [Fig. 1(B)]. The result is a mapping of sound frequency into place along the BM, illustrated by the ''best frequency'' scale in Fig. 1(B). This Fourier-like decomposition is the basis for the tonotopic representation of sound in the auditory system and was originally postulated by Helmholtz in the nineteenth century. Vibration patterns of the BM are transformed into action potential in the auditory nerve by transducer cells called hair cells which sense the motion of the BM [Fig. 1(A); Refs. 14 and 24].

The representation of sound by the action-potential patterns of auditory-nerve fibers has been studied extensively[10] and modeled using the theory of stochastic processes (e.g., see Refs. 12, 13, and 35). The auditory nerve transmits its neural code to the central nervous system, where the first processing station is the cochlear nucleus. Experimental and modeling studies of signal

Address correspondence to Murray B. Sachs, PhD, Department of Biomedical Engineering, 720 Ross Bldg., Johns Hopkins University, 720 Rutland Avenue, Baltimore, MD 21205. Electronic mail: msachs@bme.jhu.edu

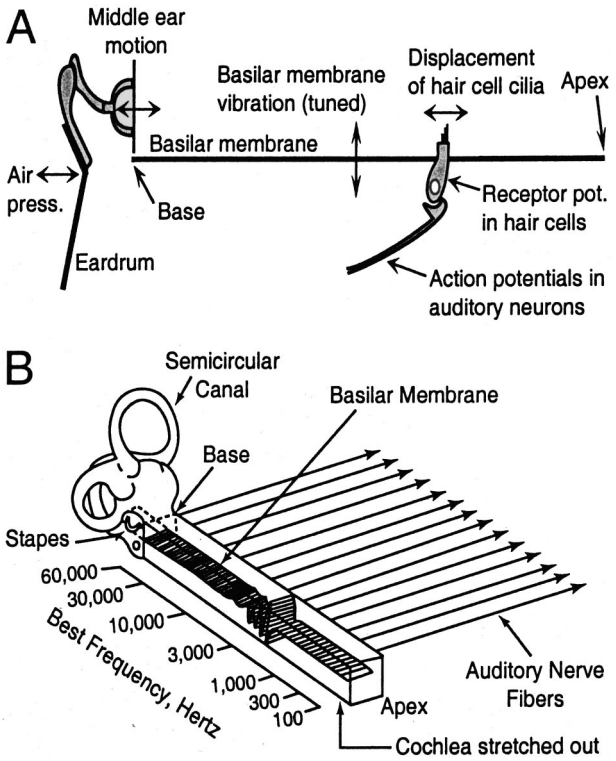

FIGURE 1. (A) Summary of the transduction process in the mammalian cochlea. Sound (air-pressure fluctuations) cause mechanical vibrations of the eardrum and middle ear bones; these couple the vibrations to the inner ear, where they produce vibration of the basilar membrane (BM). BM vibrations, in turn, displace the cilia of hair cells which transduce the vibration into electrical potentials that excite action potentials in auditory-nerve fibers. (B) Illustration of the tonotopic organization of the cochlea. The BM vibrations are tuned, so that energy at a given frequency causes a vibration which peaks at one point along the membrane. The scale at left shows the mapping of frequencies of maximum displacement (or best frequencies) into place along the BM for the cat cochlea. Auditory-nerve fibers innervate one hair cell, and so are sensitive to the BM vibration at that point [(B) redrawn from Zweig *et al.* (Ref. 44)].

processing in the cochlear nucleus and higher centers in the auditory system have been targets of biomedical engineers since the 1960s.[19,26,36] An optimistic assessment of the current state of knowledge of this peripheral auditory system suggests that specification of the auditory-nerve code for sounds as complex as speech is within our reach.[32] That is to say, given a sound-pressure wave form at the entrance to the external ear, we can accurately model the auditory-nerve discharge patterns in response to that sound.

In this paper we ask how we may exploit this knowledge to benefit the hearing impaired. Audiologists describe two types of hearing impairment: conductive hearing loss and sensorineural hearing loss. Conductive loss results from a deficiency in the transmission of acoustic energy through the external and middle ears. Such hearing loss can be corrected by relatively simple hearing aids which simply amplify sound to compensate for the frequency-dependent transmission loss. Sensorineural hearing loss results from damage to the hair cells in the inner ear. Not shown in the summary of Fig. 1(A) is that there are two kinds of hair cells in the mammalian cochlea, called inner and outer hair cells. Inner hair cells (IHCs) synapse with auditory nerve fibers and serve the transducer function described in Fig. 1(A). Outer hair cells (OHCs) participate in generating the mechanical motion of the BM; there is now good evidence that outer hair cells are responsible for the sensitivity and frequency selectivity of BM motion.[23] Hair cell damage produces a complex of effects, including an inability to hear soft sounds, disturbances of loudness, and loss of frequency selectivity.[20] Even with sophisticated hearing aids, these effects of sensorineural impairment cannot be fully compensated. The usual strategy in hearing-aid design is to attempt to restore normal perception of sound. In the work reviewed below, we have taken the alternate strategy of attempting to restore normal patterns of auditory nerve activity. The goal is to develop models of the neural representation of sound in impaired ears that are accurate enough to suggest and test new strategies for hearing-aid signal processing. We will describe one simple example of such a strategy.[18]

NORMAL AUDITORY-NERVE REPRESENTATION OF VOWELS

As a result of the mechanical tuning of the BM,[21] each auditory-nerve fiber is tuned to a best frequency (BF). This tuning is usually illustrated by a frequency tuning curve, which is a plot of the sound pressure at the threshold of response versus frequency. The solid line in Fig. 2(A) shows a schematic tuning curve for an auditory-nerve fiber. A tone with a frequency and level falling above the tuning curve (in the gray area) will cause an increase in the discharge rate of this neuron above its spontaneous rate. The tuning curve is commonly characterized by three parameters: BF, threshold and $Q_{10\,\mathrm{dB}}$. The BF is the frequency with the lowest response threshold; it corresponds to the cochlear place innervated by the fiber. The threshold value at BF is referred to as the fiber's threshold. $Q_{10\,\mathrm{dB}}$, a measure of the sharpness of tuning, is defined as the BF divided by the bandwidth of the tuning curve 10 dB above threshold.

Figure 2(B) shows thresholds plotted against BF for a population of auditory-nerve fibers from several cats.[17] At any BF there is a range of thresholds, even in normal ears. The solid line marked ''BTC'' shows the best or most sensitive thresholds and defines the normal threshold of hearing, as a function of frequency. Figure 2(C)

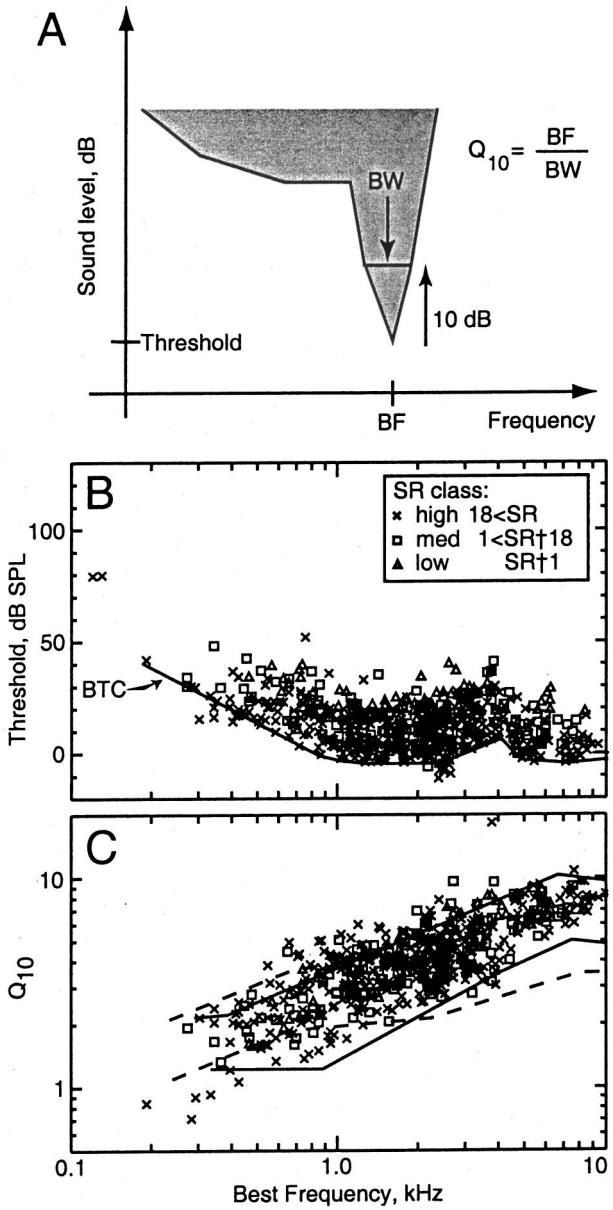

FIGURE 2. Measures of tuning curve characteristics for fibers in a pool of normal hearing cats indicate normal sensitivity and tuning. Modified from Miller *et al.* (Ref. 17) with permission from the Acoustical Society of America © 1997. Symbols denote SR class for the fiber, as shown in the legend. (A) Schematic indicating how BF, threshold, and Q_{10} are calculated from the threshold tuning curve. (B) Distribution of thresholds at BF. The solid line labeled BTC describes the best thresholds for animals used in this study. (C) Tuning curve width as measure by Q_{10} was within the range of values obtained previously by others [dashed lines from Liberman (Ref. 15); solid lines from Evans and Wilson (Ref. 9)].

shows the range of $Q_{10\ dB}$ for normal fibers; there is a trend for sharper tuning at higher BFs.

The tuning curve describes the responses of auditory-nerve fibers to pure tones. It can only predict in general terms the responses to a complex stimulus, like speech.[32]

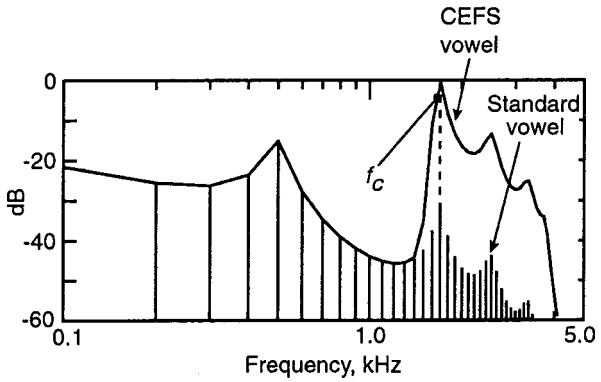

FIGURE 3. Power spectra of the standard vowel /ɛ/ and CEFS version. The line spectrum shows the unprocessed vowel's spectral shape and the solid line shows the CEFS-modified spectral envelope. Both are periodic stimuli with a fundamental frequency of 100 Hz and formants at 500 Hz, 1.7 kHz, and 2.5 kHz. The CEFS vowel was obtained by high-pass filtering the standard vowel with a cutoff frequency 50 Hz below the second-formant frequency (indicated by the vertical dashed line). Modified from Fig. 1(B) of Miller *et al.* (Ref. 18) with permission from the Acoustical Society of America © 1999.

In this paper, we consider the responses to a steady vowel-like sound, whose spectrum is shown in Fig. 3 [the "standard vowel;" the "contrast-enhanced frequency-shaped (CEFS) vowel" will be discussed later]. The sound has the spectrum of the vowel /ɛ/ as in "bet;" it is periodic with a fundamental frequency of 100 Hz, so it has a line spectrum. Decades of speech research have shown that the identity of a vowel is determined by its formant frequencies, i.e., the frequencies of the spectral peaks caused by resonances in the vocal track.[25,27] For the spectrum in Fig. 3, the formants are at 0.5 kHz (F1), 1.7 kHz (F2), and 2.5 kHz (F3). From the analysis of the responses of auditory-nerve fibers to this and similar stimuli, the fundamental idea underlying speech encoding is that fibers with BFs near a formant respond preferentially to that formant; the responses to the various formants dominate the responses to other frequencies across the population of auditory-nerve fibers.

Because the vowel stimulus has many frequency components, the analytical technique shown in Fig. 4 is necessary in order to describe fully a fiber's responses to the vowel. A 50 ms segment of the time wave form of the vowel /aɛ/ (as in "bat") is shown in Fig. 4(A). As before, the vowel is periodic with a period of 10 ms; the four cycles of oscillation in each fundamental period correspond to the first formant, as shown in the vowel's magnitude spectrum [Fig. 4(B)]. Figure 4(C) shows a train of action potentials from an auditory-nerve fiber responding to the vowel. It has been known since the 1960s (Ref. 30) that action potentials recorded from auditory-nerve fibers tend to occur in synchrony with the stimulus wave form for frequencies below about 5 kHz.

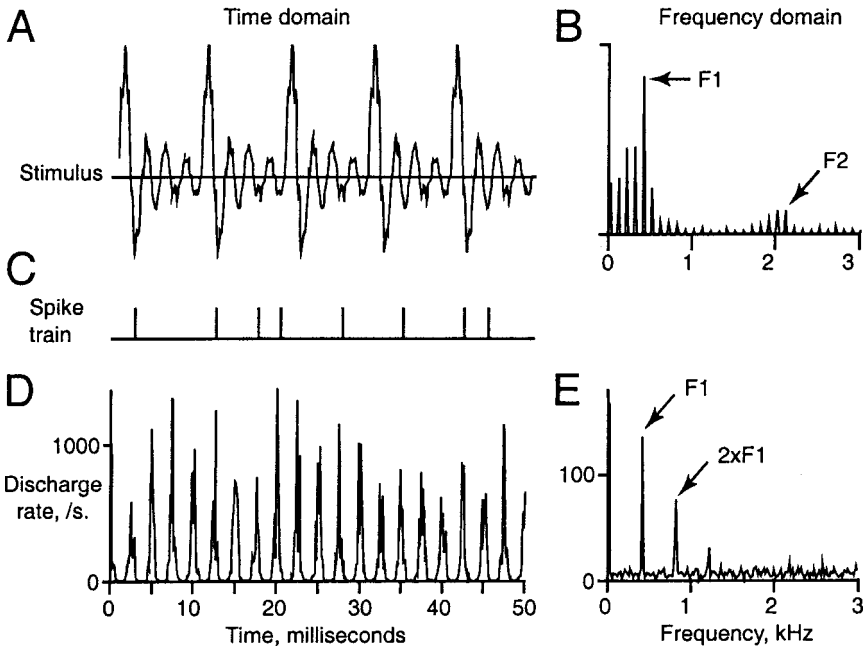

FIGURE 4. Method of analyzing auditory nerve responses to a vowel. (A) Time-domain wave form of a 50 ms segment of the vowel /aɛ/ (as in "bat"). (B) Magnitude spectrum of the vowel. Spectral peaks corresponding to the first (F1) and second (F2) formants are indicated by the arrows. (C) Example action-potential train of an auditory-nerve fiber in response to one presentation of the vowel. (D) Mean instantaneous discharge rate vs time, obtained by averaging responses like that in (C) over many repetitions of the vowel. (E) Frequency spectrum of the mean instantaneous discharge rate. The example fiber was tuned near the F1 frequency and consequently the instantaneous rate is phase locked to the F1 component of the vowel (the fourth harmonic). This dominant response is reflected in the large F1 component of the spectrum in (E). As discussed in the section on normal auditory-nerve representation of vowels, the higher harmonics in the response reflect distortion of the discharge rate response by rectification at zero rate.

This so-called phase locking can be clearly seen in the plot of the instantaneous discharge rate of the fiber versus time [Fig. 4(D)]; the stimulus frequencies to which the fiber's action-potential train are phase locked are shown by the magnitude spectrum of the instantaneous rate in Fig. 4(E) (referred to as "synchronized rate" below). This fiber's BF was near the first formant frequency and so the instantaneous rate is locked to the first formant (the fourth harmonic of the fundamental). This dominant phase locking is seen as the large peak marked "F1" in Fig. 4(E); the smaller peaks at the harmonics of F1 ("2×F1," etc.) are harmonics of the response resulting from the fact that the instantaneous rate curve in Fig. 4(D) is rectified at a zero discharge rate.[40]

Figure 5(A) shows the tuning curve of an auditory-nerve fiber with a BF exactly at the second formant frequency of the vowel (1.7 kHz). Figure 5(B) shows some important aspects of the phase-locked responses of this fiber to the vowel [analyzed as in Fig. 4(E); Ref. 17].

Responses are shown at three stimulus intensities, or loudnesses. The synchronized rates have a distinct peak at the second formant at the three different sound levels. At the lowest level (31 dB) there are also response components in a range of frequencies around the second formant, which reflect the fact that the tuning curve has a certain width even near the tip. There are also response components at higher frequencies, around the second harmonic of the second formant, due to the rectification discussed above.

As sound level increases there are two important changes in the response. First, as expected, the synchronized rate at the second formant increases, although the increase is clearly not linear. The second-formant amplitude increases at lower levels, but then saturates or even decreases slightly at the highest levels. This nonlinear behavior is due in great part to a compressive nonlinearity in the amplitude of response of the BM.[29] We will have more to say about this BM compression in the section on the computational model of normal and impaired auditory-nerve representation. As the component at the second formant increases, components in the surrounding band of frequencies disappear. This phenomenon is called "synchrony capture,"[7,8] i.e., the large spectral peak at the formant captures all of the fiber's synchrony. Synchrony capture is due to the cumulative effects of nonlinearities in the BM, the inner hair cell, the synapse, and the auditory-nerve fiber.

Figure 6 shows the responses of a population of 410 auditory-nerve fibers arrayed along the abscissa accord-

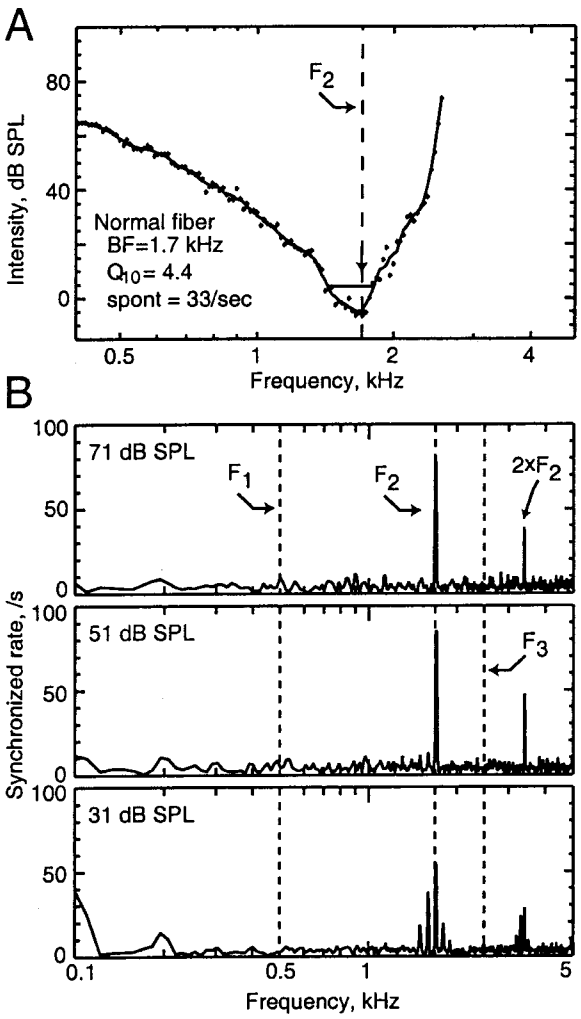

FIGURE 5. (A) Threshold tuning curve for a normal fiber with BF near F2. The horizontal line indicates the tuning curve width used to compute Q_{10}. The vertical dashed line indicates the F2 frequency. The arrow shows the fiber's BF. (B) Magnitude spectra of instantaneous rate responses (synchronized rate) of the fiber to the vowel /ϵ/ at three sound levels. The normal conversational speech level is in the range of 51–71 dB. Modified from Miller *et al.* (Ref. 17) with permission from the Acoustical Society of America © 1997.

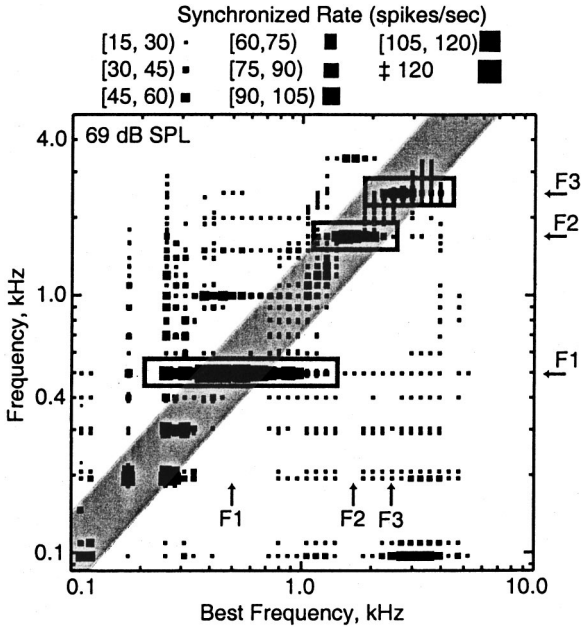

FIGURE 6. Population distribution of synchronized rates in normal fibers in response to the standard vowel. The horizontal axis is fiber BF; each column of squares shows the average synchronized rate for all fibers within a BF bin of width 0.133 octave. The size of each square represents the synchronized rate, as defined in the legend; for example, the largest squares are synchronized rates at or above 120 action potentials/s. Response components fewer than 15 action potentials/s are not plotted. Formant frequency positions are marked with horizontal arrows to the right of the figure and points where BF equals formant frequency are shown by vertical arrows at the bottom of the figure. The light-gray shaded diagonal area shows phase locking to frequencies within 0.5 octaves of BF. The three gray rectangles show where fibers are responding in a tonotopically appropriate manner to the formants. Modified from Fig. 6 of Schilling *et al.* (Ref. 33) with permission from Elsevier Science © 1998.

ing to fiber BF.[33] Each column shows the average synchronized rate of a small population of fibers grouped by BF. The amplitude of response is indicated by the size of the symbol at each frequency, as shown in the legend. This plot shows the complete tonotopic representation of the vowel, with fibers arrayed along the abscissa according to place along the BM (BF) and the strength of phase locking to different stimulus frequencies arrayed along the ordinate. The shaded region shows where fibers are responding to stimulus frequencies close to BF. Responses to formant frequencies are localized near their cochlear place, i.e., the response to the formants is strongest among fibers tuned to the formant frequency (as shown by the three gray boxes). For example, if we look at a group of fibers with BFs near the second formant [as in Fig. 5(A)], we see a large response only at the second formant and its second harmonic (middle gray box). Similarly, responses at the first formant place are dominated by the first formant and its harmonics. The first formant response spreads more widely than the second, to BFs from 0.2 to 1.2 kHz. There is also a noticeable response to the third formant located near the third-formant place. These are the defining features of the normal auditory-nerve representation for this vowel: strong narrow-band responses to the formants that dominate at the cochlear place of the formant frequency. These features depend strongly on narrow tuning and synchrony capture in the auditory nerve. This representation is the gold standard toward which we aim in the design of hearing-aid signal processing.

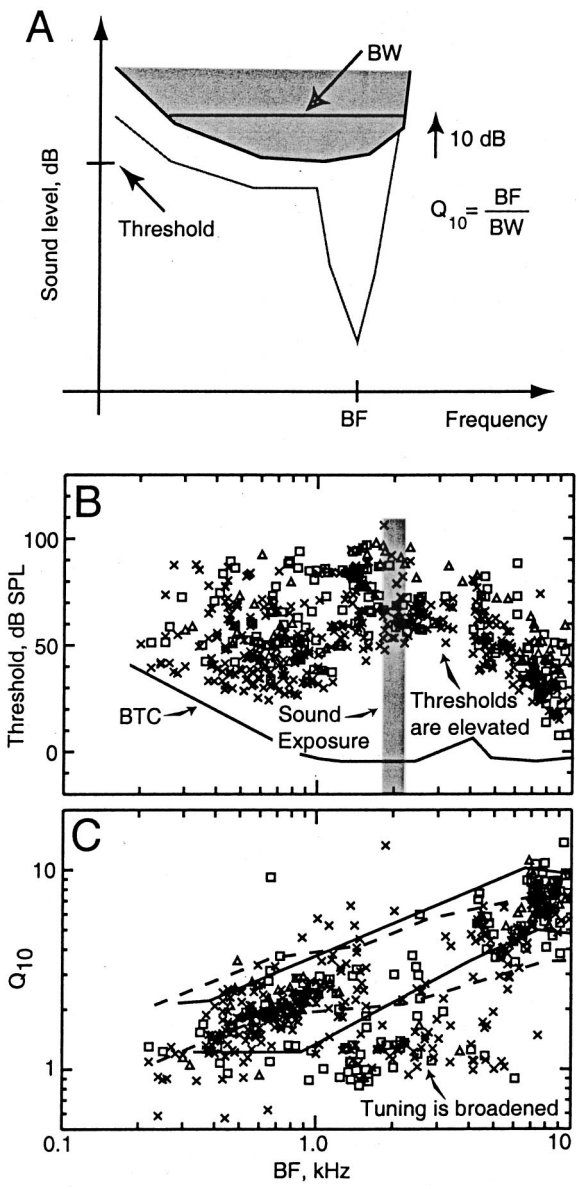

FIGURE 7. Threshold tuning curve characteristics [calculated as shown in panel (A)] for fibers in a pool of impaired cats produced by acoustic trauma. Cats were anesthetized and exposed to a 50 Hz wide band of noise centered at 2 kHz for 2 h at levels between 110–115 dB SPL (the vertical gray bar). (B) Thresholds and (C) $Q_{1\,dB}$ for fibers recorded after acoustic trauma. Solid and dashed lines are the same as in Fig. 2.

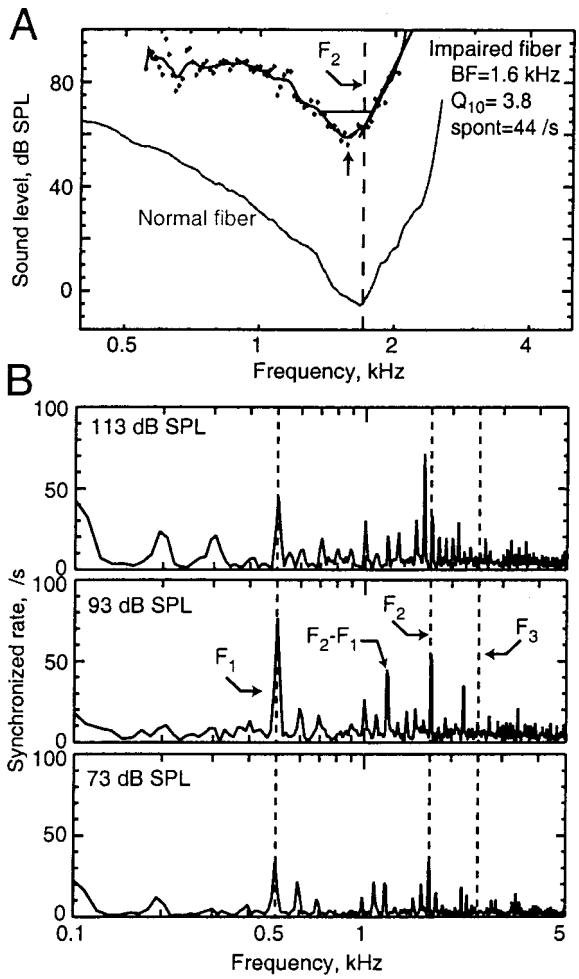

FIGURE 8. (A) Threshold tuning curve for an impaired fiber with BF near F2. The horizontal line indicates tuning curve width used to compute Q_{10}. The vertical dashed line indicates the F2 frequency. The diagonal line fit to the upper edge of the threshold tuning curve shows the high-frequency slope; for impaired fibers with broad tuning, BFs were chosen near the point of departure of the threshold tuning curve from this straight line (Ref. 15). The vertical arrow shows the fiber's BF. The impaired fiber shows a loss of sensitivity of about 60 dB. (B) Synchronized responses of the impaired fiber to the vowel /ɛ/ at three sound levels. Modified from Figs. 8 and 9 of Miller *et al.* (Ref. 17) with permission from the Acoustical Society of America © 1997.

AUDITORY-NERVE REPRESENTATION OF VOWELS IN NOISE-DAMAGED EARS

We address next what happens to the representation when the cochlea is damaged. One common cause of sensorineural hearing loss is acoustic trauma produced by exposure to loud sounds.[22] An animal model can be created by exposing the ear to high-level narrow-band noise with the results shown in Fig. 7. The schematic in panel A shows the typical effects of acoustic trauma on auditory-nerve fibers with BFs in the vicinity of the exposure frequency. These are an increase in the threshold at BF and a broadening of tuning. It has been established that these losses can be produced by damage to outer hair cells only, although usually acoustic trauma damages both outer and inner hair cells.[16] Figures 7(B) and 7(C) show the decrease in sensitivity and in the sharpness of tuning across a population of fibers; the losses are greatest for fibers with BFs close to the exposure frequency.

Figure 8 shows results from a fiber with BF close to the second formant in a noise-damaged ear.[17] The fiber's

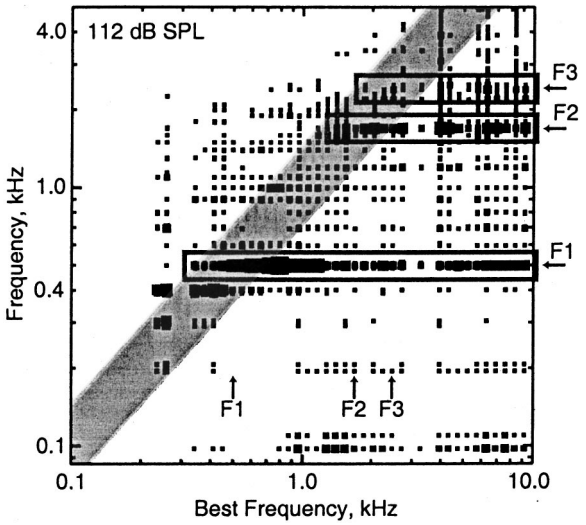

FIGURE 9. Population distribution of synchronized rates in impaired fibers in response to the standard vowel. Plotting conventions are the same as in Fig. 6. The three dark-gray rectangles emphasize that impaired fibers are *not* responding in a tonotopically appropriate manner. That is, responses to the formants spread to higher BF regions (especially F1) and the responses to F2 and F3 are weakened. Modified from Fig. 7 of Schilling *et al.* (Ref. 33) with permission from Elsevier Science © 1998.

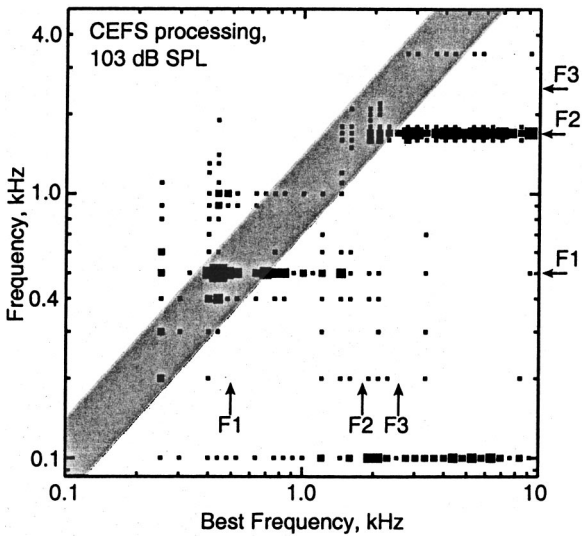

FIGURE 10. Population distribution of synchronized rates in impaired fibers in response to the CEFS vowel. Plotting conventions are the same as in Fig. 6. Note that responses are seen mainly at the first two formants and that the response to F1 is confined to its tonotopically appropriate place (within the diagonal shaded region). Responses to F2 still spread to BFs above F2 and there is no response to F3. Modified from Fig. 10 of Miller *et al.* (Ref. 18) with permission from the Acoustical Society of America © 1999.

tuning curve is compared in Fig. 8(A) with that of a normal fiber; the effect of the exposure on the threshold shift is clear. Figure 8(B) shows the responses to /ɛ/ from this fiber.[17] First, at all stimulus levels the responses are broadband. Notice, in particular, the strong responses to the first formant, and to the difference tone (F2–F1) generated by the first and second formants. Neither of these components was present in the responses of the normal fiber (cf. Fig. 5). In addition, there is phase locking to a number of other frequency components, not seen in the normal case. Clearly, there is not the same degree of synchrony capture by the second formant as we saw in the normal fiber.

The effects of the noise damage on the responses of a population of 182 fibers to the vowel are illustrated in Fig. 9.[33] In the impaired case the phase locking is much more diffuse than was seen in Fig. 6 and formant responses are no longer localized to their appropriate BF places. The impaired fibers respond to the first formant across most of the BF places. The response to the second formant is considerably weaker at its place in the impaired ear than in the normal ear and there is no place where the second formant dominates the response as in the normal ear. Virtually no separate F3 response is apparent.

We consider next what these abnormalities in the representation of speech in the impaired ear can tell us about the efficacy of signal processing schemes for hearing aids. The most straightforward and common processing scheme is simply amplification shaped to the threshold shift. This ''frequency-shaped'' amplification boosts the stimulus energy in the regions of threshold shift. The result for a vowel like /ɛ/ is that the second and third formants are amplified, but also amplified is energy at frequencies in the trough between the first and second formants. Schilling *et al.*[33] found that such amplification does restore stronger synchrony to F2 in fibers with BFs in the F2 frequency region and prevents the upward spread of synchrony to F1 into higher BF regions. However, anomalous responses to frequency components in the trough region between F1 and F2 were found in fibers with BFs in that region. These abnormal response components result from amplification of stimulus energy in this region, where the amplification is high to match the hearing loss profile. They could be misinterpreted by a hearing-impaired listener as a formant peak. Such anomalous formants are a fundamental problem for hearing aids with strongly frequency-shaped gain functions, including multiband compression aids, where the gain at the boundary between frequency bands can vary considerably.

A better strategy would be to amplify the F2/F3 region but not the trough region. Miller *et al.*[18] passed the vowel through a high-pass filter with a cutoff frequency just below the second formant frequency. The envelop of the spectrum of the contrast-enhanced frequency-shaped vowel is shown in Fig. 3 (solid line). The vertical dashed

line shows the second-formant frequency and the dot shows the high-pass filter cutoff.

Figure 10 illustrates how this signal processing improves the representation in the auditory nerve. There is no longer a strong response to the trough frequencies, the spread of F1 is restricted to frequencies near the F1 place, and there is a dominant response to F2 at the appropriate place. However, F2 spread is not reduced and there is still no F3 component in the response.

To summarize, we have described the representation of a simple speech sound in the responses of the auditory nerve in normal ears. We have further described the changes in that representation caused by noise-induced cochlear impairment and have shown that amplification of frequencies in the region of impairment improves the representation to some extent. Finally, we showed that a processing scheme that targets the formant frequencies does a better, though not perfect, job of restoring the auditory-nerve representation. Note that the CEFS algorithm presents significant problems to practical realization. With real speech, the cutoff frequency of the amplification would have to be time varying, following the second formant frequency of the speech. Algorithms to estimate the formants in real time using short segments of speech will be an interesting engineering challenge.

COMPUTATIONAL MODEL OF NORMAL AND IMPAIRED AUDITORY-NERVE REPRESENTATION

Generalizing the experimental program described above offers practical problems. Studying the representation of a variety of vowels, consonants, consonant–vowel combination, running speech, speech in noise, and on and on, is impractical in neurophysiological preparations because of the cost and time involved. To make matters worse, there are many potential processing schemes that might be pursued. In order to decrease this time and energy burden and to optimize the process of developing signal-processing schemes, we have turned to the construction of models for the normal and damaged ear that are accurate enough to allow us to test new processing schemes and new speech sounds. Such a model will allow us to eliminate schemes that prove to be of little value and thereby to reduce the cost in time, money, animals, and people needed to answer these questions experimentally.

What we require of such models is that it accurately reproduce the responses of auditory-nerve fibers to the complex, time-varying stimuli that constitute human speech. Because we currently have a large database of experimental results, we are in a position to put candidate models through rigorous testing.

Carney[3] has developed a model that can handle in some detail the time-varying speech signals of interest to us. The model incorporates many of the details of signal processing in the cochlea that we believe to be essential in predicting responses to speech, and it can be modified to account for hair-cell damage. The original Carney model for the normal cochlea is shown in Fig. 11. It consists of several sections, each modeling a different link in the chain of cochlear processes leading from a sound pressure at the ear canal to discharge patterns in the auditory nerve. Here we will focus on the first section of the model, which describes the outer hair-cell-related properties and the tuning of the BM. The second section models the nonlinear transduction and low-pass filtering of the inner hair cell. The third section describes adaptation of synaptic transmission between the inner hair cell and the auditory nerve and action-potential generation in the auditory nerve.[34]

In the model, BM tuning is represented by a time-varying narrow-band filter. Outer hair-cell control of BM motion is exerted by a nonlinear feedback path. The gain and bandwidth of the time-varying filter are controlled by the feedback signal $F(t)$. The inset in Fig. 11 shows how the gain and bandwidth of the BM filter change as a function of the feedback signal $F(t)$. $F(t)$ is the time constant of the components of the filter; $F(t)$ takes its maximum value of $3\tau_0/2$ when the feedback signal V_{fb} is zero, at low stimulus levels. This gives a sharply tuned, high-gain filter. As V_{fb} increases at higher sound levels, $F(t)$ decreases to τ_0, increasing both the bandwidth and attenuation of the filter. Because of the saturating feedback nonlinearity, the filter is sharp and linear at low levels, broader and compressive at moderate levels, and broad and linear at high levels.[3] This compressive nonlinearity has been well described experimentally in measurements of BM displacements.[29]

Damage to outer hair cells or to their stereocilia impairs their feedback on BM vibrations, which is thought to explain the broadened tips and elevated thresholds of tuning curves in impaired ears.[16,31] To model these effects we multiply the feedback signal by a scaling constant C, where C is between 0 and 1. For normal outer hair-cell function, C is set to 1 and the filter behaves as is shown in Fig. 11. For impaired OHC function, C is reduced, which reduces the maximum value of $F(t)$ to $(1+C/2)\tau_0$; $F(t)$ now varies between that maximum value and zero, which means that the BM filter has a lower gain and broader bandwidth at low-stimulus levels and less compression at moderate levels. The effect of C on model tuning curves is shown in Fig. 12 for model fibers with BFs at F2 and F3. These resemble the normal and damaged fibers in Figs. 5(A) and 8(A).

Figure 13 shows model population responses to the standard vowel for the normal model [$C=1$, Fig. 13(A)] and an impaired model [Fig. 13(B)]. In the impaired model, C for each BF is set to produce a tuning curve with $Q_{10\,\mathrm{dB}}$ equal to the twenty-fifth percentile of the

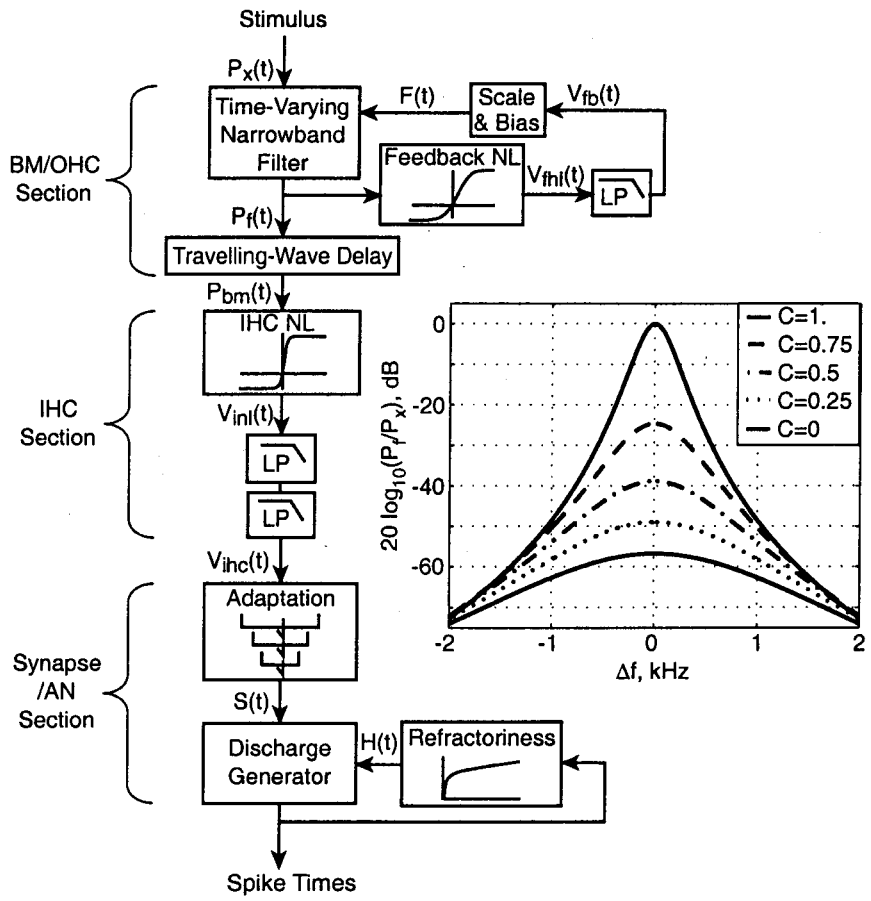

FIGURE 11. Auditory-periphery model of Carney (Ref. 3) redrawn from Fig. 1 of Carney (Ref. 3) with permission from the Acoustical Society of America © 1993. Abbreviations: basilar membrane (BM); outer hair cell (OHC); low-pass (LP) filter; nonlinearity (NL); inner hair cell (IHC); auditory nerve (AN). Inset: Gain functions for the filter, plotted as gain vs frequency deviation Δf from BF, are shown in the inset for five different values of $F(t)$ between $3\tau_0/2$ and τ_0, where τ_0 is the minimum time constant of the nonlinear filter.

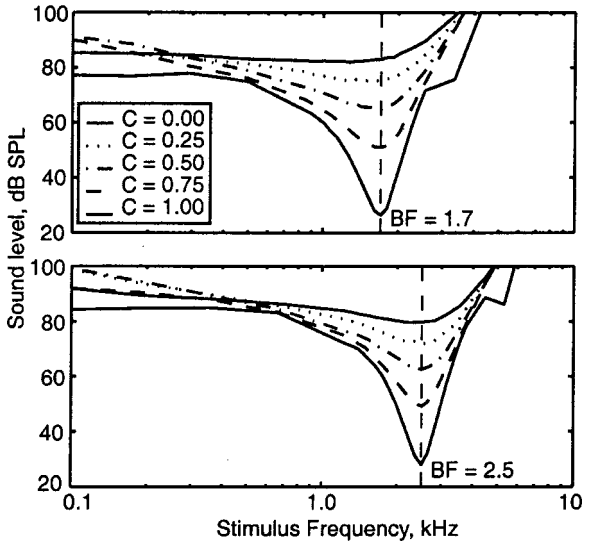

FIGURE 12. Model tuning curves as a function of OHC impairment: no impairment $\Rightarrow C=1.00$; complete impairment $\Rightarrow = C=0.00$. Top panel: BF = 1.7 kHz; bottom panel: BF = 2.5 kHz.

data in Fig. 7(C). In the normal population response [Fig. 13(A)], synchrony capture is observed at the first and second formants, as expected. The third-formant response is weaker than it should be because of inaccuracies in the thresholds of high BF fibers. With this exception, the model accurately reproduces the behavior of the data shown in Fig. 6. The impaired population response in Fig. 13(B) shows the upward spread of synchrony to F1 and the reduction in response to F2 and F3 observed in the data (cf. Fig. 9). The response to F2 is weaker than in the real data and the F1 second-harmonic response at the 1 kHz place is stronger than is seen in the data. However, the general features of the impaired ear are captured by the model. Figure 13(C) shows the modeled response to the CEFS vowel for the impaired population. The model shows response enhancements similar to the data in Fig. 10, in that the spread of F1 is limited and F2 responses dominate at the F2 place. However, as in the real data, the F2 response continues to spread to fibers with higher BFs and there is still loss of F3 response.

So, at least in the case of severely impaired OHCs, the model can reproduce reasonably well the responses of impaired auditory-nerve fibers to both natural vowels and to signal-processed vowels. These results encourage

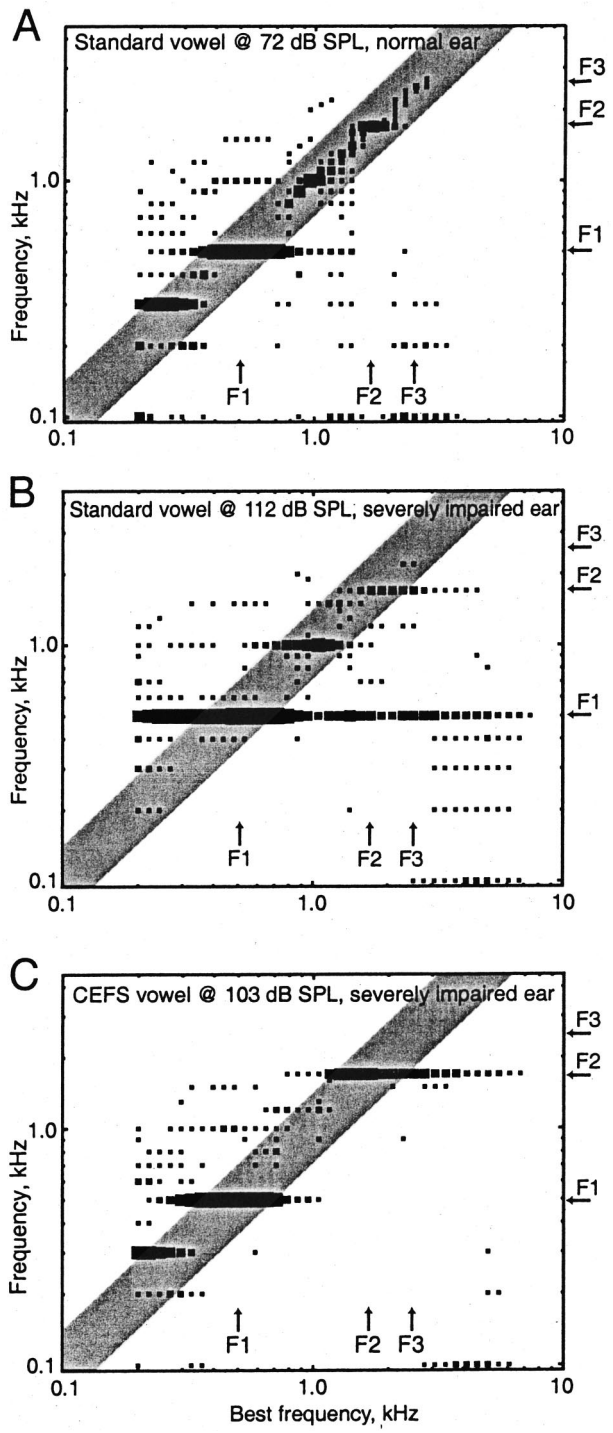

FIGURE 13. Population responses to the vowel predicted by the model. (A) Normal ear, standard vowel. (B) Severely impaired ear, standard vowel. (C) Severely impaired ear, CEFS vowel. Plotting conventions as for Fig. 6.

us to suggest that such models will be useful in suggesting and testing new signal-processing schemes for hearing aids over a wide range of stimulus and impairment conditions. We are currently developing and testing a hearing-impaired version of a newer model by Carney and colleagues that includes two-tone suppression,[41] which may be important in describing auditory-nerve responses for mild to moderate OHC impairment as well as the effects of IHC impairment.[17,39]

FUTURE TRENDS

We turn finally to our view of the future of biomedical engineering in auditory research. We foresee both steady evolutions in the field and what may be called paradigm-shifting changes. The kinds of modeling efforts we have described will evolve as more is known about the underlying cochlear mechanisms. For example, it is widely thought that the outer hair-cell control of BM motion is mediated by outer hair-cell ''electromotility.'' [2] In response to changes in membrane potential, the outer hair cell rapidly alters its length. Already, there have been attempts to model this electromotility at the subcellular level,[5,28] including some very detailed finite-element analyses at subcellular levels.[37]

Recently, a major step in understanding electromotility has been the identification of the putative motor protein in the outer hair cells.[42] Voltage-induced shape changes and a nonlinear membrane capacitance are observed in cultured human kidney cells when they are induced to express this protein. The more we are able to incorporate new understanding of the molecular mechanisms underlying hair-cell function into our models, the more confident we will be in technological developments based on those models.

A real paradigm shift may be currently in the making that someday could obviate the need for hearing aids. It has been known for some time that hair cells in nonmammalian species can regenerate following inner-ear damage.[4,38] One of the ''Holy Grails'' of hearing research is the molecular key or keys that will allow human cochlear hair cells to regenerate.

Recently, it was reported that overexpression of the transcription factor Math1 in postnatal rat cochlear explant cultures resulted in the appearance of extra hair cells.[43] The ectopic hair cells were found medial to the organ of Corti, and their source was columnar epithelial cells, located outside of the sensory epithelium. Although this result suggests one exciting path to the Holy Grail, it also suggests a number of ways in which that path will be paved with engineering advances. First, it is obviously a very long way from transfecting random cells in an explanted cochlea to transfecting cells in an intact live human cochlea, and even farther to transfecting targeted cochlear cells. This ultimately will require major advances in gene delivery systems that are likely to be developed by biomedical engineers. Second, the ectopic placement of these extra hair cells is an extreme example of a problem that faces all efforts at regeneration of the

damaged nervous system. We cannot expect to find the magic key to regeneration, be it a growth factor, a gene, a stem cell, or something now unforeseen, and simply turn the key to restore function. Instead, it is necessary to reconstruct complex tissues, perhaps by recapitulating part of the developmental sequence. This process is likely to be long and require many steps. A role for the kind of research described here is to assist in the evaluation and debugging of the results along the way.

The auditory nerve is the output of the first stage of signal processing in the auditory system. It is also the input to the central nervous system. Over the past 15–20 years a number of biomedical engineers have turned to studying speech processing in the central auditory system. For example, on the basis of a simple model for speech signal processing in the cochlear nucleus we hypothesized and then identified experimentally a population of cochlear nucleus cells that may play a crucial role in the perception of speech.[1]

But the cochlear nucleus is only one synapse into the central nervous system (CNS). In our minds the great frontier in auditory neuroscience is in unraveling the central processing responsible for auditory perception. This process will involve developing the ability to measure relevant response properties from freely moving animals involved in auditory perceptual tasks. The technology development alone needed to accomplish this goal will be a biomedical engineering challenge of first magnitude. But beyond the technology, just the words ''relevant response properties'' implies a knowledge of the issues of systems as complex as the CNS that may require a whole new conceptual framework. While we really do not have a good idea about what form that framework might take, we do believe that biomedical engineers, who do not fear complexity, will lead the way to this new frontier.

ACKNOWLEDGMENTS

This research was supported by NIH (NIDCD) Grant No. DC00109. The experimental data presented were originally taken by R. L. Miller in collaboration with K. D. Franck, J. R. Schilling, J. C. Wong, and B. M. Calhoun. Technical assistance was provided by Phyllis Taylor, Ron Atkinson, and Cynthia Aleszczyk.

REFERENCES

[1] Blackburn, C. C., and M. B. Sachs. The representations of the steady-state vowel sound /ɛ/ in the discharge patterns of cat anteroventral cochlear nucleus neurons. *J. Neurophysiol.* 63:1191–1212, 1990.

[2] Brownell, W. E., C. R. Bader, D. Bertrand, and Y. de Ribaupierre. Evoked mechanical responses of isolated cochlear outer hair cells. *Science* 227(4683):194–196, 1985.

[3] Carney, L. H. A model for the responses of low-frequency auditory-nerve fibers in cat. *J. Acoust. Soc. Am.* 93:401–417, 1993.

[4] Corwin, J. T., and J. C. Oberholtzer. Fish n' chicks: Model recipes for hair-cell regeneration? *Neuron* 19:951–954, 1997.

[5] Dallos, P., R. Hallworth, and B. N. Evans. Theory of electrically driven shape changes of cochlear outer hair cells. *J. Neurophysiol.* 70:299–323, 1993.

[6] de Boer, E. Mechanics of the cochlea: Modeling efforts. In: The Cochlea, edited by P. Dallos, A. N. Popper, and R. R. Fay. New York: Springer, 1996, pp. 258–317.

[7] Deng, L., and C. D. Geisler. Responses of auditory-nerve fibers to nasal consonant–vowel syllables. *J. Acoust. Soc. Am.* 82:1977–1988, 1987.

[8] Deng, L., C. D. Geisler, and S. Greenberg. Responses of auditory-nerve fibers to multitone complexes. *J. Acoust. Soc. Am.* 82:1989–2000, 1987.

[9] Evans, E. F., and J. P. Wilson. The frequency selectivity of the cochlea. In: Basic Mechanisms in Hearing, edited by A. R. Møller. New York: Academic, 1973, pp. 519–554.

[10] Geisler, C. D. From Sound to Synapse: Physiology of the Mammalian Ear. New York: Oxford, 1998.

[11] Helmholtz, H. On the Sensations of Tone. New York: Dover, 1954.

[12] Johnson, D. H. Point process models of single-neuron discharges. *J. Comput. Neurosci.* 3:275–299, 1996.

[13] Johnson, D. H., and A. Swami. The transmission of signals by auditory-nerve fiber discharge patterns. *J. Acoust. Soc. Am.* 68:1115–1122, 1980.

[14] Kros, C. J. Physiology of mammalian cochlear hair cells. In: The Cochlea, edited by P. Dallos, A. N. Popper, and R. R. Fay. New York: Springer, 1996, pp. 318–385.

[15] Liberman, M. C. Single-neuron labeling and chronic cochlear pathology. I. Threshold shift and characteristic-frequency shift. *Hear. Res.* 16:33–41, 1984.

[16] Liberman, M. C., and L. W. Dodds. Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hear. Res.* 16:55–74, 1984.

[17] Miller, R. L., J. R. Schilling, K. R. Franck, and E. D. Young. Effects of acoustic trauma on the representation of the vowel /ɛ/ in cat auditory-nerve fibers. *J. Acoust. Soc. Am.* 101:3602–3616, 1997.

[18] Miller, R. L., B. M. Calhoun, and E. D. Young. Contrast enhancement improves the representation of /ɛ/-like vowels in the hearing-impaired auditory nerve. *J. Acoust. Soc. Am.* 106:2693–2708, 1999.

[19] Molnar, C. E., and R. R. Pfeiffer. Interpretations of spontaneous discharge patterns of neurons in the cochlear nucleus. *Proc. IEEE* 56:993–1004, 1966.

[20] Moore, B. C. J. Perceptual Consequences of Cochlear Damage. New York: Oxford, 1995.

[21] Narayan, S. S., A. N. Temchin, A. Recio, and M. A. Ruggero. Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae. *Science* 282(5395):1882–1884, 1998.

[22] NIH Consensus Statement. *Noise Hear. Loss* 8:1–24, 1990.

[23] Nobili, R., F. Mammano, and J. Ashmore. How well do we understand the cochlea? *TINS* 21:159–167, 1998.

[24] Patuzzi, R. B. Cochlear micromechanics and macromechanics. In: The Cochlea, edited by P. Dallos, A. N. Popper, and R. R. Fay. New York: Springer, 1996, pp. 186–257.

[25] Peterson, G. E., and H. L. Barney. Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* 24:175–184, 1952.

[26] Pfeiffer, R. R. Classification of response patterns of action-potential discharges for units in the cochlear nucleus: Tone-

burst stimulation. *Exp. Brain Res.* 1:220–235, 1966.
[27] Pols, L. C., L. J. van der Kamp, and R. Plomp. Perceptual and physical space of vowel sounds. *J. Acoust. Soc. Am.* 46:458–467, 1969.
[28] Raphael, R. M., A. S. Popel, and W. E. Brownell. A membrane bending model of outer hair cell electromotility. *Biophys. J.* 78:2844–2862, 2000.
[29] Robles, L., and M. A. Ruggero. Mechanics of the mammalian cochlea. *Physiol. Rev.* 81:1305–1352, 2001.
[30] Rose, J. E., J. F. Brugge, D. J. Anderson, and J. E. Hind. Phase-locked response to low-frequency tones in single auditory-nerve fibers of the squirrel monkey. *J. Neurophysiol.* 30:769–793, 1967.
[31] Ruggero, M. A., and N. C. Rich. Furosemide alters organ of Corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane. *J. Neurosci.* 11:1057–1067, 1991.
[32] Sachs, M. B. Speech encoding in the auditory nerve. In: Hearing Science, Recent Advances, edited by C. I. Berlin, San Diego: College-Hill, 1984, pp. 263–307.
[33] Schilling, J. R., R. L. Miller, M. B. Sachs, and E. D. Young. Frequency-shaped amplification changes the neural representation of speech with noise-induced hearing loss. *Hear. Res.* 117:57–70, 1998.
[34] Sewell, W. F. Neurotransmitters and synaptic transmission. In: The Cochlea, edited by P. Dallos, A. N. Popper, and R. R. Fay. New York: Springer, 1996, pp. 503–553.
[35] Siebert, W. M. Some implications of the stochastic behavior of primary auditory neurons. *Kybernetik* 2:206–215, 1965.
[36] Smith, R. L., and J. J. Zwislocki. Responses of some neurons in the cochlear nucleus to tone-intensity increments. *J. Acoust. Soc. Am.* 50:1520–1525, 1971.
[37] Spector, A. A., M. Ameen, and A. S. Popel. Simulation of motor-driven cochlear outer hair cell electromotility. *Biophys. J.* 81:11–24, 2001.
[38] Stone, J. S., and E. W. Rubel. Cellular studies of auditory hair cell regeneration in birds. *Proc. Natl. Acad. Sci. U.S.A.* 97:11714–11721, 2000.
[39] Wong, J. C., R. L. Miller, B. M. Calhoun, M. B. Sachs, and E. D. Young. Effects of high sound levels on responses to the vowel /eh/ in cat auditory nerve. *Hear. Res.* 123:61–77, 1998.
[40] Young, E. D., and M. B. Sachs. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. *J. Acoust. Soc. Am.* 66:1381–1403, 1979.
[41] Zhang, X., M. G. Heinz, I. C. Bruce, and L. H. Carney. A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J. Acoust. Soc. Am.* 109:648–670, 2001.
[42] Zheng, J., W. Shen, D. Z. He, K. B. Long, L. D. Madison, and P. Dallos. Prestin is the motor protein of cochlear outer hair cells. *Nature (London)* 405(6783):149–155, 2000.
[43] Zheng, J. L., and W. Q. Gao. Overexpression of Math 1 induces robust production of extra hair cells in postnatal rat inner ears. *Nat. Neurosci.* 3:580–586, 2000.
[44] Zweig, G., R. Lipes, and J. R. Pierce. The cochlear compromise. *J. Acoust. Soc. Am.* 59:975–982, 1976.