

# Responses to cochlear normalized speech stimuli in the auditory nerve of cat

Alberto Recio<sup>a)</sup> and William S. Rhode

*Department of Physiology, University of Wisconsin, Madison, Wisconsin 53706*

Michael Kieft

*School of Human Communication Disorders, Dalhousie University, Halifax, Nova Scotia B3H 1R2, Canada*

Keith R. Kluender

*Department of Physiology and Department of Psychology, University of Wisconsin, Madison,*

*Wisconsin 53706*

(Received 14 December 2001; revised 18 February 2002; accepted 22 February 2002)

Previous studies of auditory-nerve fiber (ANF) representation of vowels in cats and rodents (chinchillas and guinea pigs) have shown that, at amplitudes typical for conversational speech (60–70 dB), neuronal firing rate as a function of characteristic frequency alone provides a poor representation of spectral prominences (e.g., formants) of speech sounds. However, ANF rate representations may not be as inadequate as they appear. Here, it is investigated whether some of this apparent inadequacy owes to the mismatch between animal and human cochlear characteristics. For all animal models tested in earlier studies, the basilar membrane is shorter and encompasses a broader range of frequencies than that of humans. In this study, a customized speech synthesizer was used to create a rendition of the vowel [ε] with formant spacing and bandwidths that fit the cat cochlea in proportion to the human cochlea. In these vowels, the spectral envelope is matched to cochlear distance rather than to frequency. Recordings of responses to this cochlear normalized [ε] in auditory-nerve fibers of cats demonstrate that rate-based encoding of vowel sounds is capable of distinguishing spectral prominences even at 70–80-dB SPL. When cochlear dimensions are taken into account, rate encoding in ANF appears more informative than was previously believed.

© 2002 Acoustical Society of America. [DOI: 10.1121/1.1468878]

PACS numbers: 43.64.Pg, 43.64.Sj [LHC]

## I. INTRODUCTION

Numerous studies have explored encoding of speech sounds in the peripheral and central auditory system via *in vivo* studies in animal models (e.g., Sachs and Young, 1979; Blackburn and Sachs, 1990; Recio and Rhode, 2000). Investigations of auditory-nerve fiber (ANF) responses reveal that a representation of vowel spectra based on the individual firing rates of ANFs is inadequate for transmitting information known to be conveyed in humans at levels typical of conversational speech (60–70 dB). At higher presentation levels, firing rates can saturate, resulting in nearly flat frequency-rate curves that do not resolve frequencies in complex sounds. In part, this is due to the limited dynamic range of mid- and high-spontaneous fibers (typically less than 24 dB; Schalk and Sachs, 1979). Saturation alone does not explain all of the apparent lack of frequency selectivity because peak response rates often are substantially lower than the saturation level in response to characteristic-frequency (CF) tones. It has been suggested that two-tone suppression, by which energy in one region of the spectrum (e.g., first formant  $F_1$ ) suppresses responses to energy in another region (e.g., second and third formants,  $F_2$  and  $F_3$ ), could account for the lower response rates (Sachs and Young, 1979). Inter-fiber variability, which can be due to differences in threshold

or dynamic range, can account for some of the poor performance of the rate code at high stimulus levels (Miller *et al.*, 1999). When much of this variability is factored out, as in the case of analyzing rate differences in responses to two stimuli, rate becomes more efficient for encoding differences between those two stimuli even at 70-dB SPL (Conley and Keilson, 1995).

Responses to vowel sounds similar to those in ANF have been recorded in neurons of cochlear nucleus (CN: Blackburn and Sachs, 1990; Recio and Rhode, 2000). Neuron types differed in their responses. Medium- and high-spontaneous nerve fibers provide a poorer rate-based representation of vowel spectra than low-spontaneous fibers, which can encode spectral peaks of speech sounds even at 70–80 dB SPL (Sachs and Young, 1979; Blackburn and Sachs, 1990). In CN, transient chopper neurons also maintain an accurate rate-based representation of spectra of vowels (Blackburn and Sachs, 1990; Recio and Rhode, 2000).

Whereas studies cited above included close attention to the fidelity of computer-generated vowels, they overlooked some potentially important differences between human auditory systems and those of the animals under investigation. Cat and chinchillas, which have been especially prominent animal subjects in earlier studies, have a considerably shorter cochlea than that of humans (e.g., 25 and 17.5 mm for cat and chinchilla, respectively, versus 35 mm for human). Furthermore, cats are sensitive to a considerably broader range

<sup>a)</sup>Electronic mail: recio@physiology.wisc.edu

of frequencies than humans (up to 60 kHz compared to 20 kHz for humans). These characteristics are significant when one wishes to understand neural representation of speech sounds. For example, positions of the center frequencies of the first three spectral prominences ( $F_1$ ,  $F_2$ , and  $F_3$ ) of the vowel [ε] (as in “bet”) have been synthesized at 512, 1792, and 2432 Hz, respectively, corresponding closely to means observed from natural productions (Peterson and Barney, 1952; Sachs and Young, 1979). According to Greenwood’s (1990) estimates, in a human cochlea the distances to the apex for each of those frequencies are 10.21, 17.88, and 19.93 mm, respectively. The same center frequencies in the cat will be located about 3.38, 8.03, and 9.38 mm from the apex. Consequently, the distance between the first two formant peaks ( $F_1$ ,  $F_2$ ) in the cat and chinchilla, 4.65 and 4.14 mm, respectively, is smaller than that measured in humans (7.67 mm). Although one should think of the above distances as approximations—particularly in the case of humans and chinchillas, for which accurate cochlear maps are not available—basilar-membrane (BM) length in both cats and chinchillas is unquestionably shorter relative to humans.

In contrast to the frequency range of cat, the effective frequency range of human sensitivity is considerably smaller. Particularly for the range of frequencies used in speech, human hearing benefits from a longer cochlea that encodes a narrower range of frequencies. We reason that smaller cochlear distances in which vowel sounds are mapped into the cat and chinchilla cochleae may, in part, be responsible for the lack of good spectral resolution observed in rate-based representations of vowels by ANFs.

To test this hypothesis, a modified speech synthesizer (Kiefte *et al.*, 2002) was used to create a cochlear normalized version of the vowel [ε] based on the dimensions of the cat basilar membrane. This synthesis separated spectral peaks (formants) with respect to absolute cochlear distance rather than frequency. In addition, formant bandwidth was scaled to span equal numbers of inner-hair cells for cat and human. The vowel [ε] was selected because its neuronal representation has been the object of many studies (e.g., Sachs and Young, 1979; Young and Sachs, 1979; Blackburn and Sachs, 1990; May *et al.*, 1998; Recio and Rhode, 2000).

## II. METHODS

### A. Animal preparation

Data were obtained from 11 adult cats weighing between 2.5 and 6 kg. Cats were initially anesthetized with an intraperitoneal injection of pentobarbital (50 mg/kg). Additional smaller doses were administered through a catheter inserted in the femoral vein. A thermostatically controlled heating blanket maintained body temperature at 37 °C. After insertion of a tracheal cannula, the left ear was removed and the bulla was vented with 20 cm of 1-mm plastic tube. Removal of the overlying cerebellum made visualization of the cochlear nucleus possible and gentle retraction of the CN, using saline-soaked cotton balls, then exposed the auditory nerve. After covering the nerve with warm agar, a chamber was mounted over the skull opening and filled with mineral oil. The chamber was then sealed with a glass disk holding

the electrode. KCl-filled micropipettes, with impedances of 30–70 MΩ, were used to record single-unit activity.

The protocol for these experiments was approved by the Animal Care Committee of the University of Wisconsin—Madison and meets NIH guidelines.

### B. Acoustic stimuli, experimental protocol, and data analysis

A modified Radio Shack Super Tweeter (Chan *et al.*, 1993) served as the transducer for these stimuli. After calibration of the acoustic system (100–30 000 Hz, in 100-Hz steps) using a Bruel and Kjaer 0.5-in. condenser microphone, stimuli were digitally compensated for the transfer function of the acoustic system before being presented.

Vowel stimuli, original (human) and cochlear-normalized for cat, were synthesized using the program described in Kiefte *et al.* (2002). For human vowel [ε]<sub>H</sub>, fundamental frequency ( $f_0$ ) was 128 Hz, and center frequencies of  $F_1$ ,  $F_2$ , and  $F_3$  were set at 512, 1792, and 2432 Hz, respectively. Another version of the vowel was created, the cat [ε]<sub>C</sub>, for which cochlear distance between formant peaks was held constant to the human counterpart. In creating this frequency mapping, values of  $f_0$  (128 Hz) remained the same in both versions of the vowel. Formant frequencies for the cochlear-normalized version were obtained by first determining the cochlear position in millimeters of each formant based on Greenwood’s (1990) equation

$$x = (1/a) * \log_{10}(F/A + k), \quad (1)$$

where  $a$ ,  $A$ , and  $k$  are parameters estimated from human cochlear frequency position data,  $x$  is the distance along the basilar membrane from the apex, and  $F$  is the best frequency at location  $x$ . The inverse of this function using parameters appropriate for the cat cochlea is then used to estimate the normalized frequencies. In addition, normalized  $F_1$  was adjusted downwards to 512 Hz by translating the vowel spectrum as a function of cochlear distance by an appropriate constant term in millimeters (Kiefte *et al.*, 2002). This ensured that the peak of  $F_3$  would not be too close to the base of the basilar partition. As mentioned above, the cochlear distance between  $F_1$  and  $F_2$  stimulus in humans is 7.68 mm. In the cat’s cochlea, keeping  $F_1$  at 512 Hz, the value of  $F_2$  needed to preserve the same distance is approximately 3508 Hz. Using the same procedure,  $F_3$  would be approximately 5391 Hz. However, these values were rounded to the nearest harmonic frequency—i.e., 3456 and 5376 Hz for  $F_2$  and  $F_3$ , respectively. Figure 1 shows the amplitude Fourier transform of the human (top panel) and cat (lower panel) vowels.

Duration of cat [ε]<sub>C</sub> and human [ε]<sub>H</sub> vowels was 100 ms. Stimuli were presented 100 times at a rate of 2 per s (400-ms intervening silence) at three levels: 30-, 50-, and 70-dB SPL. In the last set of experiments, stimuli were presented only at 80-dB SPL. All the stimuli were windowed using a raised cosine envelope of 5 ms at each end of the waveform. Cat and human vowels were played at a rate of 100 000 samples per second, using a 16-bit D/A.

During data acquisition, a response-area plot was initially obtained to determine best frequency (BF), threshold, and spontaneous rate (SR) of the fiber. Poststimulus time

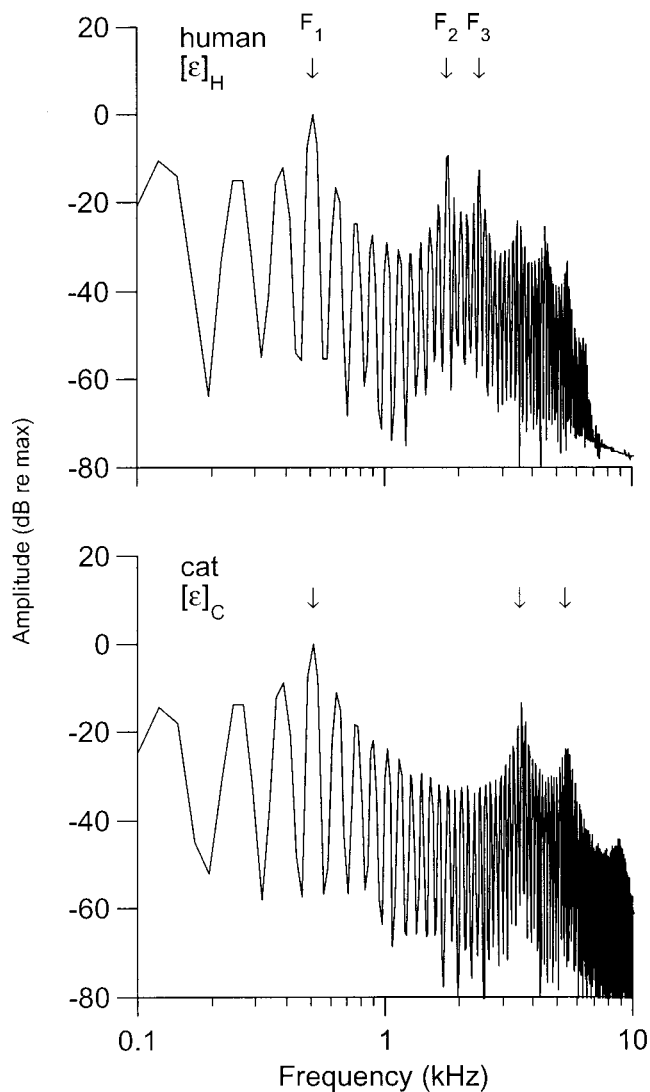


FIG. 1. Fourier transform amplitudes of human  $[\epsilon]_H$  (top panel) and cat  $[\epsilon]_C$  (low panel) sounds. Original waveforms were multiplied by a Hamming window ( $N=4096$ ) and then Fourier transformed. Localizations of  $F_1-F_3$  are indicated by arrows. Sampling rate=100 000 samples/s.

histograms (PSTHs) were routinely obtained using a 50-ms tone with frequency at BF. Rate-intensity curves using a 100-ms BF tone were also obtained to determine the maximum firing rate (MFR). Spikes were stored in the computer for later processing, which consisted mostly of determining the firing rate evoked by the speech sound. Firing rates evoked by the stimuli were estimated during the last 80 ms of the stimulus duration. To facilitate comparisons with previous work (e.g., Sachs and Young, 1979), the driven rate (rate evoked by speech sound-SR) and normalized rate (driven rate/[MFR-SR]) were also obtained for every neuron. For most neurons, the range of values of the normalized rate falls between 0 and 1. Whereas a normalized rate of 1 indicates that the neuron is driven by the speech sound at a rate identical to the maximum firing rate, a normalized rate of 0 indicates a lack of response by the neuron. Negative normalized rates can be a consequence of firing rates below the spontaneous activity level (rate suppression). Most of the analyses and plotting of figures were done in MATLAB.

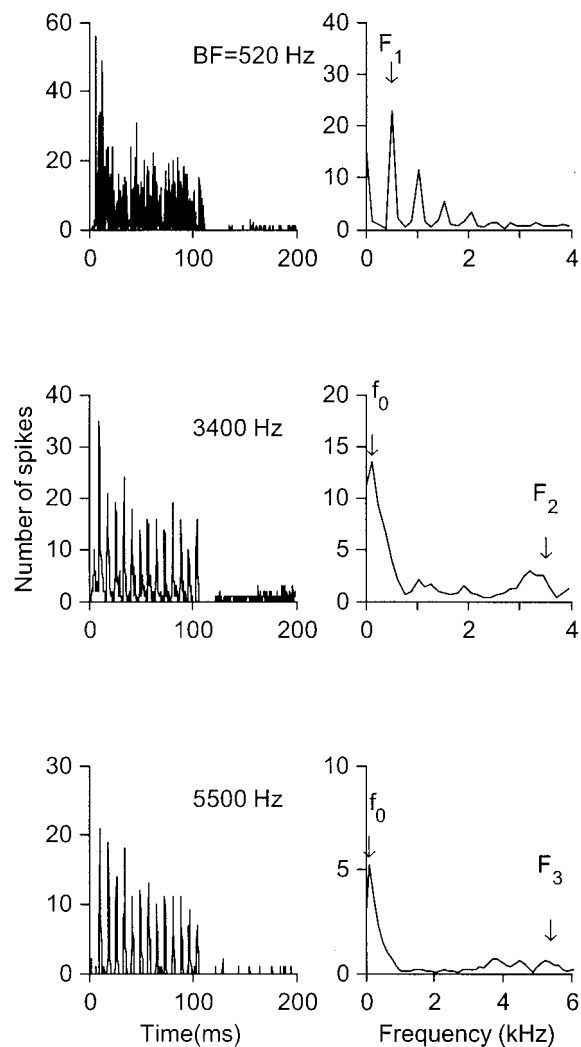


FIG. 2. Poststimulus time histograms (left column) with their corresponding Fourier transform amplitudes (right column) of the responses of three auditory-nerve fibers (ANFs) to the cat vowel  $[\epsilon]_C$ . Fourier transform amplitudes were derived from cycle histograms, obtained with a period =128 Hz ( $f_0$ ), computed from the neural responses. Unit's BF is indicated in each row. Stimuli (duration=100 ms, level=50 dB SPL) were presented 100 times with a repetition period of 0.5 s.

Based on level of spontaneous activity, units were classified as low ( $SR < 1$  spikes/s), medium ( $1 \leq SR < 18$  spikes/s), and high-spontaneous ( $SR \geq 18$  spikes/s) fibers. Data reported here are from 282 units with BFs < 10 kHz.

### III. RESULTS

#### A. Responses of individual fibers

Because ANF responses to the human vowel  $[\epsilon]_H$  have been described elsewhere (e.g., Sachs and Young, 1979; Recio and Rhode, 2000), we will concentrate on responses to the vowel normalized to the cat cochlea. Figure 2 shows the responses, in the form of PSTHs and Fourier transform amplitudes, of three ANFs to cat  $[\epsilon]_C$  at 50-dB SPL. BFs of the fibers in the top, center, and bottom rows are near-center frequencies of  $F_1$ ,  $F_2$ , and  $F_3$ , respectively. The response of low-BF neurons (i.e., neurons with BF around  $F_1$ ) shows synchronization to the peak harmonic of  $F_1$  (512 Hz), with additional phase locking to successive harmonics of 512 Hz

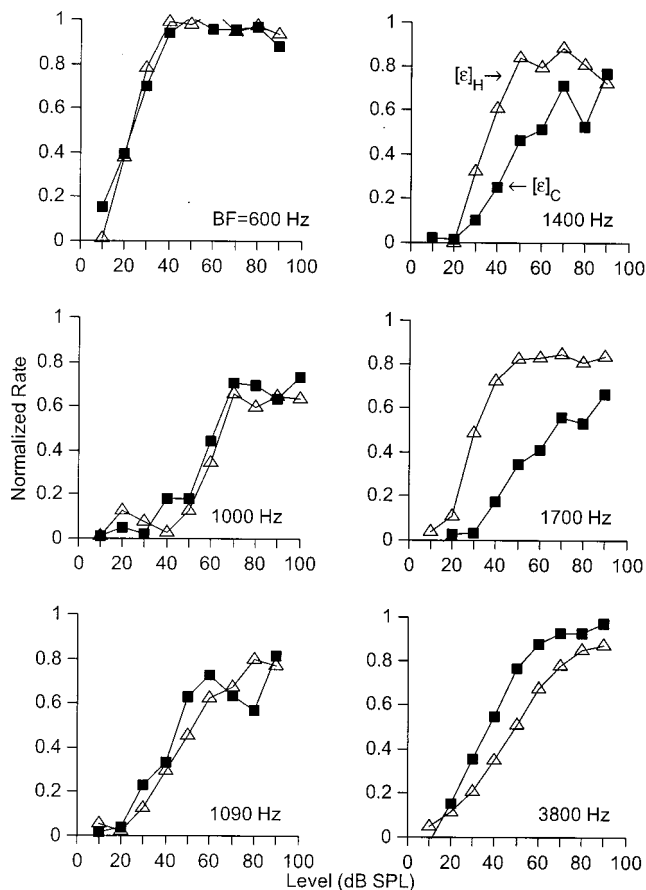


FIG. 3. Normalized rates versus stimulus level curves obtained from responses of six ANFs. Open triangles indicate responses to the human  $[\epsilon]_H$  sound and filled squares show the responses to  $[\epsilon]_C$ . Responses to vowels and BF tones were obtained from ten stimulus presentations (duration=100 ms) with a repetition rate of 0.5 s. Normalized rates were computed using steady-state (20–100-ms) spike count.

(i.e., 1024, 1536, and 2048 Hz). PSTHs of high-BF neurons (i.e., with BFs around  $F_2$  or  $F_3$ ) include peaks separated from each other at an interval equal to the period corresponding to  $f_0$ . This is confirmed in the frequency domain where the largest spectral peak occurs at  $f_0$ . Phase locking to the frequencies of  $F_2$  (3456 Hz) is almost nonexistent.

Normalized rate curves versus stimulus level in response to  $[\epsilon]_H$  and  $[\epsilon]_C$  were also obtained, as shown in Fig. 3. For neurons with BF near  $F_1$ , such as the one in the top-left panel, normalized rates in response to human and cat vowels are nearly identical and reach a maximum value of 1. This is expected because the center frequency of  $F_1$  is set to the same value for both  $[\epsilon]_H$  and  $[\epsilon]_C$ . Even at BFs around 1 kHz (center and bottom left panels), normalized rates in response to  $[\epsilon]_H$  remain similar to those obtained in responses to  $[\epsilon]_C$ . Maximum values of normalized rates, however, are around 0.8. Whether the lower maxima are due to the stimulus not being of sufficiently high amplitude or to rate suppression is unclear. Suppression by  $F_2$ , if any, appears to be independent of the center frequency (or cochlear position) of that formant. Normalized rates for neurons with BF=1400 and 1700 (top and center right) are usually higher when driven by the  $[\epsilon]_H$  (open triangles) than by  $[\epsilon]_C$ , presumably as a consequence of proximity of these BFs to  $F_2$  of  $[\epsilon]_H$

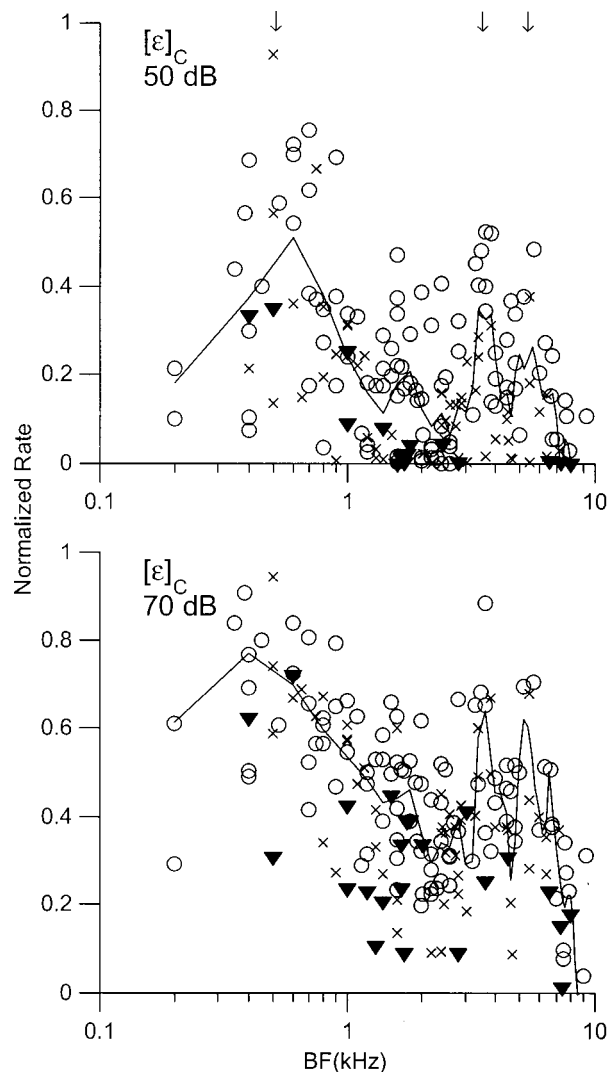


FIG. 4. Normalized rates obtained from steady-state responses of a population of 181 ANFs to  $[\epsilon]_C$  at two stimulus levels (50- and 70-dB SPL). Open circles, crosses, and filled inverted triangles indicate the responses of high-, medium-, and low-spontaneous ANFs, respectively. Continuous lines show the average normalized rate at a given BF, using frequency bins=250 Hz. Arrows indicate the value of  $F_1$ – $F_3$ .

(1792 versus 3456 Hz for  $[\epsilon]_C$ ). Dynamic range ( $\sim 60$  dB) of the responses for some of those neurons to  $[\epsilon]_C$  (filled squares) is also larger than the dynamic range obtained from responses to BF tones. The neuron with BF=3800 (lower right) is driven at a higher rate by  $[\epsilon]_C$  than  $[\epsilon]_H$ . This is not surprising because the BF of that fiber is near  $F_2$  of  $[\epsilon]_C$  (3456 Hz).

## B. Population analysis

Figure 4 shows the responses of a population of ANFs to  $[\epsilon]_C$  at two levels. Responses of high-spontaneous fibers (open circles) yield higher normalized rates than medium- and low-spontaneous units. This was expected because high-spontaneous units have lower thresholds than medium- and low-spontaneous fibers (Lieberman, 1978). Continuous lines show the average discharge rate for medium- and high-spontaneous fibers only. Individual points constituting the lines represent the average normalized rate among those neu-

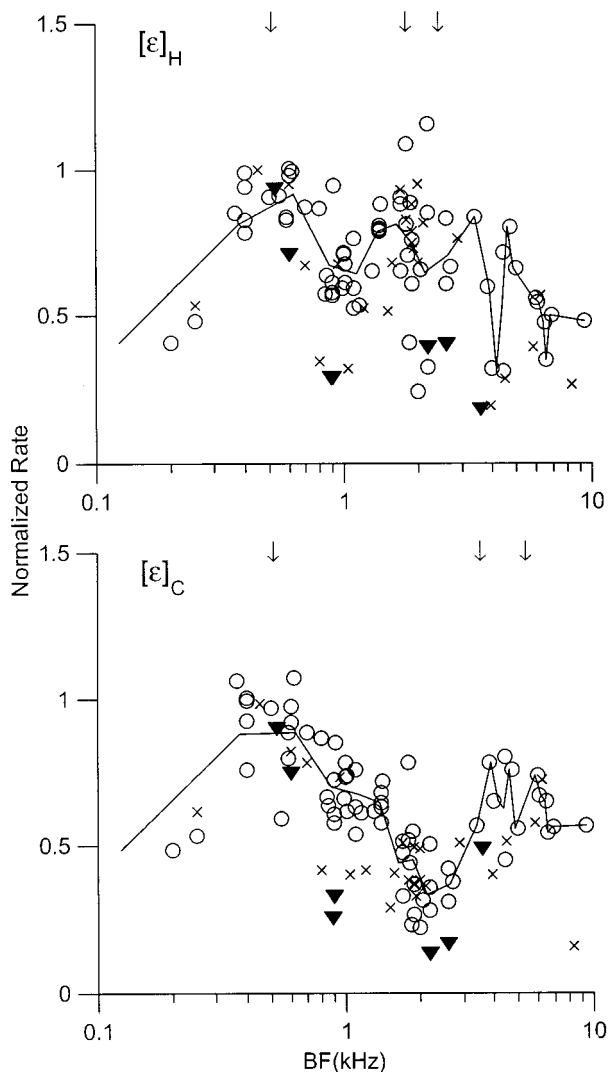


FIG. 5. Normalized rates obtained from steady-state responses of 100 ANFs to  $[\epsilon]_H$  (top panel) and  $[\epsilon]_C$  (lower panel) at 80 dB DPL. Symbols, arrows, and continuous lines represent the same information as in Fig. 4.

rons whose BF is within 125 Hz of the abscissa value (i.e., a 250-Hz window). This was done to facilitate comparisons with the work of Sachs and Young. At both stimulus levels, average rates indicate the presence of three distinct peaks, one for each of the first three formants of the vowels. Sachs and Young (1979) show similar results for responses to a 48-dB vowel sound. In contrast to the results found in the present study, Sachs and Young were unable to show distinct spectral peaks at higher stimulus levels.

Responses to 80-dB vowel sounds obtained from a different neuronal population are shown in Fig. 5, with continuous lines indicating average values. Continuous lines were obtained as described above. Normalized rates obtained from responses to  $[\epsilon]_H$  provide a poor representation of the spectrum of the vowel (i.e., there is little spectral definition), a finding that agrees with the results of Sachs and Young (1979). When the same group of neurons was stimulated using the cochlear-normalized vowel, however, normalized rates show the presence of two spectral peaks, one at the location of  $F_1$  and the other in the region of  $F_2$ – $F_3$ , with a distinct trough in between. The variability in thresholds of

units shown in Fig. 5 was smaller than the variability of the neurons in Fig. 4, which may account for the larger fluctuations observed in normalized rates of Fig. 4.

#### IV. DISCUSSION

The results of this study show that, when the spectral composition of a vowel sound is adjusted to map onto the animal cochlea as a function of cochlear distance as it would fit the human cochlea, rate-based encoding of speech stimuli provides more information than has previously been assumed, even at 70–80-dB SPL. Relatively better preservation of spectral definition is maintained in spite of potential suppressive effects at the level of the basilar membrane. In fact, rate suppression in fibers with BFs near the spectral trough between  $F_1$  and  $F_2$  appears to be independent of the center frequency of  $F_2$ . This can be observed from results from Fig. 5, in which normalized rates values around 1–1.5 kHz are similar in responses to human and cat vowel sounds.

While normalizing spectral composition of speech sounds to nonhuman cochlear dimensions has given new insights into potential encoding of speech in the human auditory system, this spectral adjustment introduces at least one disadvantage. Upward displacement of spectral peaks moves formant center frequencies beyond the range of phase locking. For example,  $F_2$  for  $[\epsilon]_C$  is at 3456 Hz and  $F_3$  is at 5376 Hz. Spectral analysis of responses of neurons with BFs in these regions show a very small spectral peak at those frequencies. For animals such as chinchilla, keeping  $F_1$  at 512 Hz results in a shift in the frequency to  $F_2$  to approximately 4.7 kHz, where phase locking is nearly zero. Although ANFs can phase lock to signals with frequencies in the 2–4-kHz range, the significance of this phase locking is not yet clear. Whereas there is an enhancement of synchrony in the CN to signals below 1 kHz (Joris *et al.*, 1994), there is also deterioration of synchronization indices for CN responses to frequencies above 1–2 kHz relative to those measured in ANFs (Rhode and Smith, 1986). Further decrease in synchrony occurs at higher levels in the auditory system, such as in the inferior colliculus and auditory cortex.

A final potential disadvantage is the implicit assumption that basilar-membrane mechanics is the same in the apex as in the base of the cochlea, or in intermediate points. With the exception of early measurements in the squirrel monkey (Rhode, 1971), direct measurements of BM motion exist only in the base and apex of certain animals. Results from the apex (Rhode and Cooper, 1996) indicate that the amount of nonlinearity is smaller than in the base, yet frequency sensitivity is broader near the apex than at the base. Indirect measurements of BM motion, such as those obtained from *revcor* functions of ANFs (Carney *et al.*, 1999) also show differences in impulse responses of fibers near the apex (BF < 500 Hz) relative to those measured from ANFs with BF > 1 kHz. Therefore, it is conceivable that mechanics for the cochlear location with BF in the human at 1792 Hz might not be strictly comparable to the corresponding location in cat cochlea (3456 Hz).

It is important to note that these shortcomings may be problems for cat or chinchilla, but this does not imply a problem for humans. For cat and chinchilla, one cannot have

cochlear distances between formants equal to those for humans and have spectral peaks for  $F_2$  and  $F_3$  within the region of reliable phase locking. For humans, with shorter frequency range and longer cochleae, it is possible to maintain cochlear distance and have phase locking, too. Of course, little is known about the mechanics of living human cochleae.

Limitations aside, the present results suggest that rate encoding of human speech sounds may be significantly more robust in the human ANF than would be expected based upon earlier measures with human speech sounds presented to nonhuman auditory systems.

## ACKNOWLEDGMENTS

Thanks to Laurel Carney and two anonymous reviewers for their comments. Work supported by NIH Grant Nos. NS-17590 and DC-04072.

Blackburn, C. C., and Sachs, M. B. (1990). "The representation of the steady-state vowel sound /ε/ in the discharge patterns of cat anteroventral cochlear nucleus neurons," *J. Neurophysiol.* **63**, 1191–1212.

Carney, L. H., McDuffy, M. J., and Shekhter, I. (1999). "Frequency glides in the impulse responses of auditory nerve fibers," *J. Acoust. Soc. Am.* **105**, 2384–2391.

Chan, J. C. K., Musicant, A. D., and Hind, J. E. (1993). "An insert earphone system for delivery of spectrally shaped signals for physiological studies," *J. Acoust. Soc. Am.* **93**, 1496–1501.

Conley, R. A., and Keilson, S. E. (1995). "Rate representation and discriminability of second formant frequencies for /ε/-like steady-state vowels in cat auditory nerve," *J. Acoust. Soc. Am.* **98**, 3223–3234.

Greenwood, D. D. (1990). "A cochlear frequency-position function for sev-

eral species—29 years later," *J. Acoust. Soc. Am.* **87**, 2592–2605.

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. T. (1994). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency," *J. Neurosci.* **71**, 1022–1036.

Kieft, M., Kluender, K. R., and Rhode, W. S. (2002). "Synthetic speech stimuli spectrally normalized for nonhuman cochlear dimensions," *ARLO* **3**, 41–46.

Liberman, M. C. (1978). "Auditory-nerve responses from cats raised in a low-noise chamber," *J. Acoust. Soc. Am.* **63**, 442–455.

May, B. J., LePrell, G. S., and Sachs, M. B. (1998). "Vowel representation in the ventral cochlear nucleus of the cat: Effect of level, background noise, and behavioral state," *J. Neurophysiol.* **79**, 1755–1767.

Miller, R. L., Calhoun, B. M., and Young, E. D. (1999). "Discriminability of vowel representations in cat auditory-nerve fibers after acoustics trauma," *J. Acoust. Soc. Am.* **105**, 311–325.

Peterson, G., and Barney, H. (1952). "Control methods used in a study of the vowels," *J. Acoust. Soc. Am.* **24**, 175–184.

Recio, A., and Rhode, W. S. (2000). "Representation of vowel stimuli in the ventral cochlear nucleus of the chinchilla," *Hear. Res.* **146**, 167–184.

Rhode, W. S. (1971). "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique," *J. Acoust. Soc. Am.* **49**, 1218–1231.

Rhode, W. S., and Cooper, N. P. (1996). "Nonlinear mechanics in the apical turn of the chinchilla cochlea *in vivo*," *Aud. Neurosci.* **3**, 101–121.

Rhode, W. S., and Smith, P. H. (1986). "Encoding timing and intensity in the ventral cochlear nucleus of the cat," *J. Neurophysiol.* **56**, 261–286.

Sachs, M. B., and Young, E. D. (1979). "Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate," *J. Acoust. Soc. Am.* **66**, 470–479.

Schalk, T. B., and Sachs, M. B. (1979). "Nonlinearities in auditory-nerve fiber responses to bandlimited noise," *J. Acoust. Soc. Am.* **67**, 903–913.

Young, E. D., and Sachs, M. B. (1979). "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.