PREPRINT: Accepted April 5, 2007, for publication in

July 2007 issue of *J. Acoust. Soc. Am.*

Representation of the vowel /ɛ/ in normal and impaired auditory nerve fibers: Model predictions of responses in cats

Zilany, Muhammad S. A.

Bruce, Ian C.[a)]

Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada L8S 4K1.

[a)] Corresponding author address: Department of Electrical and Computer Engineering, Room ITB-A213, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1. Electronic mail: ibruce@ieee.org

Running Head: Predictions of auditory nerve vowel representation

Received:

ABSTRACT

The temporal response of auditory-nerve (AN) fibers to a steady-state vowel is investigated using a computational auditory-periphery model. The model predictions are validated against a wide range of physiological data for both normal and impaired fibers in cats. The model incorporates two parallel filter paths, component 1 (C1) and component 2 (C2), which correspond to the active and passive modes of basilar membrane vibration, respectively, in the cochlea. The outputs of the two filters are subsequently transduced by two separate functions, added together, and then low-pass filtered by the inner hair cell (IHC) membrane, which is followed by the IHC-AN synapse and discharge generator. The C1 response dominates at low and moderate levels and is responsible for synchrony capture and multi-formant responses seen in the vowel responses. The C2 response dominates at high levels and produces the loss of synchrony capture observed in normal and impaired fibers. The interaction between C1 and C2 responses explains the behavior of AN fibers in the transition region, which is characterized by two important observations in the vowel responses: first, all components of the vowel undergo the C1/C2 transition simultaneously, and second, the responses to the non-formant components of the vowel become substantial.

PACS Numbers: 43.64.Sj, 43.64.Bt, 43.64.Pg, 43.64.Wn

I. INTRODUCTION

Toward the goal of understanding the mechanisms by which speech signals are represented in the brain, auditory-nerve (AN) fiber responses to speech-like sounds, especially for vowels, have been investigated by several researchers over the last few decades (Young and Sachs, 1979; Delgutte, 1980; Sinex and Geisler, 1983; Delgutte and Kiang, 1984; Palmer et al., 1986; Miller et al., 1997; Wong et al., 1998). A number of different cochlear nonlinearities prove important in encoding vowel spectra in both normal and impaired ears at sound pressure levels (SPLs) spanning the dynamic range of hearing. Responses to vowels presented at high sound pressure levels are interesting because they highlight some very striking nonlinearities in the normal cochlea and provide insight into the ability of amplification to restore normal vowel representations in the impaired cochlea (Miller et al., 1997; Wong et al., 1998; Schilling et al., 1998; Miller et al., 1999a,b). Some recent studies (Sachs et al., 2002; Bruce et al., 2003) have modelled AN responses to vowels presented at low to moderate stimulus levels with a sufficient accuracy but fail to account for a number of features of high stimulus level responses observed in the physiological experiments. Holmes et al. (2004) modelled vowel responses in guinea pigs and made comparisons with some of the cat data, but the guinea pig vowel data is substantially less extensive than the cat vowel data. In particular, no guinea pig vowel data have been published for high presentation levels or for impaired ears. In this paper, we have investigated the temporal aspects of cat AN fiber coding of a steady-state vowel in normal and impaired ears using a computational cat auditory-periphery model (Zilany and Bruce, 2006), which is suitable for representing responses to simple and complex stimuli such as tones, two-tone and broadband stimuli across a wide range of sound pressure levels.

An important observation in normal fibers at sound pressure levels corresponding to conversational speech is that the response is phase-locked almost exclusively to the formant frequency closest to the fiber's characteristic frequency (CF)[1], a phenomenon referred to as synchrony cap-

ture (Young and Sachs, 1979; Deng and Geisler, 1987b; Miller et al., 1997). However, as the sound level increases to around 80 dB SPL or above, fibers start to show synchrony to other stimulus frequency components, i.e., the responses of the fibers with CFs near one formant now also respond to other formants, particularly to the lower-frequency formants and to the distortion products of the formants (Wong et al., 1998). This is presumably due to broadening of the auditory filters with stimulus level and nonlinearity of the cochlea. At very high sound pressure levels, fibers respond in a broadband fashion, meaning that synchrony to frequencies other than formants also becomes substantial. The degradation in the temporal representation (i.e., the spread of synchrony) is also a characteristic of responses to the vowel in impaired fibers in cats (Miller et al., 1997; Wong et al., 1998; Schilling et al., 1998). Thus, the normal tonotopic representation of the vowel is substantially degraded in impaired and normal fibers at high levels.

The spread of synchrony and the synchrony capture phenomenon itself reflect nonlinear signal processing in the cochlea. Another nonlinear change in AN fiber responses seen at high levels is the component 1 (C1) - component 2 (C2) transition, which is characterized by a sharp change in the phase-level function and occasionally an accompanying dip in the rate-level function. Wong et al. (1998) reported that the C1/C2 transition occurs in AN fiber responses to a vowel as well as to tones. However, they behave differently in that in response to a vowel all components undergo the C1/C2 transition at approximately the same vowel level. This observation rules out the possibility of separate processing of the individual components of the vowel (Wong et al., 1998). On the other hand, individual tones cause AN fiber to undergo C1/C2 transition in a fairly narrow range of levels (Liberman and Kiang, 1984; Wong et al., 1998). Consequently, to reach C2 threshold (C1/C2 transition level), the level required by a particular frequency component in vowel could be significantly lower than the level required for a tone at that frequency.

A number of previous attempts had been made to predict AN responses to speech-like stimuli (e.g., Deng and Geisler, 1987a; Geisler, 1989; Jenison et al., 1991; Sachs et al., 2002; Bruce et al., 2003; Holmes et al., 2004). The cat vowel responses we attempt to describe in this paper were

previously studied by Bruce et al. (2003) using a model that incorporates many of the details of the signal processing in the cochlea and has scaling functions in the outer hair cell (OHC) and inner hair cell (IHC) sections of the model to facilitate simulation of impairment in the cochlea. The model by Bruce et al. (2003) gives good qualitative and quantitative descriptions of the vowel responses at low to moderate levels only. This is expected, because this model includes only a single mode of vibration to the IHC that produces the C1 responses only. To address high level effects such as the C1/C2 transition, peak splitting, and the shift in the best frequency (BF)[1] with level, Zilany and Bruce (2006) significantly extended the model of Bruce et al. (2003) by including a C2 filter, parallel to the signal-path C1 filter. The selection of the C2 filter followed by a separate transduction function (referred to as the C2 transduction function) has been motivated by some physiological observations in AN fibers (for details see Zilany and Bruce, 2006). The output of the C2 filter can be regarded as a secondary passive mode of vibration of the basilar membrane in the cochlea, whereas the output of C1 filter resembles the primary active mode of vibration. This model not only better describes AN fiber responses at high levels but also gives insight into the possible physical mechanisms of the C2 response and the C1/C2 interaction. Another improvement made in Zilany and Bruce (2006) that contributes to accurate representations at high levels is that the C1 filter, originally a gammatone filter in Bruce et al. (2003), has been replaced by a chirp filter, which is responsible for the level-independent instantaneous frequency glide in the impulse response of AN fibers (Carney et al., 1999; Tan and Carney, 2003). Thus the model can simulate the shift in BF the occurs at high levels, which is partially responsible for the loss of synchrony capture seen at high levels. The increased accuracy of the model predictions in this paper also results from the fact that we have taken into account the frequency response of the sound delivery system used in the physiological studies. This has been done to closely match the vowel's spectral shape at the tympanic membrane, as was presented during physiological experiments.

By using accurate cochlear models and providing quantitative simulations, we can predict how certain underlying mechanisms may affect the perception of speech. In addition, a more accurate

model of cochlear processing will lead to a deeper understanding of how the various physiological changes account for the change in the neural responses associated with hearing impairment. The ability to break down a single hearing loss to its constituent components may enable us to design strategies that could restore the normal neural representation for listeners with hearing loss. In this paper, we have investigated the various AN fiber nonlinearities seen in the cat vowel responses using a computational auditory-periphery model (Zilany and Bruce, 2006). The accuracy of the model predictions are assessed by comparison with the reported physiological data in cats (Miller et al., 1997; Wong et al., 1998). The model responses to other stimuli such as tones and broadband noise are described in Zilany and Bruce (2006). In Sec. II., the vowel stimuli are described and the model is explained briefly, along with a description of how the acoustic trauma from the physiological studies is modeled. The model predictions are described in Sec. III. and compared with the published physiological data.

II. METHODS

The representation of the synthesized vowel /ɛ/ in cat AN fibers (Miller et al., 1997; Wong et al., 1998) has been investigated in this paper using an auditory-periphery model developed by Zilany and Bruce (2006). To be consistent with the physiological experiments, the same stimuli and response analysis methods have been utilized here.

A. Stimuli

Following the methods used in the physiological studies, we have filtered the vowel /ɛ/ (Wong et al., 1998) and the stimulus "besh" (Miller et al., 1997) by the head related transfer function (HRTF) of the human head (Wiener and Ross, 1946). Note that the synthesized syllable "besh" has the vowel spectrum identical to that of the isolated vowel, and the responses are investigated only in the steady-state vowel portion. The HRTF-filtered stimulus is then passed through a fil-

ter that approximately matches the frequency response of the sound delivery system used in the experiments (Miller and Young, private communication). This correction for the sound delivery system is necessary as the magnitude of the vowel in the F2 region (delivered to the tympanic membrane) is often several dB lower than that of the nominal vowel (Wong et al., 1998). The spectral representation of the vowel is shown in Fig. 1, where the thin line shows the example spectral envelope after compensating for the acoustics of the sound delivery system used in the physiological studies.

[Figure 1 about here.]

B. Model of the auditory-periphery

The auditory-periphery model used to simulate the responses to the vowel has been developed by Zilany and Bruce (2006) and is capable of generating realistic response properties of the AN fibers in cats across a wide range of CFs and intensities spanning the dynamic range of hearing. A schematic diagram of the model is given in Fig. 2. Each section of the model provides a phenomenological description of the major functional components of the auditory-periphery, from the middle ear (ME) to the auditory nerve.

[Figure 2 about here.]

The first section models the filtering properties of the ME, which affects the relative levels of the components of wide-band stimuli and hence plays an important role in simulating responses to stimuli such as vowels. The input to the ME is the instantaneous pressure waveform of the stimulus in Pa sampled at 500 kHz. The ME filter is followed by a signal-path C1 filter, which sets up the baseline tuning for the AN fiber. A feed-forward control path regulates the gain and bandwidth of the C1 filter to account for several level-dependent properties in the cochlea. The C1 filter has been designed in such a way that it can address a range of realistic response properties

of the cochlea. The asymmetrical orientation of the poles and zeros of this filter causes inclusion of the instantaneous frequency (IF) glides in the impulse response of the AN, and thus the model is able to simulate the BF shift as a function of sound pressure level (Carney et al., 1999; Tan and Carney, 2003). The tails of the tuning curves, which are important in modeling responses to multi-component stimuli, arise from the careful selection of the poles and zeros of this filter (Zilany and Bruce, 2006). The output of the C1 filter closely resembles the primary active mode of vibration of the basilar membrane (BM) in the cochlea.

A parallel-path C2 filter has been introduced as a second mode of excitation to the IHC and is critical for simulating the transition region effects at high levels. The C2 filter is linear, static, and is the same as the C1 filter with complete OHC impairment, i.e., the tuning of the C2 filter is same as the broadest possible tuning of the C1 filter. For CFs $>$ 600 Hz, this tuning will also be the same as the C1 filter tuning in the normal model of the cochlea at high presentation levels, where the control signal has saturated. However, it is to be noted that for fibers with CFs $<$ 600 Hz, the frequency response of the C1 filter at high levels is dynamic in the normal model of the cochlea, changing cycle-by-cycle, and thus differs somewhat from the C2 filter. The two filters (C1 and C2) are followed by two separate transduction functions, referred to as the C1 and C2 transduction functions. The selection of these functions, consistent with Kiang's two factor cancellation hypothesis (Kiang, 1990), is such that at low-levels the C2 output is significantly smaller than the C1 output, whereas at high levels the C2 output dominates, and the C1 and C2 responses are out of phase. At levels within the transition region, both outputs are approximately equal in magnitude and tend to cancel each other. In addition, the C2 response is not subject to rectification, unlike the C1 response at high levels, and consequently peak splitting results from the C1/C2 interaction (Kiang, 1990; Zilany and Bruce, 2006). Note that, unlike other models having parallel filter paths with static nonlinearities (e.g., Goldstein, 1990, 1995; Lin and Goldstein, 1995; Meddis et al., 2001) where the interaction between filter paths gives rise to the tips and tails of the tuning curves, the tips and tails of this model's tuning curves arise from the C1 filter alone. Only at very

high levels (80 dB SPL or above) does the C2 tuning contribute substantially, as shown in Fig. 5 of Zilany and Bruce (2006). As described above, the C2 filter has approximately the same tuning as the C1 filter at these levels, so the C2 filter has little effect on the shape of the magnitude-frequency response of the model.

The summed output of the two transduction functions is then passed through a seventh-order IHC low-pass filter with a cut-off frequency of 3.8 kHz that describes the fall-off in pure tone synchrony with CF above 1 kHz. The IHC output drives the IHC-AN synapse which provides the instantaneous synaptic release rate as output. The synaptic gain has been adjusted such that the model produces responses of high-spontaneous rate (50 spikes/s before refractoriness), low-threshold fibers. Finally the AN discharge times are produced in the model by a renewal process that includes refractory effects. More details of the model can be found in Zilany and Bruce (2006), and model code is available from the authors on request.

1. Modification of the model

The model developed by Zilany and Bruce (2006) has been validated against a wide range of published physiological data. However, we have found that a slight adjustment in the cochlear amplifier (CA) gain improves the model's prediction of the vowel data of Wong et al. (1998). Their data show that the range of sound levels over which the C1/C2 transition occurs for tones is fairly narrow in the CF range 1–3 kHz. To obtain similar behavior from the model, the CA gains in this CF range have to be increased from the previously published values, whereas the gains at higher CFs need no further modification.

The CA gain (for cats) in Zhang et al. (2001) was chosen to vary as a function of CF in a fashion similar to the direct measurement of CA gain in guinea pig and chinchilla from BM data (Nuttall and Dolan, 1996; Cooper and Rhode, 1997; Ruggero et al., 1997), which increases from 15 dB at low CFs to 70 dB at high CF fibers (shown by the dashed line in Fig. 3). Bruce et al. (2003) modified this function to better explain the change in Q_{10} values in impaired fibers in cats

and the vowel responses in cats at high presentation levels. This was done by applying more CA gain in the CF range 500 Hz to 5 kHz, and also reducing the gain to 52 dB at high CFs (Appendix B of Bruce et al., 2003, and shown by the dotted line in Fig. 3). With further increase in CA gain at lower CFs in this paper, this function agrees more with the estimates of CA gain versus frequency in humans, which is thought to be almost flat across frequency (Oxenham and Plack, 1997; Plack and Oxenham, 2000; Lopez-Poveda et al., 2003). The modified CA gain function is shown by the solid line in Fig. 3, and is given by

$$\text{gain}_{\text{CA}}(\text{CF}) = \max\{15, 52(\tanh(2.2\log_{10}(\text{CF}/600) + 0.15) + 1.0)/2\}, \quad (1)$$

where CF has the units of Hz. However, it is to be noted that this modification does not substantially affect the basic response properties of the model AN fibers reported in Zilany and Bruce (2006).

[Figure 3 about here.]

2. Modeling a population of impaired fibers

To predict vowel responses for the population of AN fibers reported in Miller et al. (1997), the levels of OHC and IHC impairment, C_{OHC} and C_{IHC}, respectively, have been estimated as a function of frequency (normal or impaired CF[1]) to fit the impaired Q_{10}, impaired CF, and minimum threshold shift of the impaired pool of data. It is known that damage to the OHCs causes both broadening of the tuning and elevation of the threshold along with the shift in CF, whereas impairment in the IHCs primarily produces elevation of the tuning curve without substantially broadening (Liberman and Dodds, 1984). Following the method of Bruce et al. (2003), values of C_{OHC} have been chosen as a function of CF such that the model impaired fibers show approximately the same impaired Q_{10} and impaired CF as those of the cat data. Then an appropriate amount of impairment of the IHCs (C_{IHC}) has been applied to obtain the remaining threshold shift observed in the physiological data.

[Figure 4 about here.]

Figure 4(A) shows the impairment in the OHC as a function of CF that gives a close fit to the impaired Q_{10} and impaired CF of the physiological data. Normalized Q_{10} values of the experimental data and model are provided in Fig. 4(B) as a function of impaired CF. The threshold shift relative to the best threshold curve (BTC) produced due to the impairment in the OHC alone is shown by the dashed line in Fig. 4(D). The remaining threshold shift is accounted for by impairment of the IHCs.

Figure 4(C) shows the functional relationship between C_{IHC} and impaired CF to obtain the remaining threshold shift that is not accounted for by the impairment of the OHCs. The combined threshold shift due to impairment of both OHCs and IHCs is shown by the solid line in Fig. 4(D). The IHC impairment is applied such that the model impaired fibers match the minimum threshold shift of the impaired pool of data. Even severe impairment of the OHCs in the model can only account for around two-thirds of the threshold shift seen in the physiological data, and thus substantial damage to the IHCs is also necessary to fully model the impaired data (Miller et al., 1997), consistent with histological assessment of OHC and IHC damage in cats with similar acoustic trauma (Liberman and Dodds, 1984). Note that the C2 responses do not contribute to the overall responses until around 80 dB SPL or higher and are trauma insensitive. Consequently, the C2 filter does not contribute to the threshold tuning of normal or impaired fibers except in the case of very severe impairment of C1 IHC transduction, i.e., for very small values of C_{IHC}.

III. RESULTS

Here we provide a detailed comparison of the model predictions with the results of the cat physiological experiments (Miller et al., 1997; Wong et al., 1998) in response to the vowel /ɛ/.

A. Predictions of single normal fiber data at low to moderate levels and impaired fiber data at high levels

Miller et al. (1997) recorded the responses of AN fibers to a synthesized stimulus "besh" and evaluated the synchronized responses in the steady-state vowel portion /ɛ/. The synchronized rates of AN fibers are determined by taking the Fourier transform of the poststimulus time histogram (PSTH) normalized to units of spikes/s (Sec. I.D. of Miller et al., 1997).

Shown in Fig. 5 are the synchronized responses of two AN fibers (normal and impaired) with CFs near F2 for three different sound pressure levels; the left column shows the measured responses from cats (Miller et al., 1997) and the right column shows the corresponding model predictions. Normal fiber responses shown in Fig. 5(A) predominantly synchronize to the formant F2, although there is some synchrony to the second harmonic of F2, and also to the adjacent harmonics around F2 at the lowest presentation level. In contrast, for the impaired fiber [(Fig. 5(C)] the responses become broadband, and a substantial amount of power is shifted to F1 as the sound level is increased. Note that higher presentation levels are used for the impaired fibers because of the threshold shift with impairment.

[Figure 5 about here.]

The synchronized responses of a model fiber with normal OHC and IHC function are shown in Fig. 5(B). For comparison, the CF, threshold and Q_{10} are matched to those of the example fiber from Fig. 5(A). The spontaneous rate of the model fiber after the effects of refractoriness is similar to the 33 spikes/s of the example normal fiber. The normal model fiber shows the following response properties. (i) At the lowest presentation level (31 dB SPL), the fiber shows phase locking to a number of harmonics of the vowel around F2, although the peak response is at F2. Synchrony to adjacent harmonics indicates that the fiber is operating in the linear region, i.e., compression/suppression is not in effect at this level. However, the model responses at adjacent harmonics around F2 are less prominent than in the physiological data. This might be due to the

model fiber having a lower compression/suppression threshold, such that it starts to exhibit some synchrony capture by F2 at this level. (ii) There are substantial responses present even at higher frequencies near the second harmonic of the F2, which are due to the rectification of the instantaneous discharge rate in the synapse. Again, the model responses are weaker than the measured data. (iii) As the sound level increases, the phase locking to adjacent harmonics around F2 almost disappears, the responses being dominated by F2, which is referred to as synchrony capture. Compressive and suppressive nonlinearity of the BM are the main cause for synchrony capture by F2 at these levels (Bruce et al., 2003). (iv) Synchronized rates do not increase linearly with the presentation levels, partly because of the compressive nonlinearity of the BM responses at these levels and partly because of saturation in the synapse.

Shown in Fig. 5(D) are the synchronized responses of a model impaired fiber for three different sound pressure levels. No impairment in the OHC ($C_{OHC} = 1.0$), but with IHC impairment ($C_{IHC} = 0.035$) to a normal model fiber with CF of 1.6 kHz and a Q_{10} of 3.65, produces the tuning properties with an impaired CF at 1.6 kHz, $Q_{10} \approx 3.65$, and a threshold shift of ~ 55 dB, similar to those from the example impaired fiber in Fig. 5(C). This example fiber has a slightly higher spontaneous discharge rate (44 spikes/s) than the model fiber. The model responses are broadband, meaning that significant synchrony is present at a large number of harmonics of the vowel. In particular, two prominent response components, one at F1 and the other at F2$-$F1, emerge at higher presentation levels, neither of which is present in the normal fiber's response, and significant responses are also present at other frequencies. However, the model responses show stronger synchrony to F2, whereas the physiological data for this impaired fiber shows moderate synchrony to F2 at all levels studied. It is also to be noted that at 93 dB SPL, the model prediction shows significant synchrony to a large number of harmonics other than formant related harmonics, which are not seen in the measured data. This smearing of the synchrony in the model prediction arises from the C1/C2 interaction, as the C1 and C2 responses are not completely out of phase at this presentation level (the C1/C2 transition occurs above 100 dB SPL for this fiber). Note that the output of the C1 and

C2 filters become identical only when the control path output saturates at very high levels (Zilany and Bruce, 2006).

Figure 6 shows the synchronized responses for two fibers (normal and impaired) with CFs close to the third formant (F3); both measured (Miller et al., 1997) and model responses are shown side-by-side for comparison. Here the normal fiber responses shown in Fig. 6(A) are synchronized to F3 and its adjacent harmonics. However, the synchrony capture is weaker than what was observed in the previous case [Fig. 5(A)]. The impaired fiber [Fig. 6(C)] again shows a broadband response with little synchrony to F3 and substantial responses at F1 and F2.

[Figure 6 about here.]

Shown in Fig. 6(B) are the synchronized responses for a model fiber with normal OHC and IHC function and CF, threshold and Q_{10} approximately matching the normal fiber from Fig. 6(A). Again, this example fiber has a slightly higher spontaneous discharge rate (48 spikes/s) than the model fiber. At all levels shown, the model fiber is phase locked to F3 and adjacent harmonics. As the presentation level increases, synchrony to F3 increases although some synchrony to adjacent harmonics still exists. At the highest presentation level, a small amount of synchrony to F1 and F2 emerges that is not present in the measured data, which may result from the tuning of the model fiber at this stimulus level being broader than that of the particular cat fiber.

Appropriate impairment in the OHC ($C_{OHC} = 0.18$) and IHC ($C_{IHC} = 0.035$) to a normal model fiber with CF of 2.95 kHz and a Q_{10} of 4.87, produces the tuning properties with impaired CF of 2.6 kHz, $Q_{10} \approx 1.4$, and a threshold shift of ~ 60 dB, similar to those of the example impaired fiber from Fig. 6(C). This example fiber has a somewhat higher spontaneous discharge rate (66 spikes/s) than the model fiber. The responses of this model impaired fiber for three sound levels are shown in Fig. 6(D). Phase locking to F3 is mostly lost, while synchrony to other formants is seen at all presentation levels. At the two higher presentation levels, synchrony to F1 and F2 becomes more prominent. Like the example impaired fiber, some of the significant responses arise at frequencies

that cannot be explained by the distortion products of the prominent formant frequencies.

The predictions of normal AN fiber responses to the vowel shown in this paper utilizing the model of Zilany and Bruce (2006) are quite similar to those of Bruce et al. (2003). However, the impaired fiber responses at high levels are better described by the newer model. Using Bruce et al. (2003), the synchrony to the formants in impaired fibers increases in almost all cases with presentation level [Fig. 9 and 12 of Bruce et al. (2003)], whereas the physiological data and also the model responses using Zilany and Bruce (2006) show somewhat different trends at higher levels. This improvement with the newer model comes from the interaction between the C1 and C2 responses at high levels.

B. Predictions of single normal fiber data at high levels

Wong et al. (1998) studied the temporal responses of normal fibers for the vowel /ɛ/ presented at high sound pressure levels. The vowel was resampled so that the second formant fell exactly at each fiber's CF. Figure 7 shows the spectra of responses to a vowel for a fiber with CF at 2.0 kHz. Both measured (Fig. 2 of Wong et al., 1998) and model responses are shown for four different sound pressure levels. The synchronized rates to the first 20 harmonics of the vowel are shown by solid circles in each case.

At a vowel level of 68 dB SPL, both measured and model responses are synchronized to F2, consistent with the observation that the response of a fiber with a CF near a vowel's formant frequency is tightly phase-locked to that formant frequency. A weak response to F1 emerges at this level as the tuning of the fiber becomes broader. However, it is to be noted that the F1 response for the vowel is much weaker than that of the tonal (F1) response (Fig. 1 of Wong et al., 1998), which is due to the suppression of F1 by F2 in the vowel response (Wong et al., 1998).

[Figure 7 about here.]

As the level increases to 82 dB SPL, synchrony to F2 decreases slightly in the model responses,

whereas responses to F1 and formant related harmonics (such as F2$-2\times$F1, $2\times$F1, and F2$-$F1) increase. In fact, at this level the large components of the response are at F1, F2 and its distortion products, and hence are called multi-formant response (Wong et al., 1998). This indicates that the cochlea still retains some frequency selectivity. Another important observation is that the growth of response to F1 occurs at the expense of F2 before the C1/C2 transition, i.e., the growth of the response to F1 is not related to the C1/C2 interaction. So, it can be concluded that the multi-formant response is not related in any way to the C1/C2 transition (Wong et al., 1998). This is further supported by the model predictions to the vowel observed in Bruce et al. (2003). The model of Bruce and colleagues does not include a C2 filter and consequently is capable of generating only C1 responses. The vowel responses for the Bruce et al. (2003) model also show the fall of synchrony to F2 at high levels (although only for higher CF fibers) while the synchrony to F1 increases.

At 102 dB SPL in the measured data (or 104 dB SPL in model responses), neither F1 nor F2 is dominant, and the responses become broadband, meaning that responses to the harmonics that are not related to the formants are substantial at this level. All components of the vowel for the model fiber undergo the C1/C2 transition at 107 dB SPL. Note that the C1 responses at this level are dominated by the formants and related harmonics, with the nonformant components being suppressed by the saturating nonlinearity of the C1 transduction function. In contrast, the C2 responses at this presentation level include relatively larger nonformant components than the C1 responses, because the C2 transduction function is not saturating at this level. Thus, the cancellation of the formant components in the anti-phase C1 and C2 responses allows the non-formant components in the C2 response to dominate. At very high levels above the C2 threshold (115 dB SPL in model responses or 110 dB SPL in measured data), the synchrony to F2 is completely lost and goes to F1, although significant responses are seen in frequency components which are not related to the distortion products from F1 and F2, i.e., the fiber is responding in a broadband fashion.

It is to be noted that the two higher levels (104 and 115 dB SPL) used to determine model

responses are slightly higher than the levels (102 and 110 dB SPL) of the corresponding physiological data. Since the model fiber in this case undergoes the C1/C2 transition at a slightly higher level (107 dB SPL) than the example fiber from the physiological data [102 dB SPL; *cf.* Fig. 12(C,F)], the two relatively higher levels have been used in determining model responses to illustrate the qualitative match between the model responses and the physiological data.

C. Power ratios versus level

In order to quantify the degree of synchrony capture at different presentation levels, Wong et al. (1998) computed power ratios (PRs), which separate the responses into components related to the formants (including distortion products) and other components. Total power is defined as the sum of powers of the synchronized rates over the first 20 harmonics of the vowel response. The F2 PR is the fraction of total power of the synchronized rate that is phase-locked to the second formant only. The F1 & F2-related PR is the fraction of total power of the synchronized rate that is related to the F1 (5^{th} harmonic), F2 (17^{th} harmonic) and their distortion products (7^{th}, 10^{th}, 12^{th}, 15^{th}, and 20^{th} harmonics).

Shown in Fig. 8 are the power ratios of four normal fibers with CFs in the range 1–3 kHz as a function of sound pressure level. The left and right columns show the measured and model responses, respectively. Normal model fiber responses are simulated with $C_{OHC} = 1$, and $C_{IHC} = 1$. The 400-ms synthetic vowel is sampled such that in each case the second formant falls exactly at the fiber's CF. The lower and upper bounds of the shaded region represent the levels at which synchrony capture by F2 is lost (i.e., vector strength[2], VS$\leq$ 0.5) and the C2 threshold for F1, respectively. The C2 threshold is defined as the stimulus level at which the phase shift just exceeds 90°.

[Figure 8 about here.]

The measured data and model responses exhibit three distinct response areas. Before the

shaded region (at low and moderate levels), responses are phase-locked to F2 (VS$\geq$ 0.5), whereas in the shaded region, synchrony capture by F2 is lost and a multi-formant response to F1 and F2 arises (see Fig. 7). After the shaded region (i.e., above the C2 threshold), the synchrony to F1 becomes dominant, and sometimes the unselective broadband response of the fiber is seen at these levels as judged by the drop in the F1 & F2-related PR.

Both the measured data and model responses show that synchrony lost by F2 at high level is in no way related to the C1/C2 transition as the synchrony to F2 becomes progressively weaker before the C1/C2 transition (the upper bound of the shaded region). So, the C1 response itself shows gradual loss of synchrony by F2 at these levels while at the same time the response to F1 emerges for two reasons: (i) broadening of the C1 filter, and (ii) the shift in BF of the C1 filter. Thus, in the shaded region, the PR of F2 drops, whereas the F1 & F2-related PR stays at the same level meaning that the F2 response is replaced by a response to F1. However, the fall of F1 & F2-related PR at or above C2 threshold seen in the model responses (also shown in Fig. 9 of Wong et al., 1998) reflects the cancellation of the dominant formant components in the C1 and C2 responses, as the outputs of the C1 and C2 filters become comparable around this C1/C2 transition level. At very high levels, the dominant formant response is seen at F1 only (see Fig. 7), and thus F1 & F2-related PR returns to a relatively high value.

With increasing CF, the loss of synchrony to F2 occurs at progressively lower stimulus levels and thus increases the width of the shaded region. It was argued in Wong et al. (1998) that the lower level switch of synchrony from F2 to F1 is due to the stronger two-tone suppression with increasing CF. This has been realized in the model as the CA gain is higher at higher CFs, and thus higher CF tuning becomes relatively broader (i.e., the relative change in Q_{10} is larger) than the tuning of a lower CF fiber at the same level. Consequently, the gain of the C1 filter is reduced, and then the low-pass filtering of the responses above 1 kHz in the IHC section further attenuates F2 more than F1 responses. Another potential reason for this early switch is that the frequency-scaled vowel has more power in F1 for higher CFs due to less attenuation by the ME filter. Bruce et al.

(2003) found that this contributes to the switch of synchrony from F2 to F1 at an earlier level with increasing CF. Note that the levels at which the model predicts loss of synchrony capture by F2 are slightly lower than those of the experimental data for all CFs studied. This might be due to the fact that the model responses can only fit to the lower limits of the vector strength versus CF in the cat data (see Fig. 11 of Zhang et al., 2001).

[Figure 9 about here.]

Figure 9 shows the PRs of four impaired fibers with substantial threshold shifts and broad tuning for sound pressure levels ranging from moderate to very high levels. The left column shows the data obtained experimentally (Wong et al., 1998), and the right column is for model responses in which impaired Q_{10}, impaired CF and threshold shifts for the fibers are matched with those of the physiological data. To do this, appropriate impairment in the OHC (C_{OHC}) and IHC (C_{IHC}) is applied to the model fibers with particular CF and Q_{10} values to obtain similar tuning curves to the experimental fibers. Like Fig. 8, the F2 PR and F1 & F2-related PR as a function of level are shown by dotted and solid lines, respectively.

Both the measured data and the model predictions for impaired fibers show that the synchrony to F2 is mostly lost in all cases studied. The largest component of the response is at F1, which can be inferred from the observation that the F1 & F2-related PR is typically more than twice the corresponding F2-related PR in each case. However, all of these responses are quite broadband as judged by the fall of F1 & F2-related PR from the values near 1.0 observed in normal fibers (compare to Fig. 8).

It is to be noted that the model responses for the three fibers with higher CFs show slightly higher synchrony to F2 than those of measured data for sound levels up to 80 dB SPL, which is caused by the C1 responses. In all of these cases, the synchrony to F2 is near zero at $\sim$ 60 dB SPL because of the loss of sensitivity (threshold shift) of these fibers, and then synchrony to F2 increases with level. Since these fibers are not completely impaired, the corresponding C1 responses

still show some selectivity. However, above 80 dB SPL, the synchrony to F2 for all of these fibers again approaches zero, because at these levels the responses of the fibers are dominated by the broadband C2 response. In the model, the C2 filter has been implemented as the completely OHC-impaired version of the C1 filter, and thus has very broadband response properties. In addition, the C2 responses are not affected by the impairment of OHCs or IHCs. Thus, at very high levels, the synchrony behavior for the formants or harmonics is similar to that of normal fibers. So, the responses are dominated by the formant F1 with other substantial non-formant residual components (see the lowest panel of Fig. 7).

D. Synchrony capture as a function of CF

Miller et al. (1997) measured synchrony capture by the formants (F1, F2 and F3) of the vowel /ɛ/ in "besh" in terms of power ratios for a population of normal and impaired fibers across a wide range of CFs. Here the PR of a particular formant is defined as the fraction of total power (sum of powers of the first 50 harmonics) that is phase-locked to the first, second and third harmonics of that formant, as long as they are less than or equal to 5 kHz. Measured data and model predictions for populations of normal and impaired fibers are shown in the Fig. 10 and 11, respectively. Thick solid lines indicate the model predictions and gray shaded areas show the range of PR values observed in the physiological experiments (Miller et al., 1997).

[Figure 10 about here.]

Model predictions of PRs for normal fibers at 49 and 69 dB SPL shown in Fig. 10 fall predominantly within the range of values observed physiologically. In both the measured data and the model predictions, AN fibers are almost exclusively synchronized to the formant frequency nearest to the fiber's CF, and the peak of the synchrony to a particular formant becomes wider at higher levels which reflects a broader tuning at higher presentation levels.

To predict responses for a population of impaired fibers (Miller et al., 1997), we utilized the functional relationships between C_{OHC} and CF and C_{IHC} and CF described in Section II.B., such that the model population has impaired Q_{10} and threshold shift as a function of impaired CF similar to the experimental data. In order to compensate for the loss of sensitivity, higher presentation levels are used for the impaired fibers. Model predictions of PRs for impaired fibers shown in Fig. 11 also fall reasonably within the range of values seen in the physiological data. Since the acoustic trauma caused impairment predominantly around the F2 and F3 region, the synchrony to F2 and F3 falls drastically from those of the normal values. Higher presentation levels could not restore the normal response properties.

[Figure 11 about here.]

Both measured data and model predictions show that the synchrony to F3 stays near zero for both presentation levels, whereas fibers with CFs at or near F3 show some synchrony to F1 and F2, which is consistent with the results presented in Fig. 6. At 92 dB SPL, model F1 PRs exhibit lower values than the measured data for CFs higher than 800 Hz, because for these fibers the C1/C2 transition occurs at or around this presentation level. So, the response components contain noticeable amount of non-formant residual components, and thus the PR of F1 falls substantially. However, at 112 dB SPL, C2 responses dominate, and again F1 becomes the dominant component in the response. Consequently, the model fibers' synchrony to F1 stays within the range of physiological data in that CF region. At 112 dB SPL, the CF of peak synchrony to F1 shifts upward in the physiological data somewhat more than in the model predictions.

E. C1/C2 transitions in tonal and vowel responses

Wong et al. (1998) showed that the C1/C2 transition occurs even with complex stimuli. In fact, all components of the vowel undergo C1/C2 transition simultaneously, and the transition level is determined by the level at which the largest component reaches its C2 threshold. Note that the C2

threshold (or transition level) is defined as the level at which the phase is shifted by 90°. Figure 12 shows the comparison of the C2 thresholds for the first two formants in vowel responses and for tones at the frequencies of F1 and F2. For comparison, both measured responses from cats and model predictions are shown side by side; the left column shows the measured data and the right column shows the model predictions.

The model responses are shown for 21 fibers, and to be consistent with Wong et al. (1998), the playback sampling rate of the vowel /ɛ/ was changed in each case such that the second formant of the vowel falls exactly at the fiber's CF. The CFs of the model AN fibers range from 1 to 3 kHz (in steps of 100 Hz), which covers the usual range of second formant frequencies encountered in speech. It is worth noting that in determining C2 thresholds for the formants, the phase is only computed when the vector strength is greater than 0.15. It has been observed that when the synchrony is weak, the phase estimates are erratic and hence cannot be computed reliably. To make a comparison with the published data, the model results are analyzed based on the assumption that the levels of F1 and F2 are always 2 and 18 dB, respectively, below the overall vowel level[3].

[Figure 12 about here.]

Figures 12(A) and 12(D) show the comparison of the C2 thresholds (in dB SPL) for the tone at F1 (along the abscissa) and for the formant F1 in vowel responses (along the ordinate). The model predictions, consistent with the measured data, fall within a band of ± 5 dB SPL of the equality line. A similar comparison is shown for F2 in Figs. 12(B) and 12(E). Here the F2 level in the vowel is 5–20 dB lower than the level of the F2 tone to reach the C2 threshold. This early transition is attributed to the presence of other harmonics in the vowel. However, the vowel level when F2 reaches its C2 threshold is within ± 5 dB of the vowel level when F1 reaches its C2 threshold, as shown in Figs. 12(C) and 12(F). Since the level of F1 in the vowel is almost same as the level of the F1 tone to reach C2 threshold, and for F2 the formant reaches its C2 threshold at a level substantially lower than the level required for the corresponding tone, it can be concluded

that it is the largest component in the vowel (in this case F1) that determines the level for C2 threshold for all components of the vowel. Both physiological data and model predictions rule out the possibility of separate processing of each component of the vowel, otherwise individual component might undergo C1/C2 transition separately. So, the overall vowel level determines the transition behavior for each component of the vowel.

This behavior can be understood qualitatively from the model architecture. The design of the model is such that the output of the C1 and C2 filter becomes equal when the control path output saturates, and the two separate transduction functions following the C1 and C2 filters make them completely out of phase. In response to a single tone, a very high sound pressure level is required to generate a large enough control signal to reduce the gain of the C1 filter to that of the C2 filter. In contrast, for a wideband stimulus such as a vowel, the control signal depends on the overall level of the stimulus, because the bandwidth of the control-path filter is wider than that of the signal-path C1 filter. Consequently, the level at which the outputs of the C1 and C2 filter become equal is determined by the overall level of a multi-component stimulus. The overall level of a harmonic complex such as a vowel is determined primarily by the largest harmonic component. Following the IHC transduction functions, the C1 and C2 responses are out of phase for all components of the stimuli, and the steep growth of C2 transduction allows each component of the stimuli to undergo the C1/C2 transition at the same high presentation level, which is determined by the largest harmonic component, as observed above.

IV. DISCUSSION

The normal and impaired AN fiber responses in cats to the vowel /ɛ/ are described in this paper using the computational model proposed by Zilany and Bruce (2006). The vowel presented to the tympanic membrane during physiological experiments is replicated here, and then applied as the input to the model. To do this, the stimulus is first filtered by the HRTF of the human head and then

modified by the frequency response of the acoustic delivery system. This is necessary, because the change in the relative levels of the stimulus components (especially the formants) affects the AN responses substantially, and part of the improvements in predictions over the Bruce et al. (2003) results from this. The model predictions are compared to the responses from the physiological experiments (Miller et al., 1997; Wong et al., 1998), which clearly demonstrates the ability of the model to closely represent the measured data in both normal and impaired ears. The measured data and model predictions include responses of individual fibers as well as a population of fibers over a wide range of CFs.

A. C1/C2 transitions for the vowel

The observation that all components of the vowel undergo the C1/C2 transition simultaneously (Wong et al., 1998) has been accurately modeled in this paper, and the transition level is determined by the level of the largest component of the vowel. To provide further evidence, we have determined the transition levels in the model AN responses to a contrast enhanced frequency-shaped (CEFS) version of the vowel /ɛ/. The CEFS vowel a spectrum (i.e., the relative magnitude of the formants and harmonics) that is substantially different from that of the standard /ɛ/, such that it is dominated by the second formant (Miller et al., 1999a). The line spectrum of the CEFS vowel is shown in the upper panel of Fig. 13. The playback sampling rate of this vowel is changed so that F2 always falls at the fiber's CF. Again, model AN fibers with CFs ranging from 1 to 3 kHz with steps of 100 Hz have been considered. The lower panel of Fig. 13 shows the comparison of the C2 thresholds (in dB SPL) in model AN fiber responses for a tone at the F2 frequency (along the abscissa) and for the formant F2 in the CEFS vowel (along the ordinate). The result falls within a band of ± 5 dB SPL from the equality line, indicating that it is the largest component F2 in the CEFS vowel that determines the C1/C2 transition for all components of the vowel [*cf.* Fig. 12(E)]. It is to be noted that the synchrony to F1 for this vowel is not enough to produce a steady phase

response before the C1/C2 transition, and that is why the C1/C2 transition for F1 is not shown here. However, in the C2 responses (i.e., at very high levels), the synchrony to F1 is substantial and thus gives a steady phase response.

[Figure 13 about here.]

B. Comparison with previous models

A number of earlier computational models attempted to predict the AN fiber responses to speech-like sounds (e.g., Deng and Geisler, 1987a; Geisler, 1989; Jenison et al., 1991; Sachs et al., 2002). However, most of them lacked several important aspects of nonlinearity observed in the cochlea that are essential in predicting responses to speech. For example, Geisler (1989) studied the effects of broadened tuning on responses to speech using a computational model that did not include two-tone suppression and BM compression. Recently, Bruce et al. (2003) and Holmes et al. (2004) attempted a more systematic investigation of AN responses to vowels in cats and guinea pigs, respectively. Both studies employed more complete AN models, in that they are able to produce a wide range of physiological phenomena in response to simple and complex stimuli.

The predictions of AN fiber responses using the model by Zilany and Bruce (2006) are almost same as those of Bruce et al. (2003) when operating at low to moderate sound pressure levels. However, the input to the model is modified here by the frequency response of the sound delivery system to closely match the vowel's spectral shape at the tympanic membrane, as was presented during experiments. The change in relative levels of the formants explains the onset of loss of synchrony capture by F2 at moderate levels for CFs above 2 kHz (see Fig. 8), which was not observed by Bruce et al. (2003). At high levels, the predictions presented in this paper are superior to those of Bruce et al. (2003) due to several reasons. First, the C1 filter, originally a gammatone filter in Bruce et al. (2003), has been replaced by a chirp filter in Zilany and Bruce (2006), and thus the impulse response of the model AN fiber shows the instantaneous frequency glide in it, which

is responsible for producing the shift in BF at high levels in normal and impaired fibers. The loss of synchrony capture in impaired fibers as well as in normal fibers at high levels is partly attributed to this. Second, the inclusion of the C2 filter in Zilany and Bruce (2006) enables the model to address the AN responses seen at the C1/C2 transition and above. For instance, the significant loss of synchrony capture by the dominant formant components observed around the C1/C2 transition level cannot be described without the C1/C2 interaction. Third, the CA gains for fibers with lower CFs in Bruce et al. (2003) are not sufficient to account for the significant loss of synchrony capture by F2 seen at very high levels in normal and impaired fibers.

Holmes et al. (2004) predicted guinea pig data for single and double vowels (Palmer et al., 1986; Palmer, 1990) using a cochlear model applying the dual resonance nonlinear (DRNL) filter bank approach (Meddis et al., 2001; Sumner et al., 2003). They showed that the BM nonlinearity has a very small effect in the representation of speech at low frequencies, and it is the filter shape not the nonlinearity that is responsible for the vowel formant representation in their model predictions. This observation strongly contrasts with the model responses in cats (Deng and Geisler, 1987a; Sachs et al., 2002; Bruce et al., 2003), where BM nonlinearity plays an important role in the representation of speech over a wide range of CFs, and thus synchrony capture is relatively stronger. Holmes et al. (2004) suggest that the differences in the guinea pig and cat model predictions may arise from a lower degree of BM nonlinearity in guinea pigs at lower CFs. Indeed, compressive nonlinearity appears to be absent from BM vibration recordings made in the apex of the guinea pig cochlea (Cooper and Rhode, 1995; Khanna and Hao, 1999). Unfortunately, no similar BM data are available in cat, so we must infer the degree of BM nonlinearity from the cat AN data. In this work, we have further increased nonlinearity (CA gain) at lower CFs than was employed in Bruce et al. (2003)—see Fig. 3. Human psychoacoustic data shows substantial nonlinearity in low frequency hearing (Oxenham and Plack, 1997; Plack and Oxenham, 2000; Lopez-Poveda et al., 2003). If our estimates of cat BM nonlinearity are at least qualitatively correct, then this may suggest that the cat is a more suitable animal model than the guinea pig for

predicting neural responses to speech in humans.

The main focus of this paper is to predict AN responses to vowels at high levels (Miller et al., 1997; Wong et al., 1998), especially the spread of synchrony and the C1/C2 transition. Lin and Goldstein (1995) incorporated the two-factor hypothesis in their model to explain high-level nonlinearities such as the C1/C2 transition and peak splitting, and conducted a detail analysis of responses to tones. As their model incorporates only static nonlinearities, Lin and Goldstein (1995) found it necessary to introduce an *ad hoc* stimulus-dependent phase adaptation mechanism to produce the correct stimulus-dependent phase changes as a function of time. In addition, the C1/C2 transition occurs in the model BM responses, inconsistent with most of the physiological observations (Ruggero et al., 1997; Cheatham and Dallos, 1998; Robles and Ruggero, 2001). It is unlikely that this model can predict vowel data at high levels, especially the C1/C2 transition. Holmes et al. (2004) predicted guinea pig data to vowels at low to moderate levels, and in fact, only cat data is available at high levels. However, the model used in Holmes et al. (2004) is of the type with parallel filter paths with static nonlinearities like that of Lin and Goldstein (1995), and hence it is not clear if this form of model could predict the vowel C1/C2 transition behavior observed in cats.

C. Implications for speech intelligibility and hearing aids

Although the predictions of normal AN fiber responses at low to moderate levels shown in this paper utilizing the model of Zilany and Bruce (2006) are quite similar to those of Bruce et al. (2003), the normal and impaired fiber responses at high levels are better described by the newer model. It is of interest that the method of implementing impairment in this paper was the same as that of Bruce et al. (2003). Consequently, the improved predictions of the impaired AN fiber data arise solely as a result of the improved description of the normal AN fiber behavior that was obtained with the new C1 filter properties and the addition of the C2 filter.

In predicting impaired fiber responses to the vowel using Bruce et al. (2003), the synchrony

capture by F2 remains substantial even at very high levels (Figs. 14 and 16 of Bruce et al. (2003)). In contrast, measured data and predictions using the new model (Figs. 9 and 11 of this paper) show significant loss of synchrony to F2 at high levels. The new model also now better predicts the loss of synchrony to F2 in normal fibers at high levels—compare Fig. 13 of Bruce et al. (2003) with Fig. 8 of this paper. However, in both the measured data and the new model predictions, the degradation of the representation of F2 at high presentation levels is more severe in the impaired fibers. Wong et al. (1998) suggested that these results could at least partly explain the roll-off observed in speech intelligibility at high levels that is stronger in hearing impaired listeners than normal hearing listeners, even when both are operating at the same presentation level (Stelmachowicz et al., 1985; Glasberg and Moore, 1986). We have tested this hypothesis with a speech-intelligibility predictor utilizing the Zilany and Bruce (2006) model and have found that the model can indeed explain the effects of presentation level and hearing impairment on intelligibility (Zilany and Bruce, 2007). Consequently, this model appears to be a useful tool for the design and evaluation of hearing aid amplification schemes (e.g., Miller et al., 1999a; Bruce, 2004; Bondy et al., 2004).

D. Future directions

The results of this study show good qualitative and quantitative match of the model predictions with the vowel data in cats. The model responses to other simple and complex stimuli are described in Zilany and Bruce (2006). Now the model appears accurate enough to be useful as a front end in many research areas, such as speech recognition in noisy conditions (e.g., Ghitza, 1988; Tchorz and Kollmeier, 1999), modeling of auditory scene analysis (e.g., Brown and Cooke, 1994), modeling of neural circuits in the auditory brain-stem (e.g., Hewitt and Meddis, 1992), design of speech processors in cochlear implants (e.g., Wilson et al., 2005), and design of hearing aid amplification schemes (e.g., Miller et al., 1999a; Bruce, 2004; Bondy et al., 2004).

Although the model is primarily designed to replicate AN fiber data, it does give some insight

about the possible physical mechanisms producing the C1/C2 interaction in the cochlea. The model architecture has been designed such that two modes of BM vibration, active and passive, are transduced by two separate IHC transduction functions. This separate transduction mechanism seems consistent with two important observations: (i) C1 responses are highly affected by the impairment in the cochlea, while C2 is resistant to trauma, and (ii) all components of a multi-component stimuli undergo C1/C2 transition simultaneously irrespective of their individual levels. However, further experimental and modeling studies of cochlear micromechanics and hair cell transduction are clearly required in order to better determine the details of the physical basis for the C2 response.

In addition, our model provides an explanation of why all the formants of a vowel undergo the C1/C2 transition together. Consistent with the hypothesis of Wong et al. (1998), we found that the model's C1/C2 transition for the vowel formants is dependent on the overall level of the vowel, which is dominated by the formant with the greatest amplitude. We tested this hypothesis further using the CEFS vowel, a variant on the standard /ɛ/ vowel. However, in physiological experiments the CEFS vowel has only been presented to cats with impaired ears (Miller et al., 1999a). Further experiments in normal-hearing cats with a range of different vowels or multi-tone complexes would be beneficial in further understanding the C1/C2 transition behavior for such stimuli.

ACKNOWLEDGMENTS

The authors would like to thank Eric Young and Roger Miller for supplying acoustic calibration and auditory nerve data from the physiological studies and Laurel Carney, Michael Heinz, Eric Young and three anonymous reviewers for giving invaluable feedback on earlier versions of the manuscript. This research was supported by NSERC Discovery Grant 261736 and the Barber-Gennum Chair Endowment.

NOTES

[1]In this paper, best frequency (BF) is the frequency at which the fiber response is maximum, which can vary with sound level, whereas the characteristic frequency (CF) is the stimulus frequency at which the fiber has the lowest threshold. Note that the model parameter CF is invariant to the stimulus sound pressure level and hair cell impairment. However, the value for CF that is measured from a model tuning curve may change as a function of hair cell impairment (see Fig. 5 of Zilany and Bruce, 2006), although the model parameter CF has not changed. In such cases, we refer to the CF measured from the tuning curve as the *impaired CF*. Note that in this paper the CFs (normal or impaired) are estimated from the corresponding tuning curves following the method described in Liberman (1984).

[2]Vector strength is a dimensionless measure of phase locking, and is defined, for a particular frequency, as the ratio of the synchronized rate at that frequency and the average discharge rate of the fiber; $\mathrm{VS}(kf_0) = |R(kf_0)|/R(0)$, where $|R(kf_0)|$ is the magnitude of the Fourier transform of the period histogram at frequency kf_0, the k^{th} harmonic of the stimulus fundamental f_0 (Wong et al., 1998).

[3] Wong et al. (1998) erroneously stated that the levels of F1 and F2 in the vowel driving the headphone were always 2 and 18 dB, respectively, below the overall vowel level. However, this actually corresponds to the average levels of the first two formants of the vowel *after* being modified by the acoustic delivery system (Miller and Young, private communication). In fact, the acoustic delivery system changes the levels of the various components of the vowel to a extent of $\pm$6 dB for frequencies below 3 kHz (Wong et al., 1998). So, the relative levels of the various vowel components, including F1 and F2, change as the vowel is resampled. Thus, the analysis should be based on the exact levels of F1 and F2 in the vowel at the tympanic membrane. However, when

the model responses were analyzed with the exact levels of F1 and F2 in the vowel at the tympanic membrane (input to the model), the results were qualitatively the same as when using the average levels of F1 and F2 in the vowel used in Wong et al. (1998). Consequently, to avoid reanalyzing the data, we have used the same analysis (average rather than exact levels) as Wong et al. (1998).

REFERENCES

Bondy, J., Becker, S., Bruce, I., Trainor, L., and Haykin, S. (2004). "A novel signal-processing strategy for hearing-aid design: Neurocompensation," Signal Processing **84**, 1239–1253.

Brown, G. J. and Cooke, M. (1994). "Computational auditory scene analysis," Comput. Speech Lang. **8**, 297–336.

Bruce, I. C. (2004). "Physiological assessment of contrast-enhancing frequency shaping and multi-band compression in hearing aids," Physiol. Meas. **25**, 945–956.

Bruce, I. C., Sachs, M. B., and Young, E. D. (2003). "An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses," J. Acoust. Soc. Am. **113**, 369–388.

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.

Cheatham, M. A. and Dallos, P. (1998). "The level dependence of response phase: Observations from cochlear hair cells," J. Acoust. Soc. Am. **104**, 356–369.

Cooper, N. P. and Rhode, W. S. (1995). "Nonlinear mechanics at the apex of the guinea-pig cochlea," Hear. Res. **82**, 225–243.

Cooper, N. P. and Rhode, W. S. (1997). "Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea," J. Neurophysiol. **78**, 261–270.

Delgutte, B. (1980). "Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers," J. Acoust. Soc. Am. **68**, 843–857.

Delgutte, B. and Kiang, N. Y. S. (1984). "Speech coding in the auditory nerve: V. Vowels in background noise," J. Acoust. Soc. Am. **75**, 908–918.

Deng, L. and Geisler, C. D. (1987a). "A composite auditory model for processing speech sounds," J. Acoust. Soc. Am. **82**, 2001–2012.

Deng, L. and Geisler, C. D. (1987b). "Responses of auditory-nerve fibers to nasal consonant-vowel syllables," J. Acoust. Soc. Am. **82**, 1977–1988.

Geisler, C. D. (1989). "The responses of models of "high-spontaneous" auditory-nerve fibers in a damaged cochlea to speech syllables in noise," J. Acoust. Soc. Am. **86**, 2192–2205.

Ghitza, O. (1988). "Temporal non-place information in the auditory-nerve firing patterns as a front-end for speech recognition in a noisy environment," J. Phonetics **16**, 109–123.

Glasberg, B. R. and Moore, B. C. J. (1986). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," J. Acoust. Soc. Am. **79**, 1020–1033.

Goldstein, J. L. (1990). "Modeling rapid waveform compression on the basilar membrane as multiple- bandpass-nonlinearity filtering," Hear. Res. **49**, 39–60.

Goldstein, J. L. (1995). "Relations among compression, suppression, and combination tones in mechanical responses of the basilar membrane: Data and MBPNL model," Hear. Res. **89**, 52–68.

Hewitt, M. J. and Meddis, R. (1992). "Regularity of cochlear neucleus stellate cells: A computational modeling study," J. Acoust. Soc. Am. **93**, 3390–3399.

Holmes, S. D., Sumner, C. J., O'Mard, L. P., and Meddis, R. (2004). "The temporal representation of speech in a nonlinear model of the guinea pig cochlea," J. Acoust. Soc. Am. **116**, 3534–3545.

Jenison, R. L., Greenberg, S., Kluender, K. R., and Rhode, W. S. (1991). "A composite model of the auditory periphery for the processing of speech based on the filter response functions of single auditory-nerve fibers," J. Acoust. Soc. Am. **90**, 773–786.

Khanna, S. M. and Hao, L. F. (1999). "Reticular lamina vibrations in the apical turn of a living guinea pig cochlea," Hear. Res. **132**, 15–33.

Kiang, N. Y.-S. (1990). "Curious oddments of auditory-nerve studies," Hear. Res. **49**, 1–16.

Liberman, M. C. (1984). "Single-neuron labeling and chronic cochlear pathology. I. Threshold shift and characteristic-frequency shift," Hear. Res. **16**, 33–41.

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

Liberman, M. C. and Kiang, N. Y.-S. (1984). "Single-neuron labeling and chronic cochlear pathology. IV. Stereocilia damage and alterations in rate- and phase-level functions," Hear. Res. **16**, 75–90.

Lin, T. and Goldstein, J. L. (1995). "Quantifying 2-factor phase relations in non-linear responses from low characteristic-frequency auditory-nerve fibers," Hear. Res. **90**, 126–138.

Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (2003). "Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing," J. Acoust. Soc. Am. **113**, 951–960.

Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (2001). "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**, 2852–2861.

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

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

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

Nuttall, A. L. and Dolan, D. F. (1996). "Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig," J. Acoust. Soc. Am. **99**, 1556–1564.

Oxenham, A. J. and Plack, C. J. (1997). "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **101**, 3666–3675.

Palmer, A. R. (1990). "The representation of the spectra and fundamental frequencies of steady-state single- and double-vowel sounds in the temporal discharge patterns of guinea pig cochlear-nerve fibers," J. Acoust. Soc. Am. **88**, 1412–26.

Palmer, A. R., Winter, I. M., and Darwin, C. J. (1986). "The representation of the steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons," J. Acoust. Soc. Am. **79**, 100–113.

Plack, C. J. and Oxenham, A. J. (2000). "Basilar-membrane nonlinearity estimated by pulsation threshold," J. Acoust. Soc. Am. **107**, 501–507.

Robles, L. and Ruggero, M. A. (2001). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (1997). "Basilar-membrane responses to tones at the base of the chinchilla cochlea," J. Acoust. Soc. Am. **101**, 2151–2163.

Sachs, M. B., Bruce, I. C., Miller, R. L., and Young, E. D. (2002). "Biological basis of hearing-aid design," Ann. Biomed. Eng. **30**, 157–168.

Schilling, J. R., Miller, R. L., Sachs, M. B., and Young, E. D. (1998). "Frequency-shaped amplification changes the neural representation of speech with noise-induced hearing loss," Hear. Res. **117**, 57–70.

Sinex, D. and Geisler, C. (1983). "Responses of auditory-nerve fibers to consonant-vowel syllables," J. Acoust. Soc. Am. **73**, 602–615.

Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P., and Mott, J. (1985). "Speech perception ability and psychophysical tuning curves in hearing- impaired listeners," J. Acoust. Soc. Am. **77**, 620–627.

Sumner, C. J., O'Mard, L. P., Lopez-Poveda, E. A., and Meddis, R. (2003). "A non-linear filter-bank model of the guinea-pig cochlear nerve: Rate responses," J. Acoust. Soc. Am. **113**, 3264–3274.

Tan, Q. and Carney, L. H. (2003). "A phenomenological model for the responses of the auditory-nerve fibers. II. Nonlinear tuning with a frequency glide," J. Acoust. Soc. Am. **114**, 2007–2020.

Tchorz, J. and Kollmeier, B. (1999). "A model of auditory perception as a front end for automatic speech recognition," J. Acoust. Soc. Am. **106**, 2040–2050.

Wiener, F. M. and Ross, D. A. (1946). "The pressure distribution in the auditory canal in a progressive sound field," J. Acoust. Soc. Am. **18**, 401–408.

Wilson, B. S., Schatzer, R., Lopez-Poveda, E. A., Sun, X., Lawson, D. T., and Wolford, R. D. (2005). "Two new directions in speech processor design for cochlear implants," Ear & Hearing **26**, 73S–81S.

Wong, J. C., Miller, R. L., Calhoun, B. M., Sachs, M. B., and Young, E. D. (1998). "Effects of high sound levels on responses to the vowel /ε/ in cat auditory nerve," Hear. Res. **123**, 61–77.

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.

Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (2001). "A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression," J. Acoust. Soc. Am. **109**, 648–670.

Zilany, M. S. A. and Bruce, I. C. (2006). "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," J. Acoust. Soc. Am. **120**, 1446–1466.

Zilany, M. S. A. and Bruce, I. C. (2007). "Predictions of speech intelligibility with a model of the normal and impaired auditory-periphery," accepted for publication in *Proceedings of 3rd International IEEE EMBS Conference on Neural Engineeering* (IEEE, Piscataway, NJ).

Figure Captions

1 Line spectrum of the synthesized vowel /ɛ/ after filtering by the HRTF of the human head. The fundamental frequency is 100 Hz, and the first three formants indicated by the vertical dashed lines are at 0.5, 1.7, and 2.5 kHz, respectively. A representative spectral envelope of the vowel, as delivered to the tympanic membrane, is shown by the thin line.

2 Schematic diagram of the auditory-periphery model, reprinted from Zilany and Bruce (2006) with permission from the Acoustical Society of America © (2006). The input to the model is an instantaneous pressure waveform of the stimulus in units of Pa and the output is the spike times in response to that input. The model has a middle-ear filter, a feed-forward control-path, a signal-path C1 filter and a parallel-path C2 filter, an inner hair-cell (IHC) section followed by a synapse model and a discharge generator. Abbreviations: outer hair cell (OHC), low-pass (LP) filter, static nonlinearity (NL), characteristic frequency (CF), inverting nonlinearity (INV). C_{OHC} and C_{IHC} are scaling constants that indicate OHC and IHC status, respectively. The bold and grey lines in the filter functions represent the tunings at low and high sound pressure levels, respectively. The wideband C2 filter shape is fixed and is the same as the broadest possible C1 filter. The bold and grey line in the stage following the C1 filter (C1 transduction function) indicate the nonlinearity in the IHC in normal and impaired (scaled down according to C_{IHC}) conditions, respectively. The stage following the wideband C2 filter, referred to as the C2 transduction function, is an inverting nonlinearity and is insensitive to trauma.

3 Cochlear amplifier (CA) gain as a function of CF. The dashed line shows the gain function used in Zhang et al. (2001), the dotted line shows the gain used in Bruce et al. (2003), and the solid line indicates the gain function used in this paper.

4 Modeling of impaired fibers by applying impairment in the OHC (C_{OHC}) and IHC (C_{IHC}). A: Functional relationship between C_{OHC} and CF (in the model) to fit the impaired Q_{10} and impaired CF of the physiological data. B: Normalized Q_{10} vs. impaired CF. Q_{10} values of the impaired fibers are normalized by their corresponding normal Q_{10} values. Here impaired CF corresponds to the model fiber after applying impairment in the OHC. Gray symbols indicate measured data (Miller et al., 1997) from individual fibers with different spontaneous rates (SRs), and the solid line shows the model predictions for the function shown in A. The dashed line represents the normalized Q_{10} for the model normal fibers. C: Impairment in the IHC as a function of impaired CF to account for the remaining threshold shift. D: Impaired thresholds relative to the respective normal BTC for the physiological data and model as a function of impaired CF. Gray symbols indicate measured data (Miller et al., 1997) from individual fibers with different spontaneous rates (SRs). The dashed line represents the model predictions of threshold shifts for impairment in the OHC alone (as in A), and the solid line is for impairment of both OHCs and IHCs.

5 Synchronized responses of normal (A, B) and impaired (C, D) fibers, all with CFs around F2, to the vowel /ɛ/ for three sound pressure levels as labeled. Left panels show the measured responses, reprinted from Fig. 9 of Miller et al. (1997) with permission from the Acoustic Society of America © (1997), and model predictions are shown in the right panels. A: Measured data from a normal fiber with CF of 1.7 kHz and Q_{10} of 4.4. B: Model predictions from a normal model fiber with CF of 1.7 kHz and Q_{10} of 4.5. C: Measured data from an impaired fiber with CF of 1.6 kHz, a Q_{10} of 3.8, and a threshold shift $\sim$ 60 dB. D: Model predictions from an impaired model fiber; no impairment in the OHC ($C_{OHC} = 1.0$), but with IHC impairment ($C_{IHC} = 0.035$) to a normal model fiber with CF of 1.6 kHz and a Q_{10} of 3.65, produces the tuning properties with an impaired CF at 1.6 kHz, $Q_{10} \approx 3.65$, and a threshold shift of $\sim$ 55 dB, similar to those from the example fiber in Fig. 5(C).

6 Synchronized responses of normal (A, B) and impaired (C, D) fibers, all with CFs around F3, to the vowel /ɛ/ for three sound pressure levels as labeled. Left panels show the measured responses, reprinted from Fig. 10 of Miller et al. (1997) with permission from the Acoustic Society of America © (1997), and model predictions are shown in the right panels. A: Measured data from a normal fiber with CF of 2.5 kHz and Q_{10} of 4.3. B: Model predictions from a normal model fiber with CF of 2.5 kHz and Q_{10} of 4.5. C: Measured data from an impaired fiber with CF of 2.6 kHz, a Q_{10} of 1.4, and a threshold shift of $\sim$ 60dB. D: Model predictions from an impaired model fiber; appropriate impairment in the OHC ($C_{OHC} = 0.18$) and IHC ($C_{IHC} = 0.035$) to a normal model fiber with CF of 2.95 kHz and Q_{10} of 4.87, produces the tuning properties with impaired CF at 2.6 kHz, $Q_{10} \approx 1.4$, and a threshold shift of $\sim$ 60 dB, similar to those from the example fiber in Fig. 6(C).

7 Frequency spectra of responses to the vowel for a fiber with CF at 2.0 kHz for four sound pressure levels spanning from moderate to high levels as labeled. Left four panels show the measured synchronized rates, reprinted from Fig. 2 of Wong et al. (1998) with permission from Elsevier Science © (1998), and the corresponding model predictions are shown in the right four panels. Synchronized rates to the first 20 harmonics are shown by solid circles in each case. Here, the original synthesized vowel (Fig. 1) is resampled such that the second formant is positioned exactly at the fiber's CF. Note that the levels at which the experimental and model fibers undergo the C1/C2 transition are 102 and 107 dB SPL, respectively.

8 Power ratio vs. level functions for four normal fibers with CFs as labeled. Left column shows the measured PRs, reprinted from Fig. 4 of Wong et al. (1998) with permission from Elsevier Science © (1998), and model predictions are displayed in the right column. Dashed lines with solid square symbols indicate the F2 PRs and solid lines with circles represent the F1 & F2 PRs, as shown in the legend. The lower and upper bounds of the shaded regions represent, respectively, the sound pressure levels at which synchrony capture by F2 is lost (VS$\leq$ 0.5) and the C2 threshold for F1.

9 Power ratio vs. level functions for four impaired fibers with CFs as labeled. Left column shows the measured PRs, reprinted from Fig. 6 of Wong et al. (1998) with permission from Elsevier Science © (1998), and model predictions are displayed in the right column. Dashed lines with solid square symbols indicate the F2 PRs and solid lines with circles represent the F1 & F2 PRs, as shown in the legend. Experimental impaired fibers are modeled by applying appropriate impairment (C_{OHC}, C_{IHC}) to the OHC and IHC of the normal model fibers: (i) A normal model fiber with CF of 1.8 kHz and a Q_{10} of 3.86, after being impaired with $C_{OHC} = 0.0$ and $C_{IHC} = 0.0001$, gives the tuning properties of the fiber with impaired CF at 1.4 kHz, impaired $Q_{10} \approx 0.8$, and a threshold shift of ~ 75 dB. (ii) A normal fiber with CF of 1.8 kHz and a Q_{10} of 6.0, after being impaired with $C_{OHC} = 0.4$ and $C_{IHC} = 0.01$, gives the tuning properties with impaired CF at 1.7 kHz, impaired $Q_{10} \approx 3.5$, and a threshold shift of ~ 55 dB. (iii) A normal fiber with CF of 2.5 kHz and a Q_{10} of 4.5, after being impaired with $C_{OHC} = 0.6$ and $C_{IHC} = 0.008$, gives the tuning with impaired CF at 2.3 kHz, impaired $Q_{10} \approx 2.8$, and a threshold shift of ~ 60 dB. (iv) A normal fiber with CF of 2.8 kHz and a Q_{10} of 4.75, after being impaired with $C_{OHC} = 0.4$ and $C_{IHC} = 0.01$, gives the tuning with impaired CF at 2.5 kHz, impaired $Q_{10} \approx 2.0$, and a threshold shift of ~ 60 dB.

10 Model predictions of F1, F2 and F3 power ratios for normal fibers as a function of CF at two sound pressure levels: 69 dB SPL (upper panels) and 49 dB SPL (lower panels). Thick solid lines indicate model predictions and gray shaded areas show the range of PR values observed in the physiological experiments (Miller et al., 1997). Vertical dashed lines are aligned along the formant frequencies.

11 Model predictions of F1, F2 and F3 power ratios for impaired fibers as a function of impaired CF at two sound pressure levels: 112 dB SPL (upper panels) and 92 dB SPL (lower panels). Impaired fibers are modeled by developing a functional relationship between C_{OHC}, C_{IHC} and impaired CF—see Section II.B.*2*. Thick solid lines show the model predictions and gray shaded areas indicate the range of PR values observed in the physiological experiments (Miller et al., 1997). Vertical dashed lines indicate the formant frequencies.

12 Comparison of C2 thresholds for the two formants (in vowel responses) and for tones at the frequencies of F1 and F2. Left column (A–C) shows the measured responses, reprinted from Fig. 3 of Wong et al. (1998) with permission from Elsevier Science © (1998), and model predictions are shown in the right column (D–F). Model predictions are based on the responses from 21 fibers with CFs ranging from 1 to 3 kHz. In each part, the solid line indicates where the abscissa and ordinate values are equal. A,D: Comparison of the C2 thresholds (in dB SPL) for the formant F1 in vowel responses and for the frequency F1 in tonal responses. Dashed lines represent ± 5 dB SPL deviations from the equality line. B,E: Comparison of the C2 thresholds for the formant F2 in vowel responses and the frequency F2 in tonal responses. Dashed line shows 10 dB SPL difference. C,F: Comparison of the overall vowel levels at which the two formants F1 and F2 undergo their corresponding C1/C2 transitions. Again, dashed lines indicate ± 5 dB SPL deviations from the equality line.

13 Upper panel: line spectrum of the CEFS vowel (Miller et al., 1999a). Lower panel: comparison of the C2 thresholds for the formant F2 in vowel responses and the frequency F2 in tonal responses. Dashed lines indicate ± 5 dB SPL deviations from the equality line.

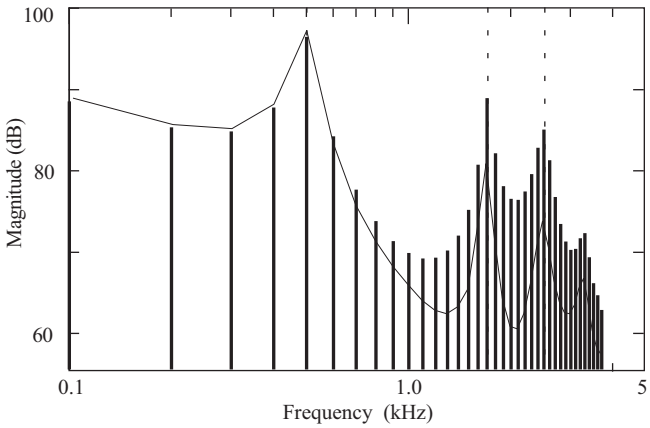

FIG. 1.

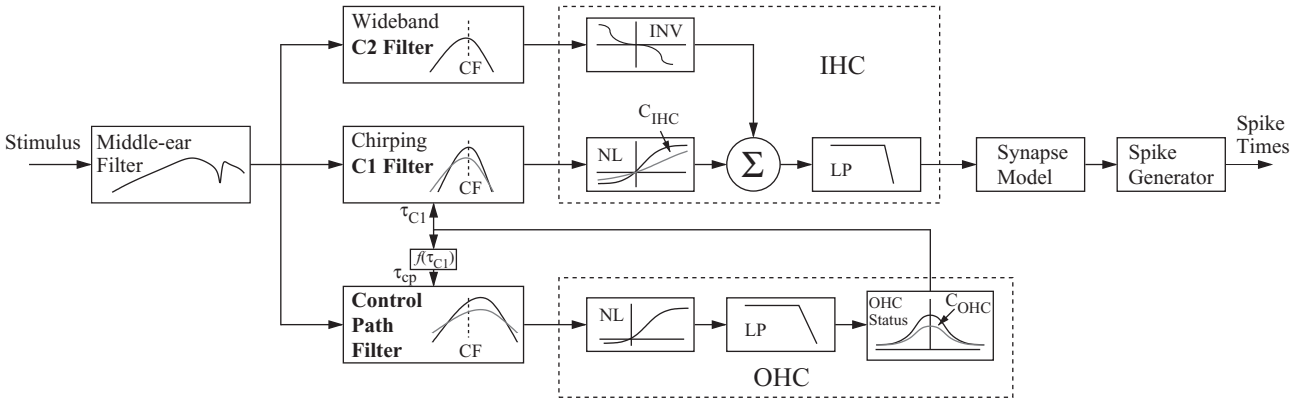
Stimulus
Middle-ear Filter
Wideband C2 Filter
CF
Chirping C1 Filter
CF
Control Path Filter
CF
τ_{C1}
τ_{cp}
$f(\tau_{C1})$
INV
IHC
C_{IHC}
NL
Σ
LP
Synapse Model
Spike Generator
Spike Times
NL
LP
OHC Status
C_{OHC}
OHC

FIG. 2.

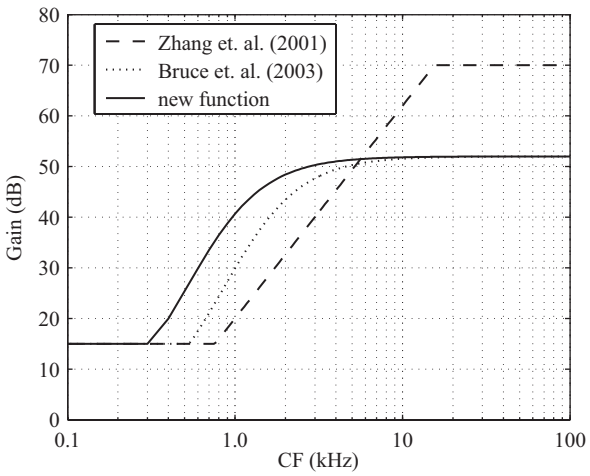
Zhang et. al. (2001)
Bruce et. al. (2003)
new function
Gain (dB)
CF (kHz)
80
70
60
50
40
30
20
10
0
0.1
1.0
10
100

FIG. 3.

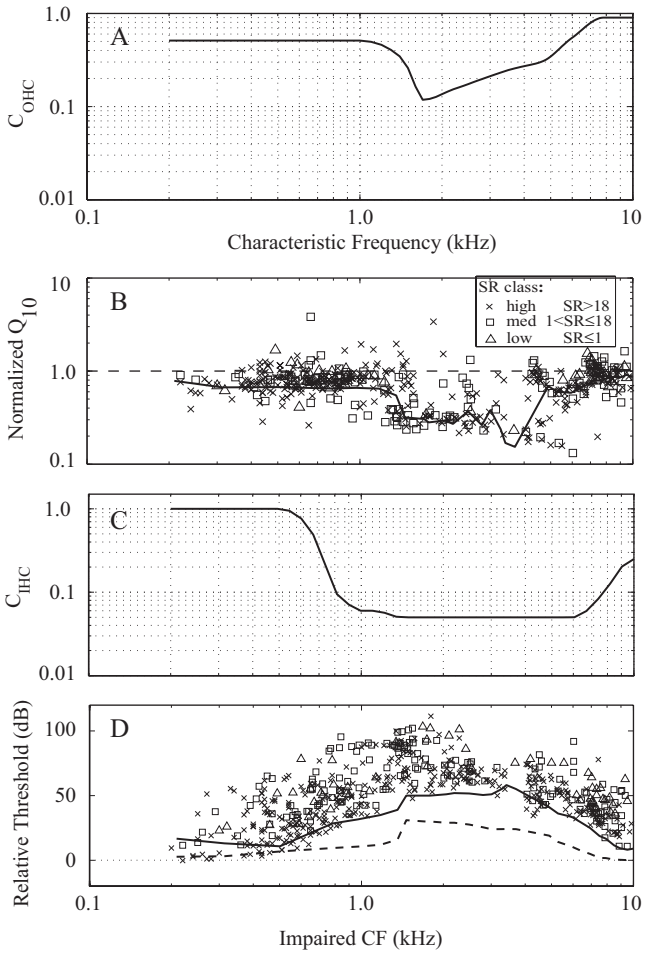
A
B
C
D
C_{OHC}
Normalized Q_{10}
C_{IHC}
Relative Threshold (dB)
Characteristic Frequency (kHz)
Impaired CF (kHz)
SR class:
high SR>18
med 1<SR≤18
low SR≤1
1.0
0.1
0.01
10
100
50
0

FIG. 4.

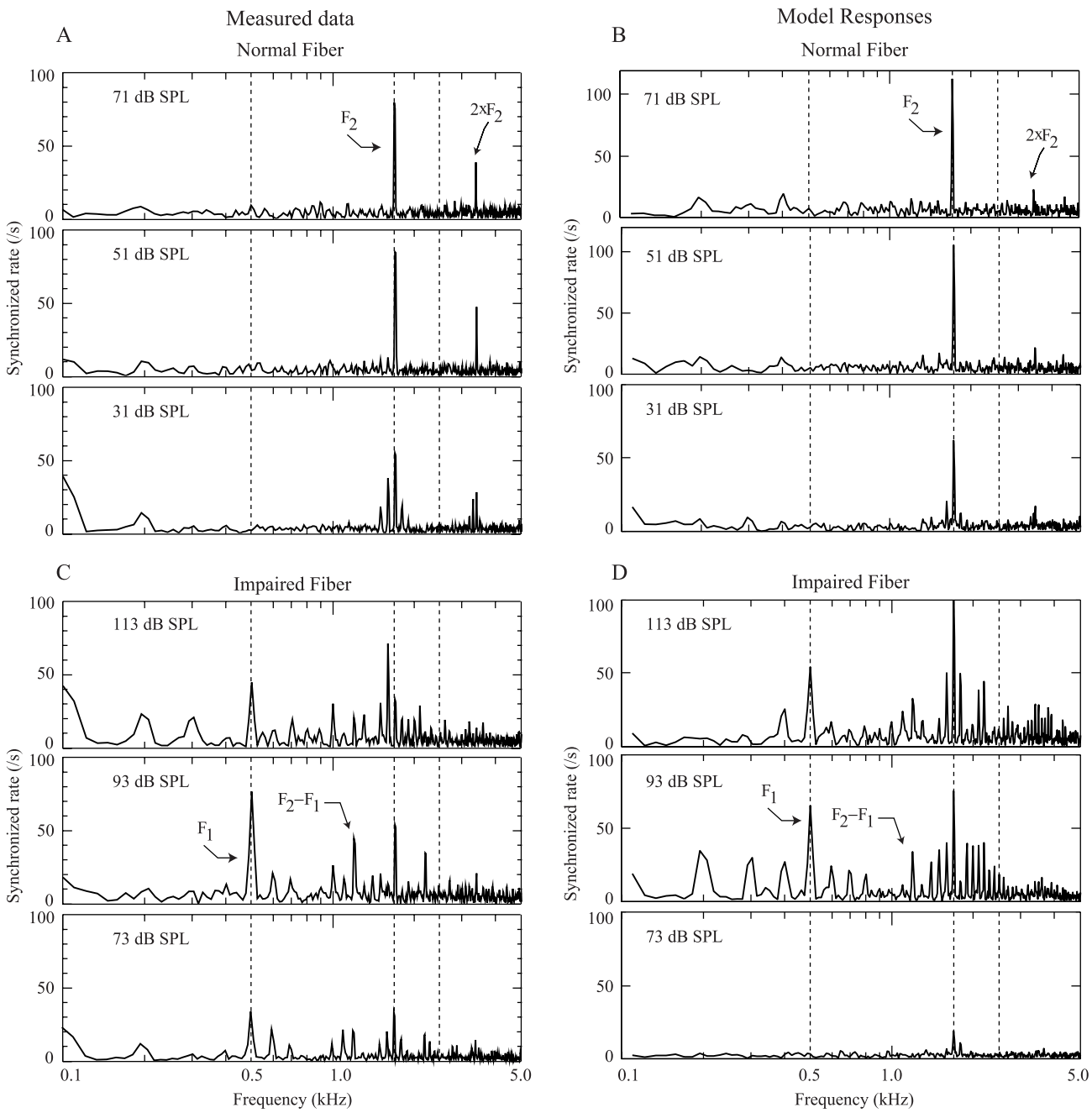
Measured data
Model Responses
A
Normal Fiber
B
Normal Fiber
71 dB SPL
51 dB SPL
31 dB SPL
F2
2xF2
Synchronized rate (/s)
C
Impaired Fiber
D
Impaired Fiber
113 dB SPL
93 dB SPL
73 dB SPL
F1
F2−F1
0.1
0.5
1.0
5.0
Frequency (kHz)

FIG. 5.

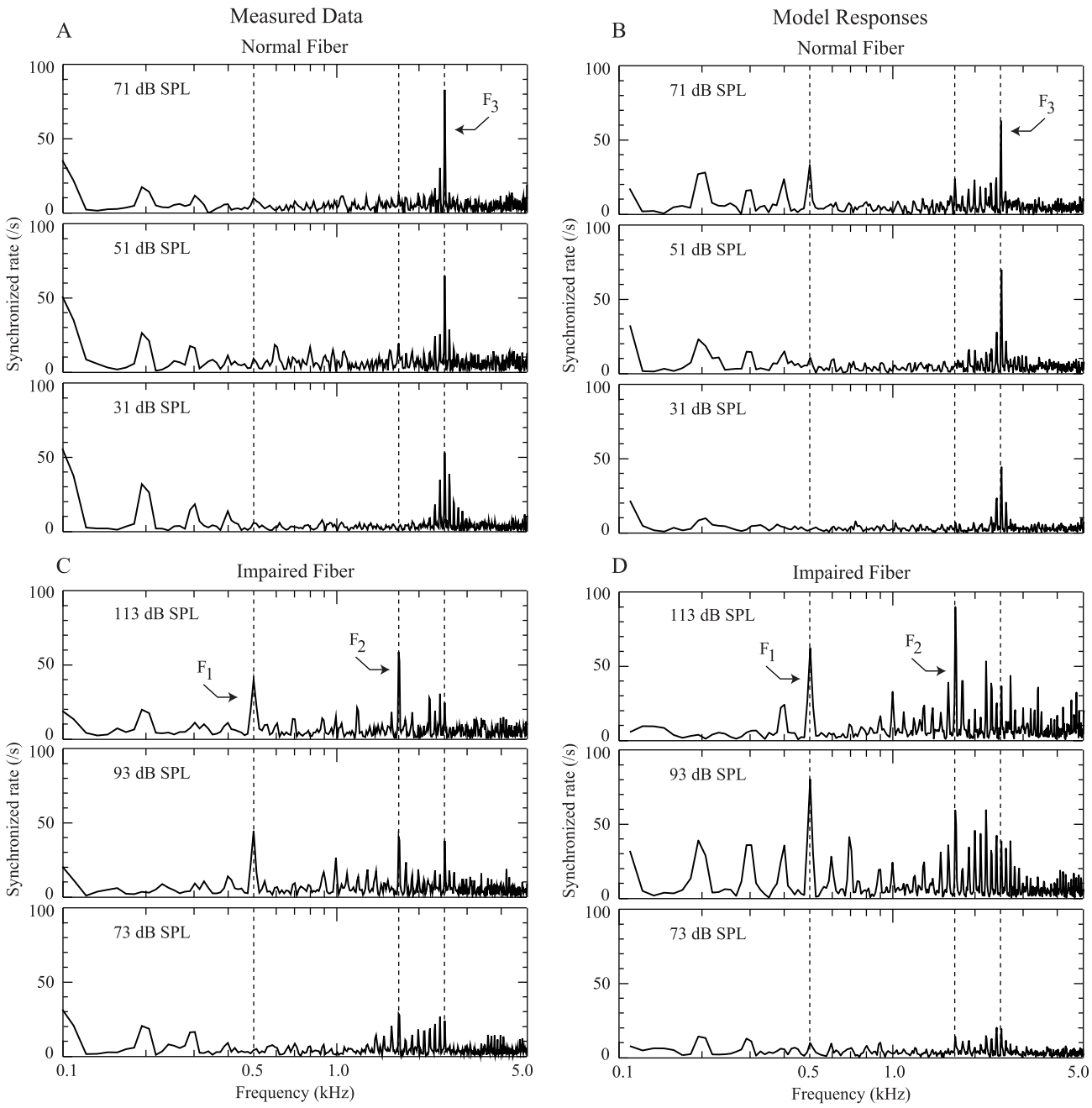
Measured Data
Model Responses
A
Normal Fiber
B
Normal Fiber
71 dB SPL
51 dB SPL
31 dB SPL
F3
Synchronized rate (/s)
C
Impaired Fiber
D
Impaired Fiber
113 dB SPL
93 dB SPL
73 dB SPL
F1
F2
0.1
0.5
1.0
5.0
Frequency (kHz)

FIG. 6.

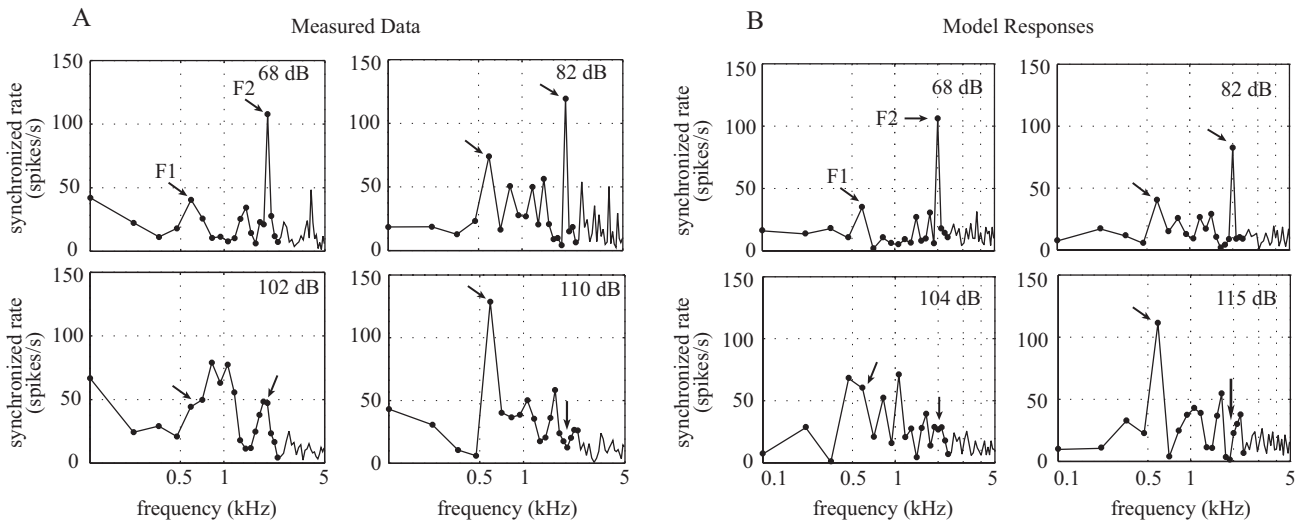
A
Measured Data
B
Model Responses
synchronized rate (spikes/s)
68 dB
82 dB
102 dB
110 dB
104 dB
115 dB
F1
F2
frequency (kHz)

FIG. 7.

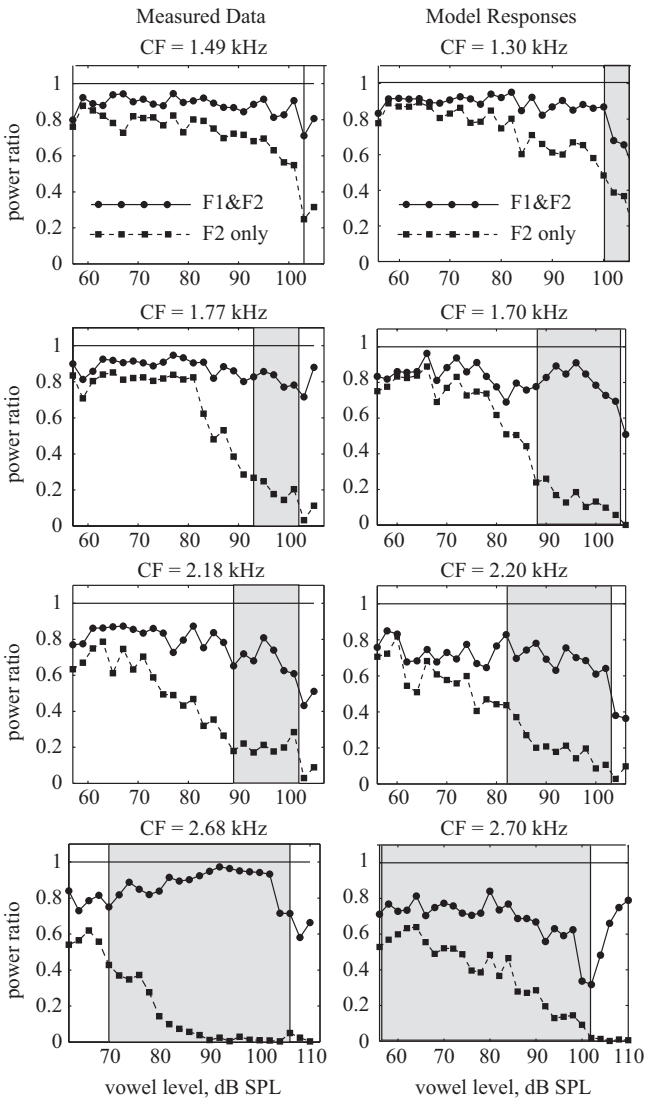
Measured Data
Model Responses
CF = 1.49 kHz
CF = 1.30 kHz
CF = 1.77 kHz
CF = 1.70 kHz
CF = 2.18 kHz
CF = 2.20 kHz
CF = 2.68 kHz
CF = 2.70 kHz
power ratio
F1&F2
F2 only
vowel level, dB SPL

FIG. 8.

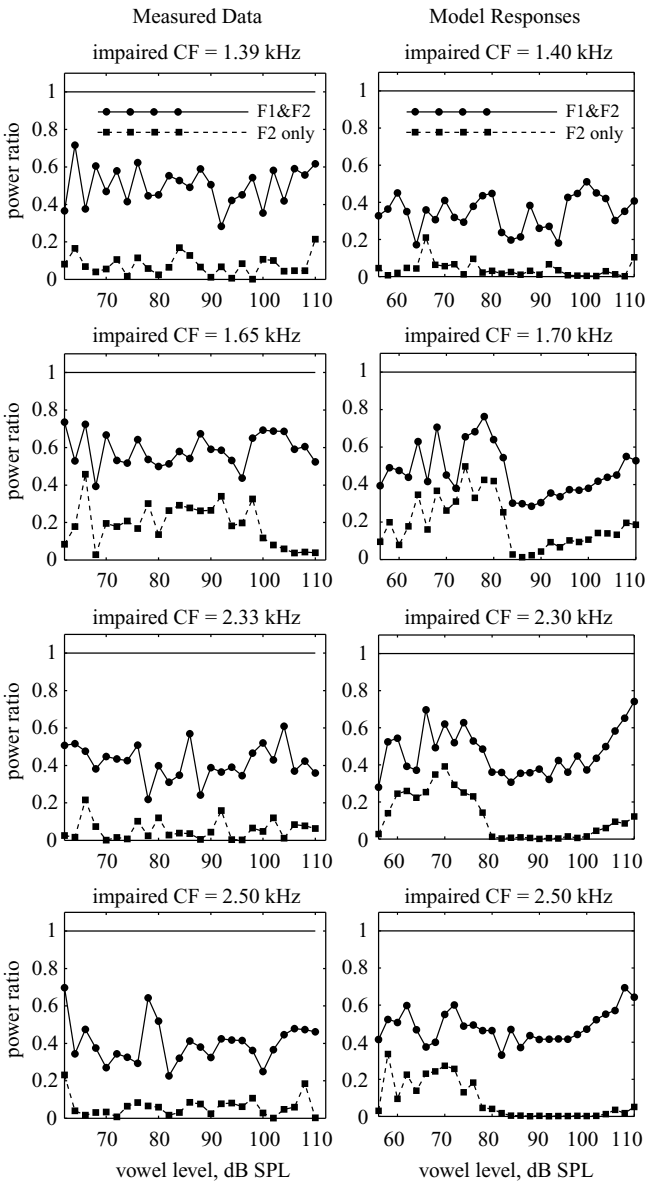
Measured Data
Model Responses
impaired CF = 1.39 kHz
impaired CF = 1.40 kHz
impaired CF = 1.65 kHz
impaired CF = 1.70 kHz
impaired CF = 2.33 kHz
impaired CF = 2.30 kHz
impaired CF = 2.50 kHz
impaired CF = 2.50 kHz
F1&F2
F2 only
power ratio
vowel level, dB SPL

FIG. 9.

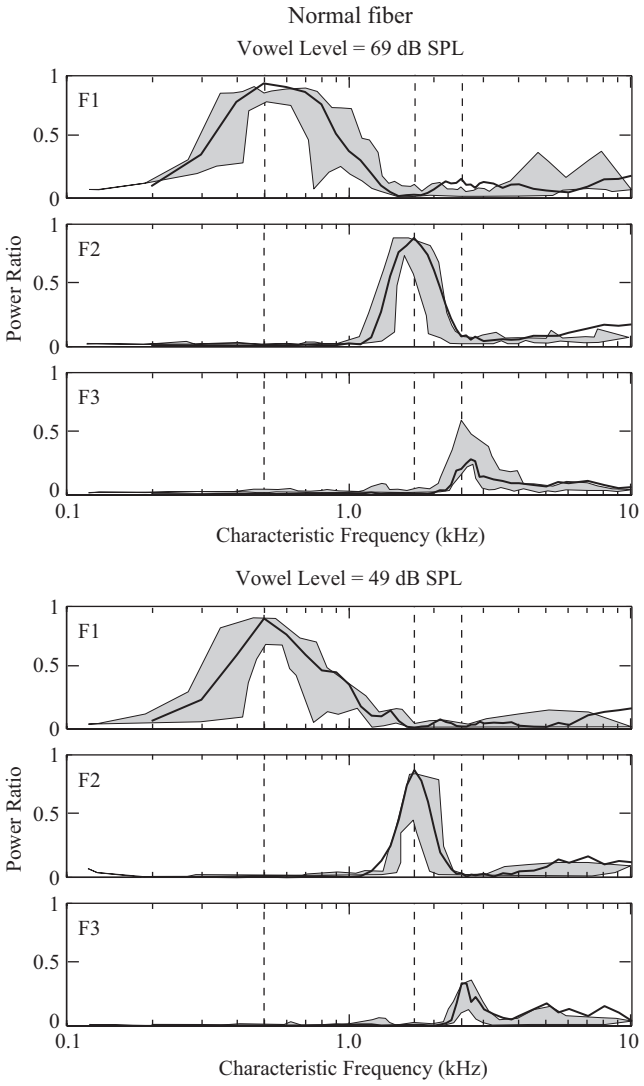
Normal fiber
Vowel Level = 69 dB SPL
F1
F2
F3
Power Ratio
1
0.5
0
0.1
1.0
10
Characteristic Frequency (kHz)
Vowel Level = 49 dB SPL
F1
F2
F3
Power Ratio
Characteristic Frequency (kHz)

FIG. 10.

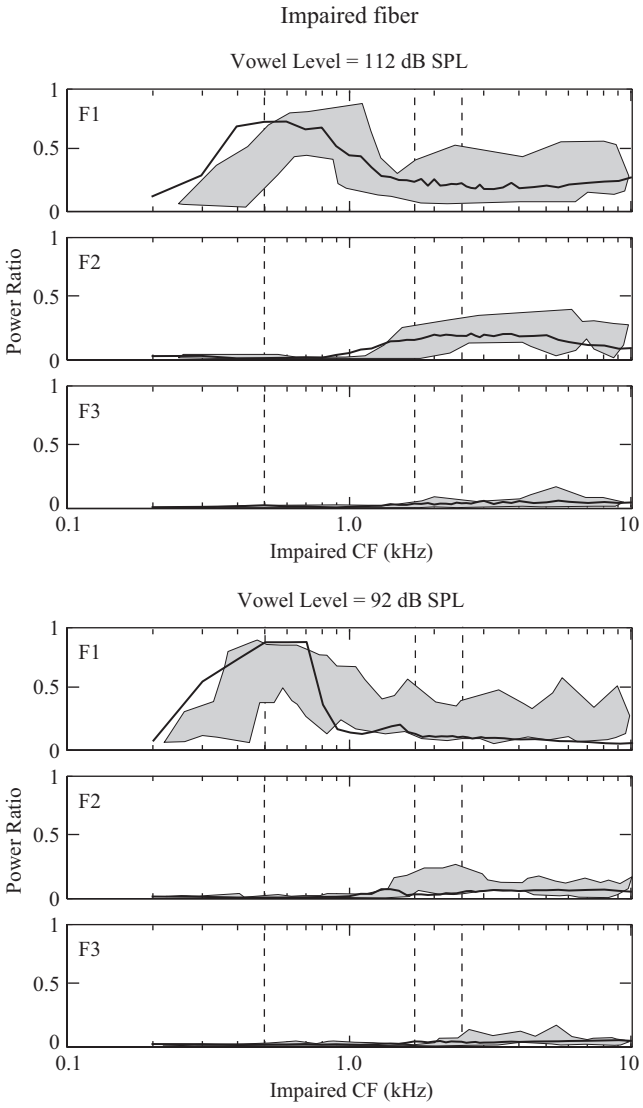
Impaired fiber
Vowel Level = 112 dB SPL
F1
F2
F3
Power Ratio
1
0.5
0
0.1
1.0
10
Impaired CF (kHz)
Vowel Level = 92 dB SPL
F1
F2
F3
Power Ratio
Impaired CF (kHz)

FIG. 11.

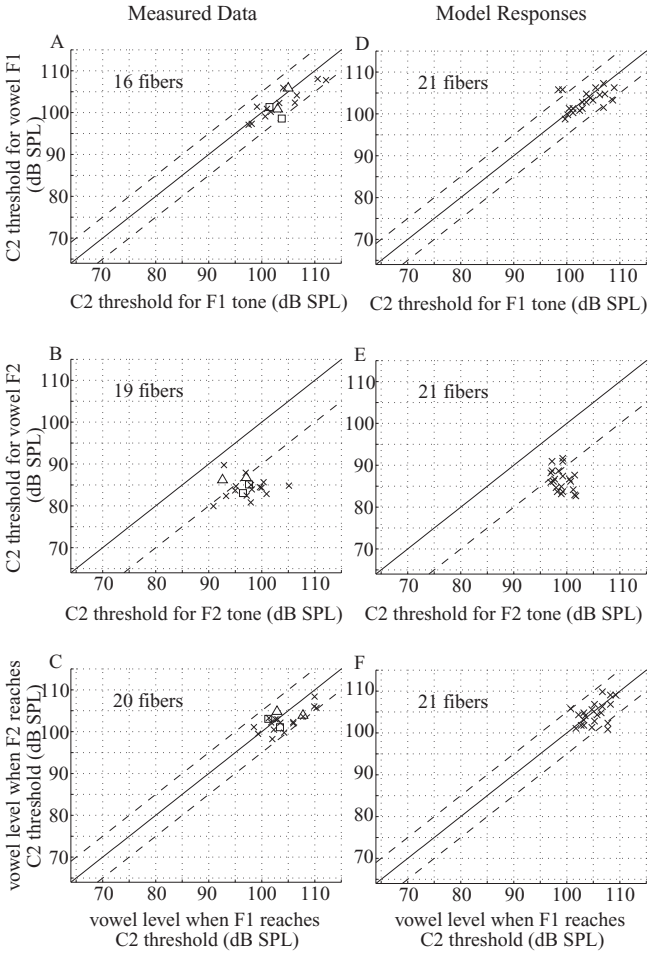
Measured Data
Model Responses
A
D
16 fibers
21 fibers
C2 threshold for vowel F1 (dB SPL)
C2 threshold for F1 tone (dB SPL)
B
E
19 fibers
21 fibers
C2 threshold for vowel F2 (dB SPL)
C2 threshold for F2 tone (dB SPL)
C
F
20 fibers
21 fibers
vowel level when F2 reaches C2 threshold (dB SPL)
vowel level when F1 reaches C2 threshold (dB SPL)
70
80
90
100
110

FIG. 12.

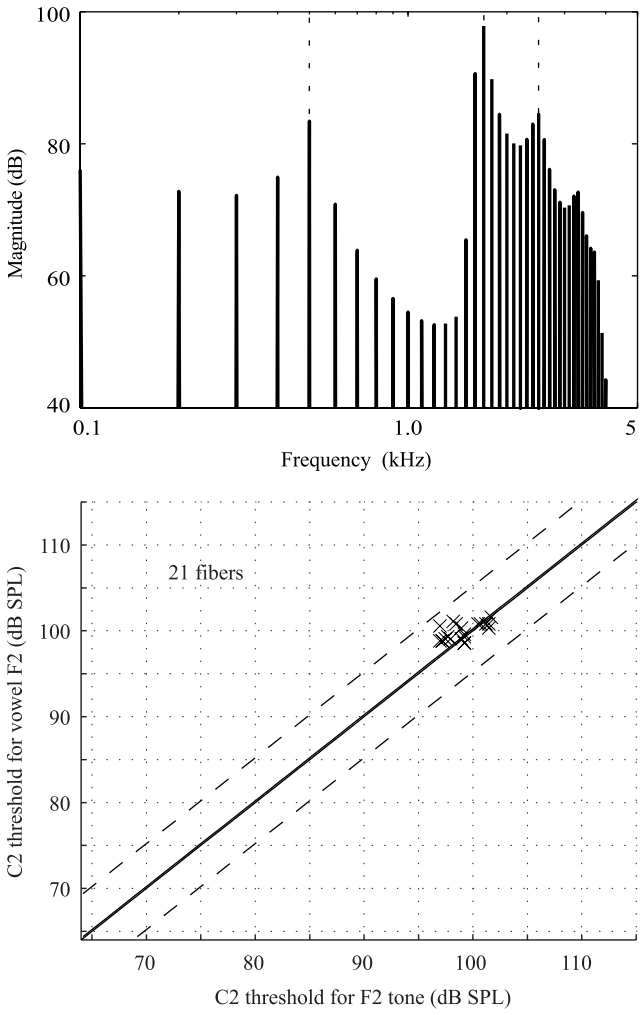
100
80
60
40
Magnitude (dB)
0.1
1.0
5
Frequency (kHz)
21 fibers
110
100
90
80
70
C2 threshold for vowel F2 (dB SPL)
70
80
90
100
110
C2 threshold for F2 tone (dB SPL)

FIG. 13.