Physiological Modeling for Hearing Aid Design

Ian C. Bruce, Eric D. Young and Murray B. Sachs

Center for Hearing Sciences and Dept. of Biomedical Engineering

The Johns Hopkins University School of Medicine, Baltimore, MD 21205

JOHNS HOPKINS
UNIVERSITY

0 Abstract

Perceptual models of impaired hearing are being used increasingly in hearing aid design. The most successful application has been compression schemes to compensate for loudness recruitment. Less fruitful have been attempts to counteract degraded cochlear filtering with spectral shaping. Perceptual models are developed from psychophysical measures that reflect both peripheral and central processing. More physiological detail may be required to describe the effects of a cochlear lesion on peripheral coding of speech.

For example, physiological data from hearing-impaired cats indicate that conventional hearing aid signal processing schemes do not restore normal auditory nerve responses to a vowel (Miller et al., JASA 101:3602, 1997) and can even produce anomalous and potentially confounding patterns of activity (Schilling et al., Hear. Res. 117:57, 1998). These deficits in the neural representation may at least partially account for poor speech perception in some hearing aid users. An amplification scheme has been developed that produces neural responses to a vowel more like those seen in normal cats and that minimizes confounding responses (Miller et al., JASA 106:2693, 1999).

A physiological model of the normal and impaired auditory periphery would provide simpler and quicker testing of such potential hearing aid designs. Details of a physiological model will be presented. Model predictions of vowel responses suggest that degraded cochlear filtering can indeed account for a good deal of the data from hearing-impaired cats described above. However, some response properties appear to result from physiological features that are not considered in perceptual models. In particular, auditory nerve responses to speech stimuli are very sensitive to wide-band nonlinearities in the basilar membrane mechanics known as two-tone suppression. These nonlinearities strongly affect how well auditory nerve fibers synchronize to specific formants at different stimulus intensities (Wong et al., Hear. Res. 123:61, 1998). Such factors will be important in generalizing the speech-processing algorithm described above to running speech. Model predictions of a prospective amplification scheme will be discussed.

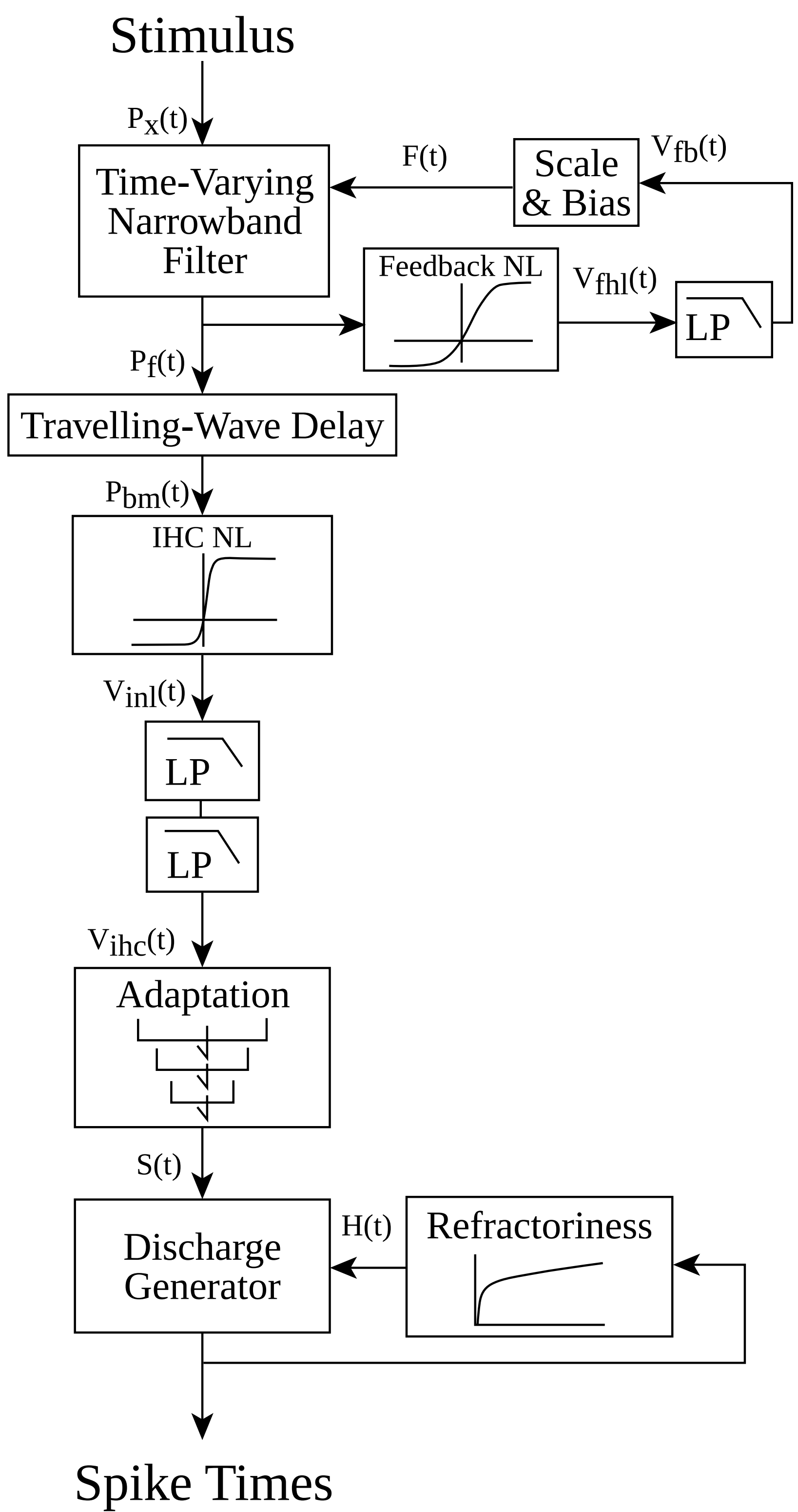
1 Modeling normal outer hair cell control of basilar membrane filtering

The auditory-periphery model of Carney (1993) comprises several sections, each modeling a different part of cochlea function. This is illustrated in the schematic diagram below.

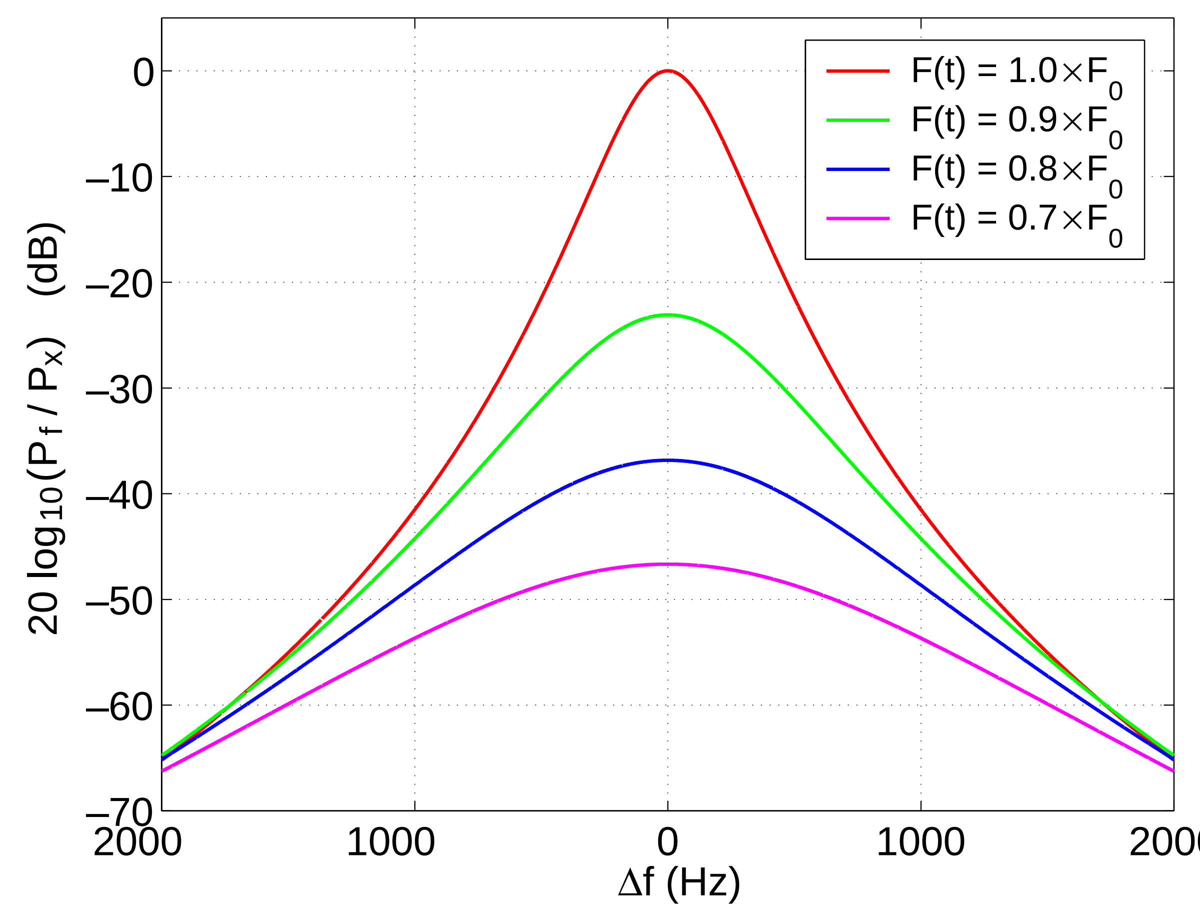
The first section describes the outer hair cell (OHC) controlled filter properties and the traveling wave delay of the basilar membrane (BM).

The second section models the nonlinear (NL) transduction function and low-pass (LP) filtering of the inner hair cell (IHC).

The third section describes adaptation of synaptic transmission between the IHC and the auditory nerve (AN) and spike generation and refractoriness in the AN.



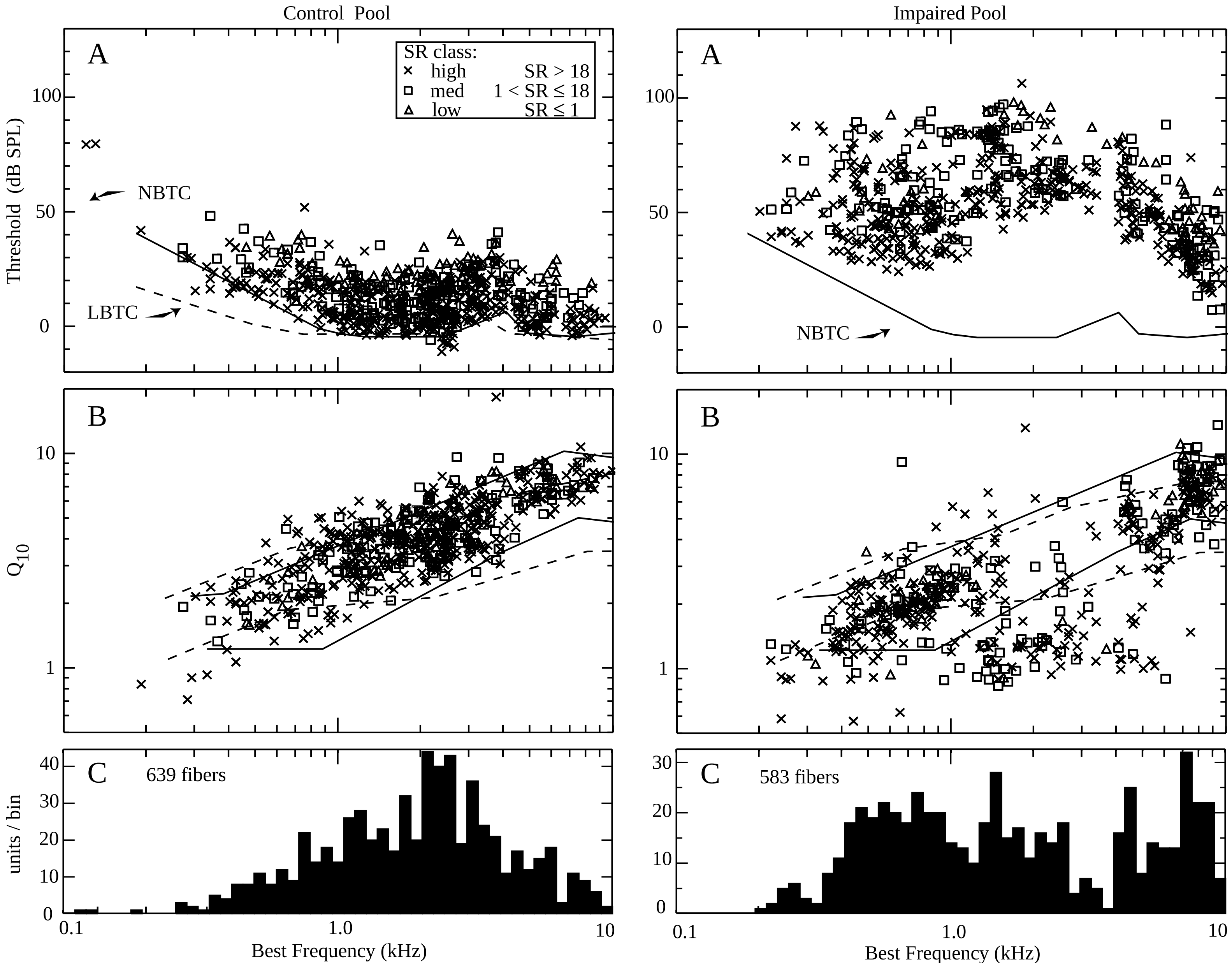
The figure below shows how the gain and bandwidth of the normal BM filter changes as a function of the feedback signal F(t), where F_0 is the resting value of F(t), i.e., when $P_x = 0$. At low stimulus intensities F(t) is ~100%; as the stimulus intensity grows, F(t) is reduced, with a maximum reduction of just under ~30%. Decreasing F(t) increases both the bandwidth and the attenuation. Consequently, the filter is sharp and linear at low intensities and broad and compressive at high intensities.

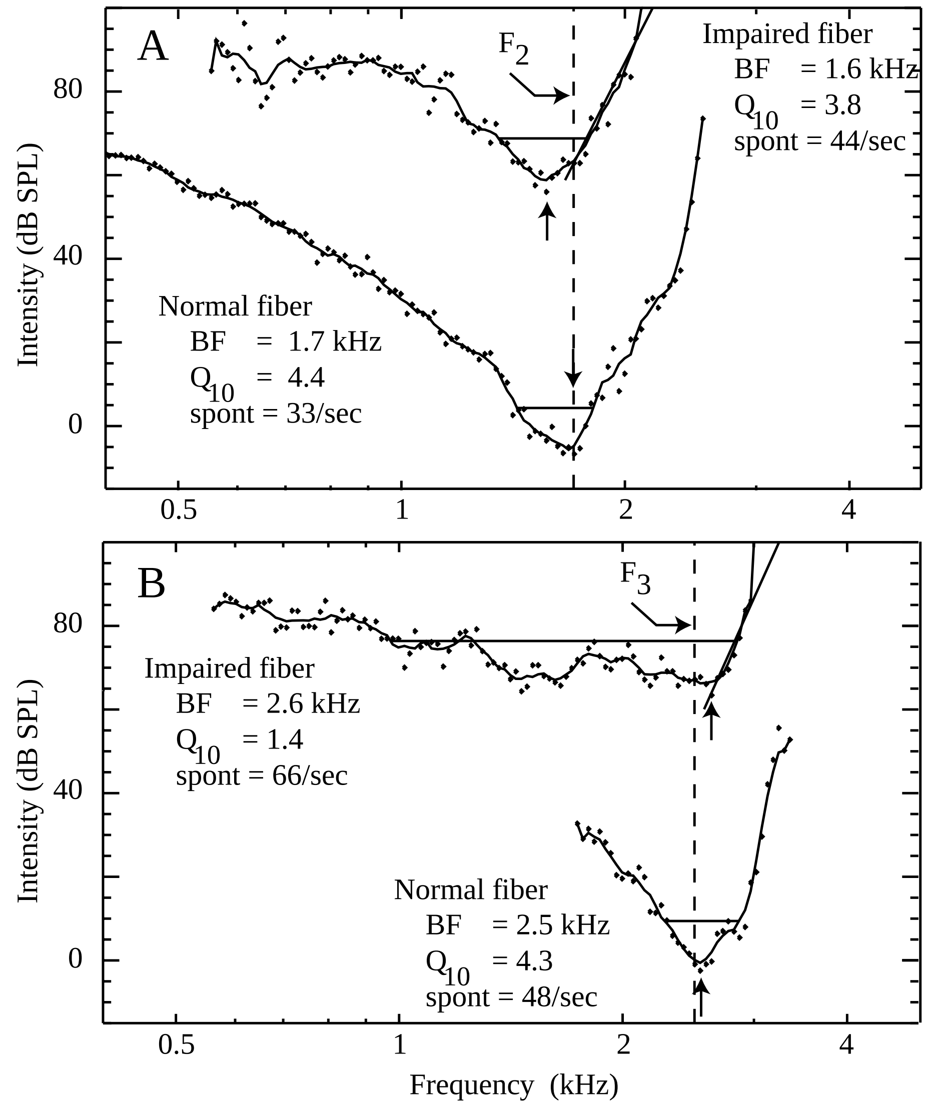


2 Tuning in normal and impaired auditory nerve fibers

Auditory nerve population data from Miller et al. (1997) are presented below. The “Control Pool” comprises data from normal cats. The “Impaired Pool” consists of data from acoustically traumatized cats (2 hours of a 110–115 dB SPL, 50 Hz noise band centered at 2 kHz). Fibers impaired by acoustic trauma exhibit elevated thresholds in the region of the sound exposure (see figures A below) and broadened tuning (see figures B below).

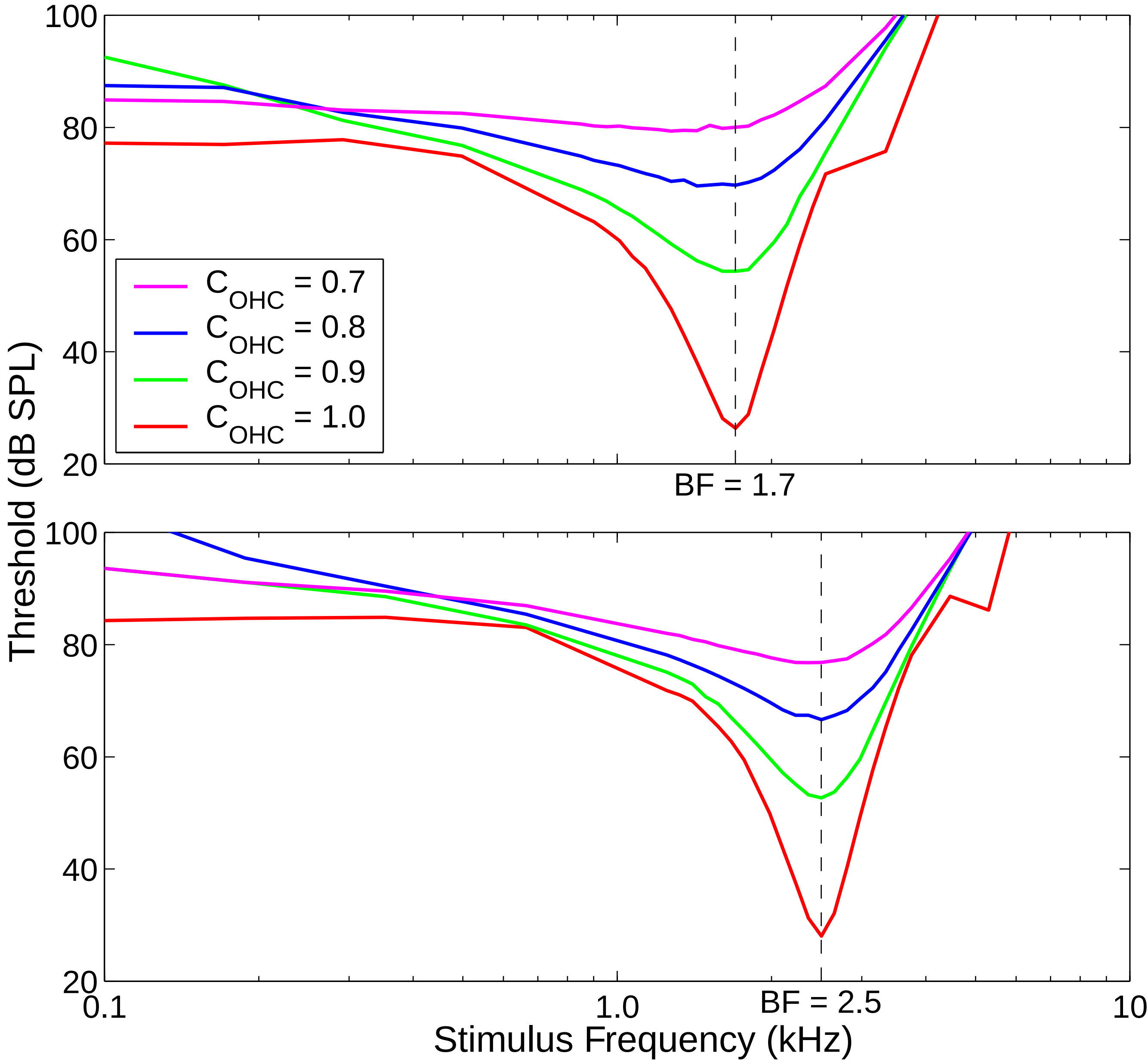
Shown below are tuning curves for normal and impaired fibers with BFs near the 2nd (figure A) and 3rd (figure B) formant of the vowel /eh/. Note that both formants fall within the region of acoustic trauma in the impaired pool.





3 Impaired outer hair cell function can describe elevated thresholds and broadened tuning

Damage to OHCs or their stereocilia impairs their mechanical amplification of BM vibrations and can consequently explain the broadened and elevated tips of threshold tuning curves as shown in Panel 2 (Liberman and Dodds, 1984). In order to model the effects of OHC status on the BM filter, we multiply the feedback signal F(t) by a scaling constant C_{OHC}, where $0 < C_{OHC} \leq 1$. For normal OHC function, C_{OHC} is set to 1 and the filter output remains unaffected. For impaired OHC function, C_{OHC} is set between 0 and 1. The lower the value of C_{OHC}, the broader the tuning and the greater the attenuation of the BM filter at low intensities, and the less compressive (i.e., more linear; Harrison, 1981; Patuzzi et al., 1989) is the filter at high intensities. The effect of C_{OHC} on model tuning curves is illustrated to the right for two model fibers with BFs at the 2nd and 3rd formants, respectively, of the vowel /eh/.



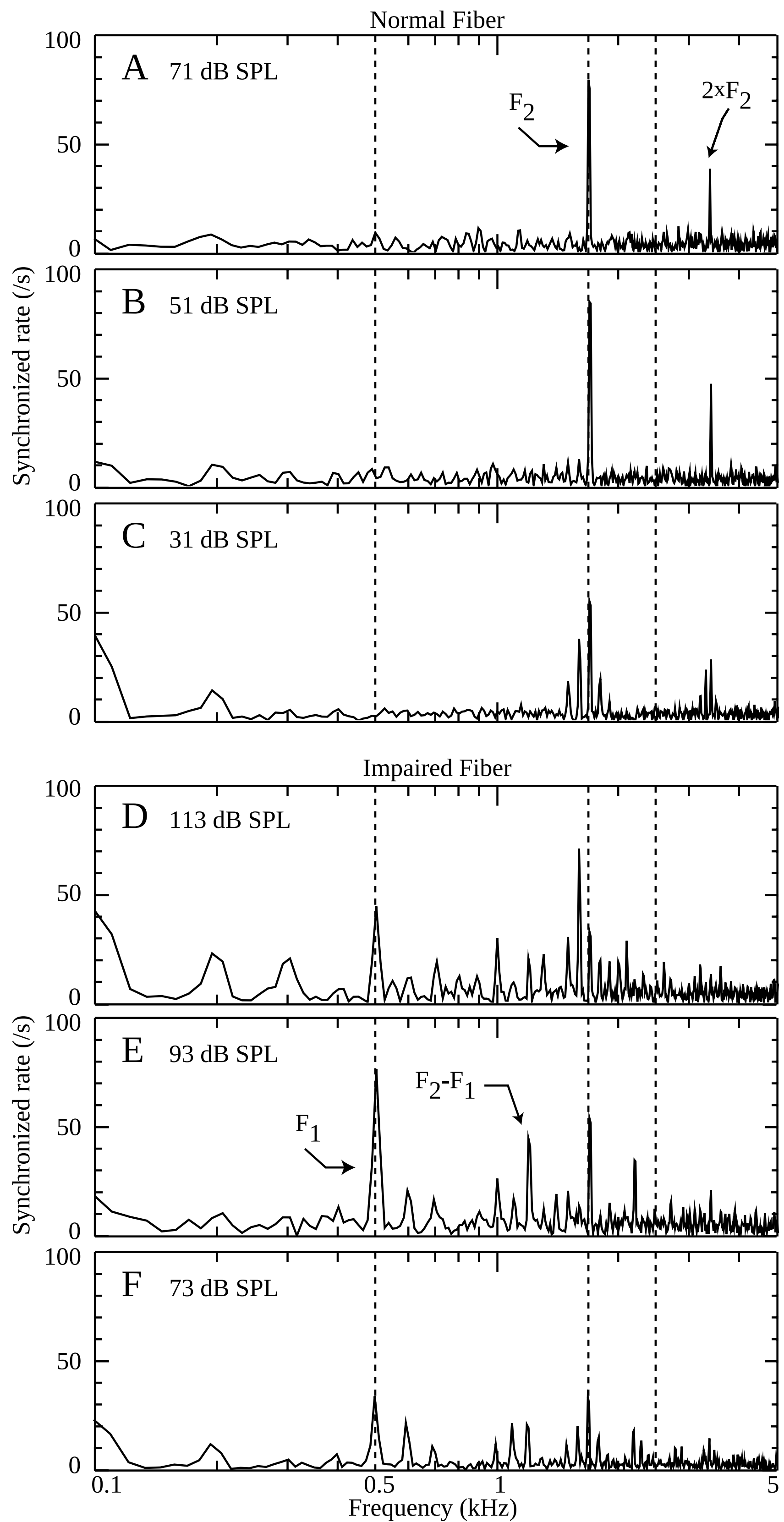
4 Synchronized response to the vowel /eh/

The synchronized response of the four fibers from Panel 2 to the vowel /eh/ is evaluated by taking the Fourier transform of the PSTH (i.e., the instantaneous discharge rate in response to the vowel) normalized to units of spikes/s (Miller et al., 1997).

BF ≈ F2

The normal fiber with BF near F2 exhibits synchrony capture, i.e, as the stimulus increases above the fiber threshold, the response becomes synchronized almost exclusively to F2.

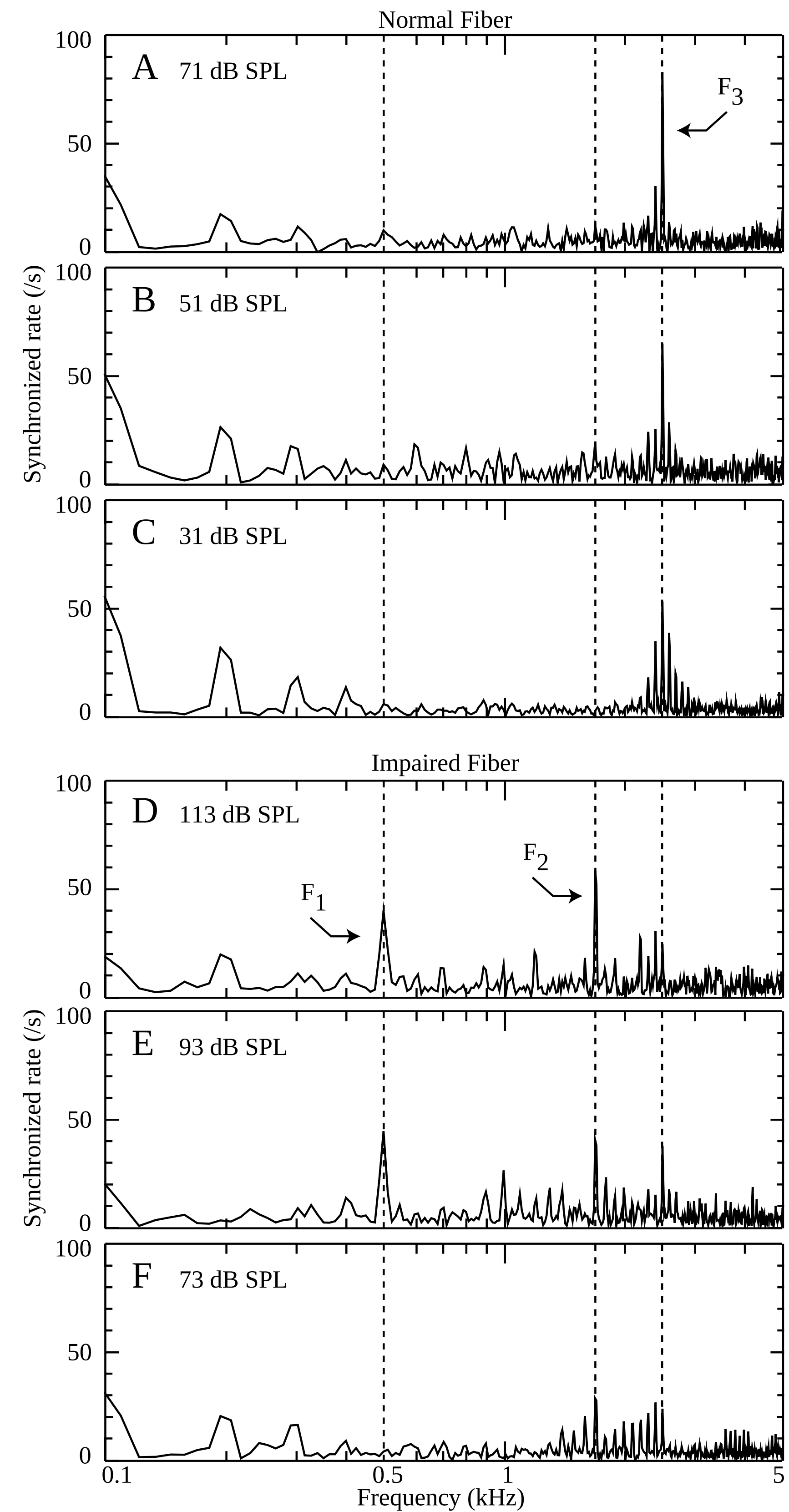
In contrast, the impaired fiber shows much broader tuning, particularly to the higher-intensity 1st formant.



BF ≈ F3

Some synchrony capture is also observed in the normal fiber with BF near F3. Synchrony capture is not observed in all fibers near F3, although they do exhibit a strong response to F3.

The impaired fiber again shows much broader tuning, synchronized particularly to F1 and F2.



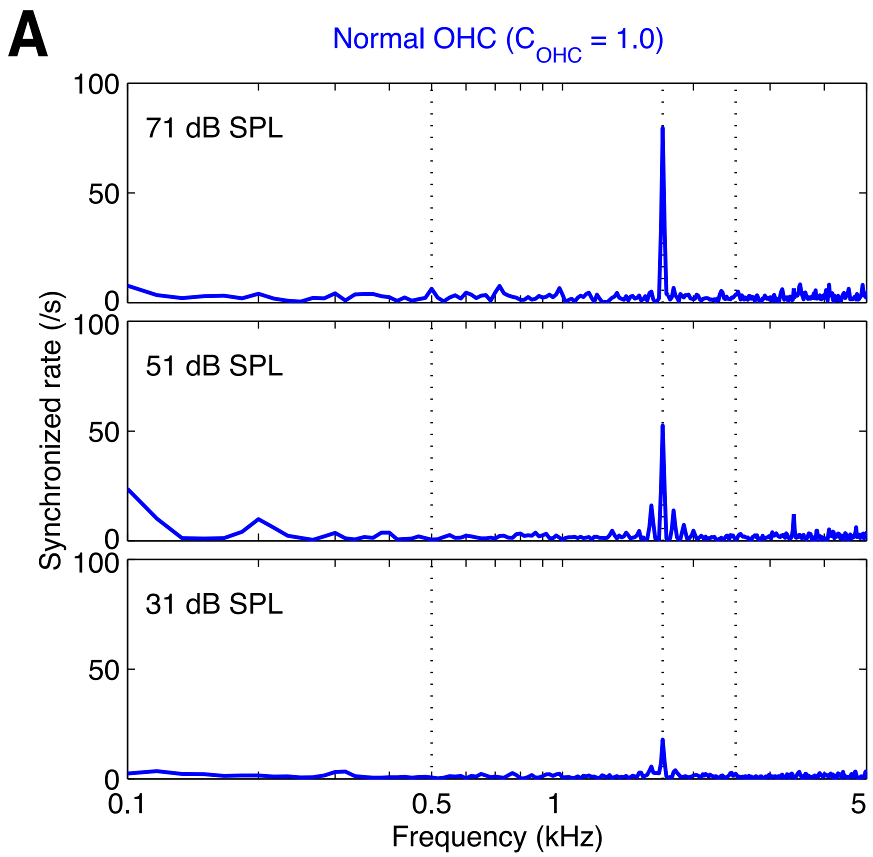
5 Model predictions of the synchronized response to the vowel /eh/

The synchronized responses of two model fibers (with BFs matching the fibers of Panel 4) with and without OHC impairment are compared below.

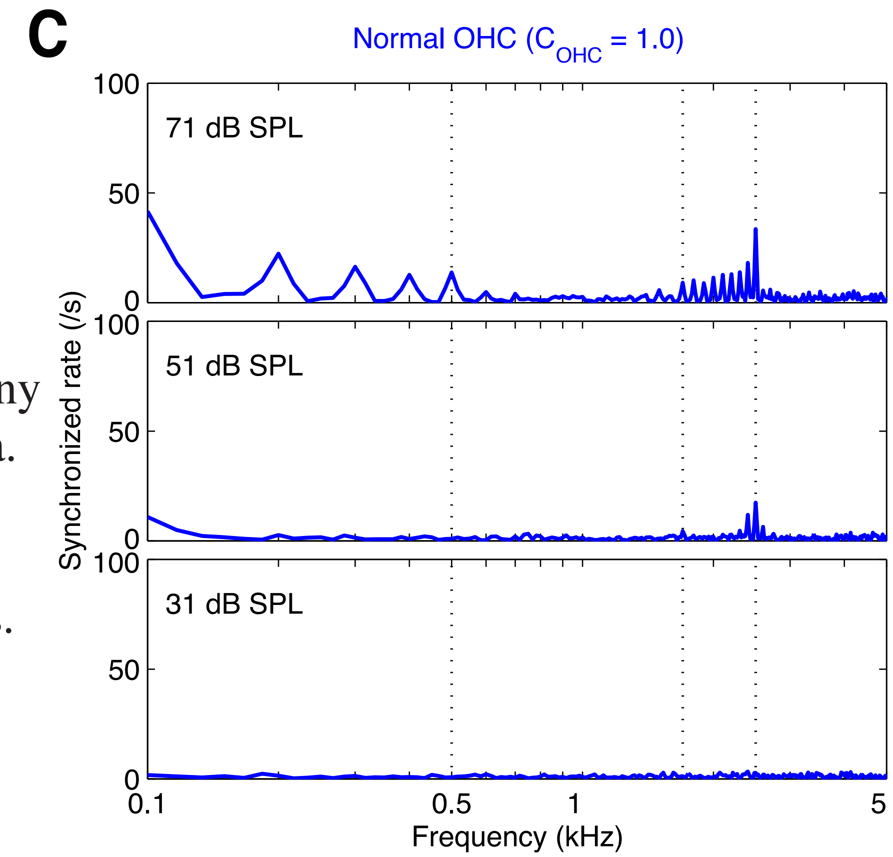
BF ≈F2

BF ≈F3

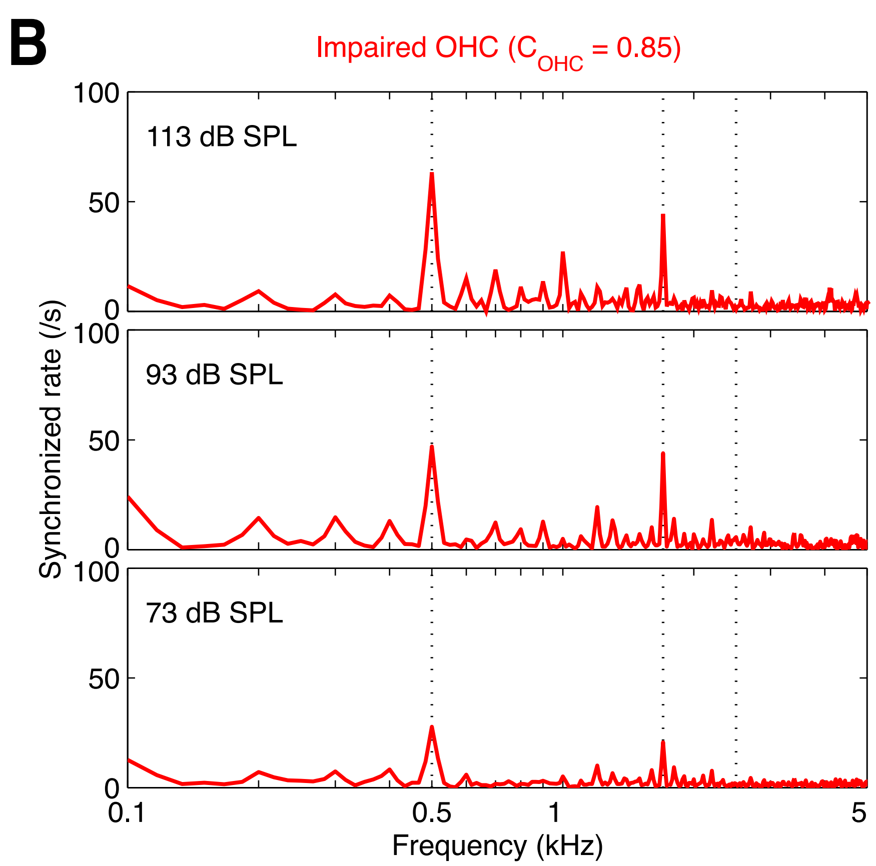
The model fiber with normal OHC function ($C_{OHC} = 1$) and BF near F2 exhibits synchrony capture, consistent with the physiological data.



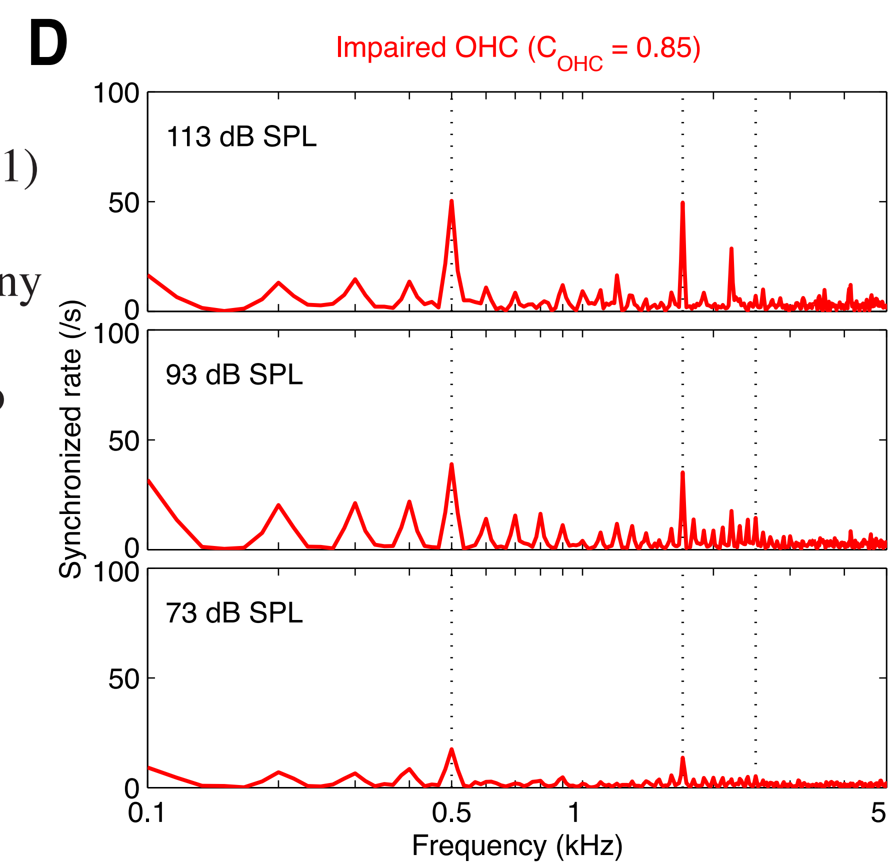
The model fiber with normal OHC function and BF near F3 does not exhibit the same degree of synchrony capture as the data. This is due to limitations of the model at high BFs.



Likewise, the model fiber with impaired OHC function ($C_{OHC} < 1$) shows the broader tuning and large F1 response observed in the AN data.

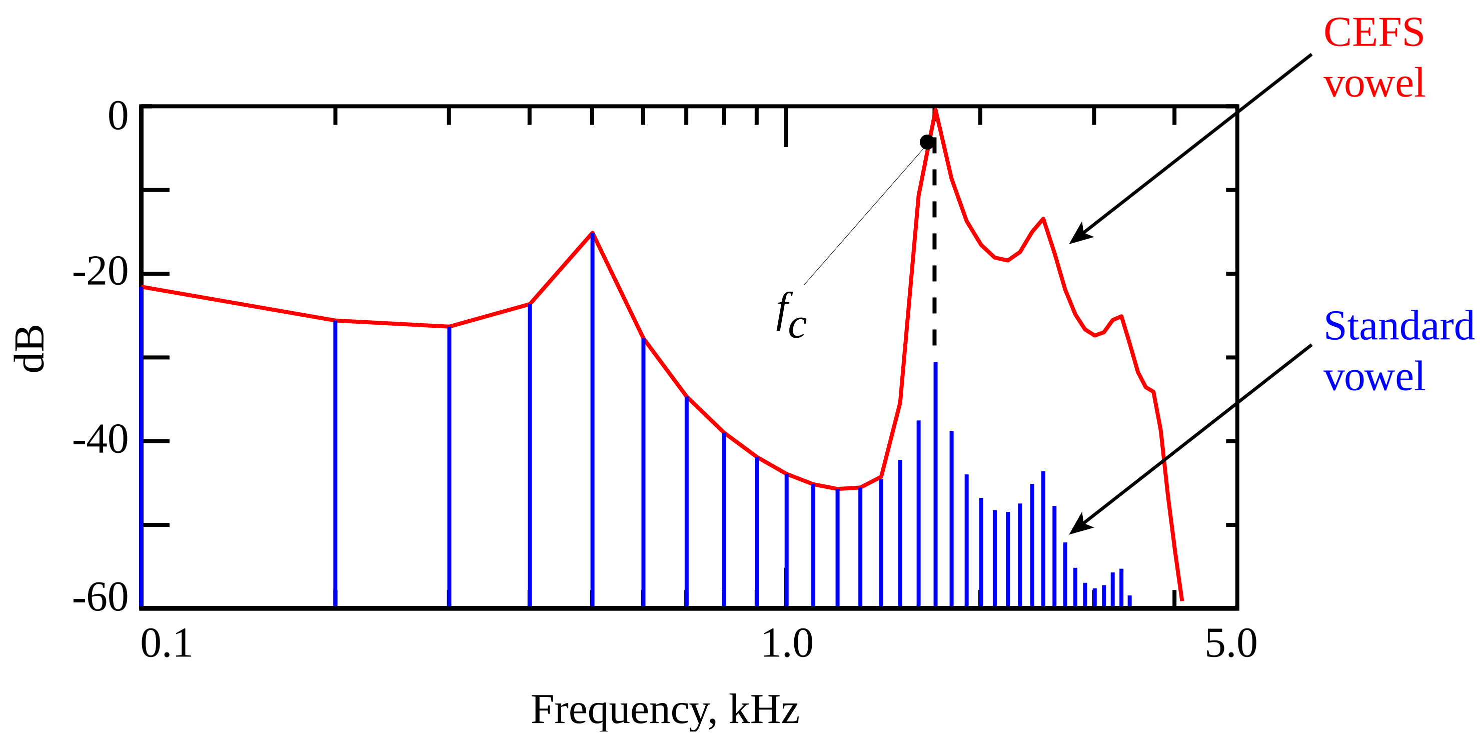


Impaired OHC function ($C_{OHC} < 1$) does lead to increased synchrony to F1 and F2 and loss of response to F3.



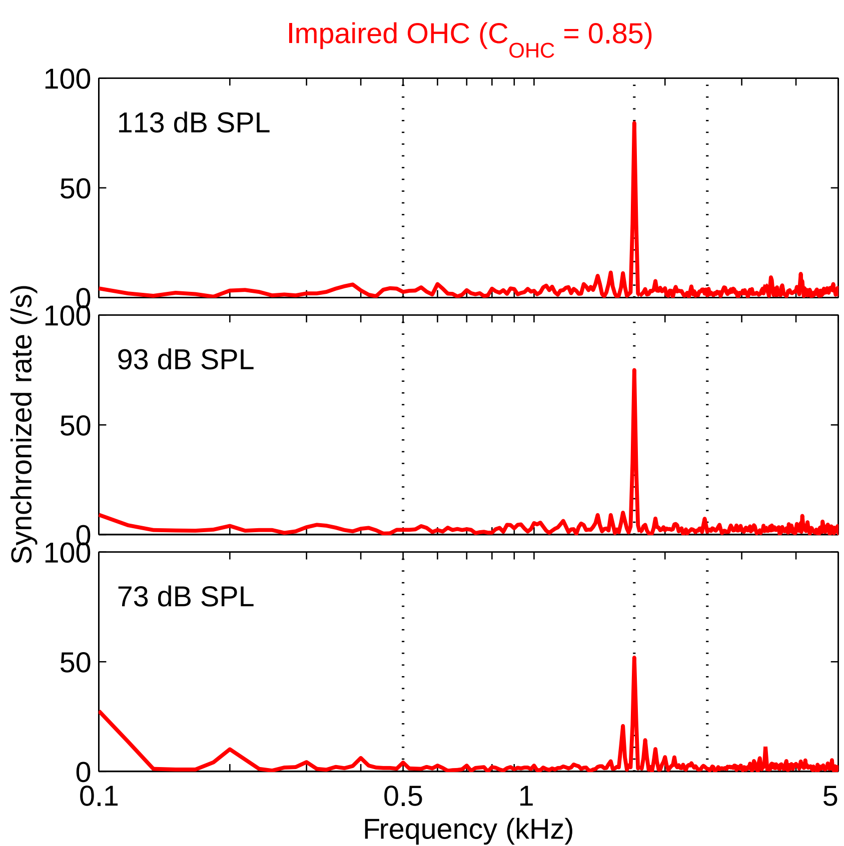
6 Model predictions for a contrast-enhanced frequency shaped (CEFS) vowel

In Schilling et al. (1998) and Miller et al. (1999), shaping of the vowel spectrum was investigated to see if a modified vowel could cause an impaired fiber to regain synchrony capture. It was found that it could indeed be recovered by enhancing the second and third formants and suppressing the first formant and the components in the "trough" between F1 and F2. Spectra of the standard and CEFS-modified vowel /eh/ are shown below, along with model responses for fibers with BFs at F1 and F2.



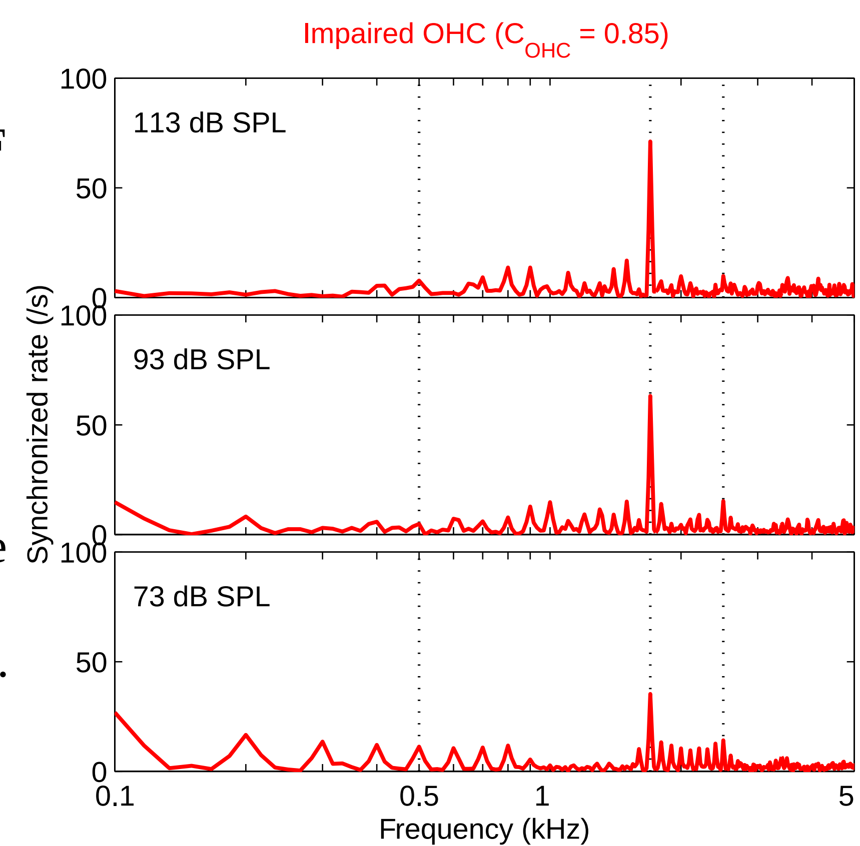
BF ≈ F2

For an OHC-impaired model fiber with BF near F2, synchrony capture is indeed regained with the modified vowel (*cf.* Panels 5A&B).



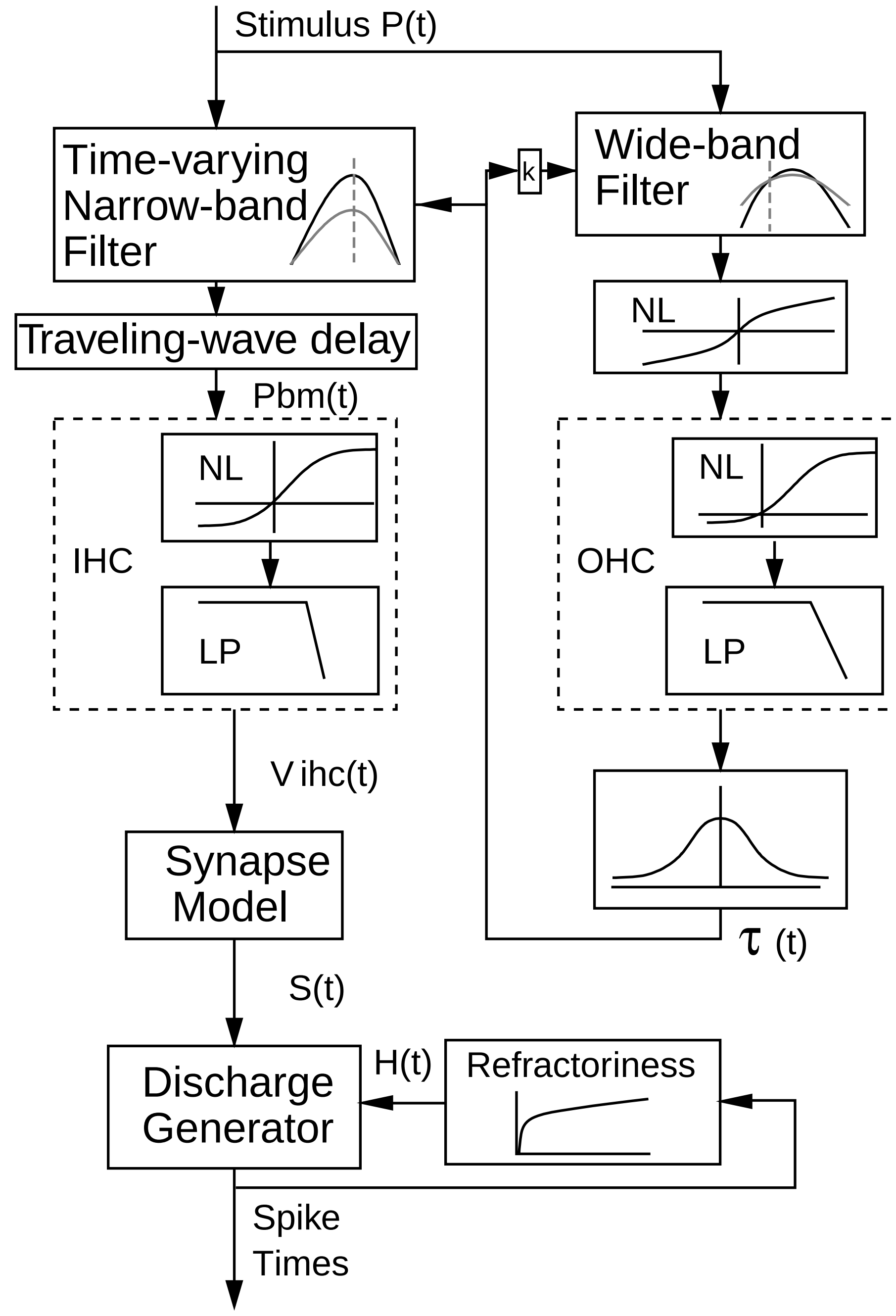
BF ≈ F3

The OHC-impaired model fiber with BF near F3 also exhibits synchrony capture to the 2nd formant of the modified vowel, instead of to F3 (*cf.* Panels 5C&D). The same effect is seen in the AN fiber data.



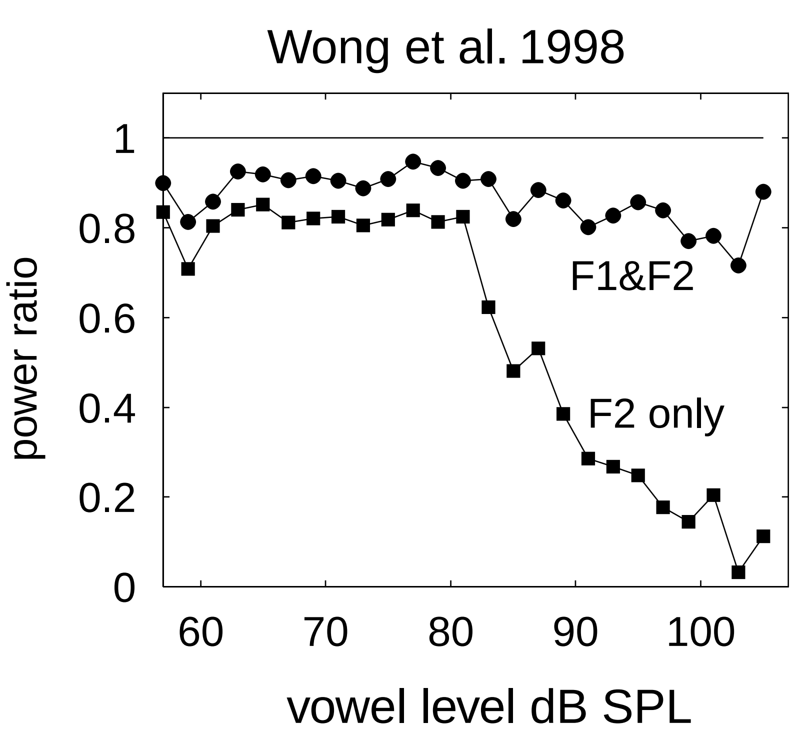
7 Including a basilar membrane wide-band nonlinearity better predicts synchrony in response to the vowel /eh/

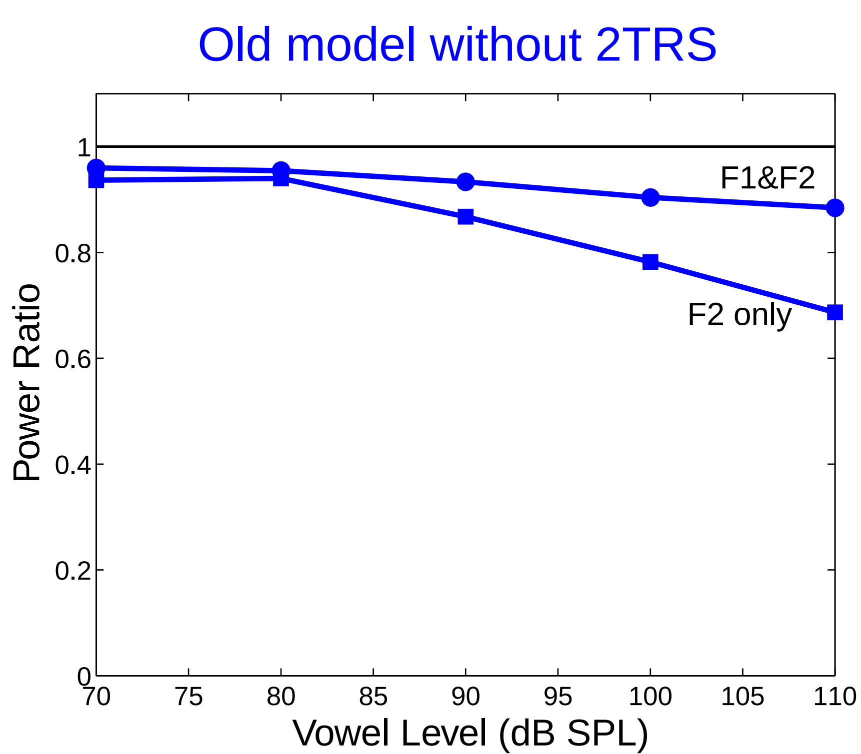
The Carney (1993) model (Panel 1) does not include any wide-band nonlinearities and is therefore unable to predict effects such as two tone rate suppression (2TRS; Sachs and Abbas, 1976). This is because the BM filter bandwidth and gain are controlled by a feedback mechanism, such that stimuli with little energy passing through the BM filter cannot affect the filter behavior. In Bruce et al. (1999) we postulated that this is also the explanation of why the model does not predict loss of synchrony capture by the vowel formant closest to the fiber's BF at very high stimulus intensities (Wong et al., 1998). This is important because of the high stimulus intensities used to compensate for elevated thresholds in the case of a hearing loss. Zhang et al. (2000) have developed a model, shown to the right, that does include a wide-band nonlinearity and consequently predict 2TRS. This is achieved by having the OHC control of the BM filter use a filter of its own that is wider than the BM filter. Consequently, stimuli that have little energy passing through the narrow-band BM filter can still modify the behavior of the filter, if they have sufficient energy passing through the wide-band OHC filter. This is illustrated below for synchronized responses to the vowel /eh/.

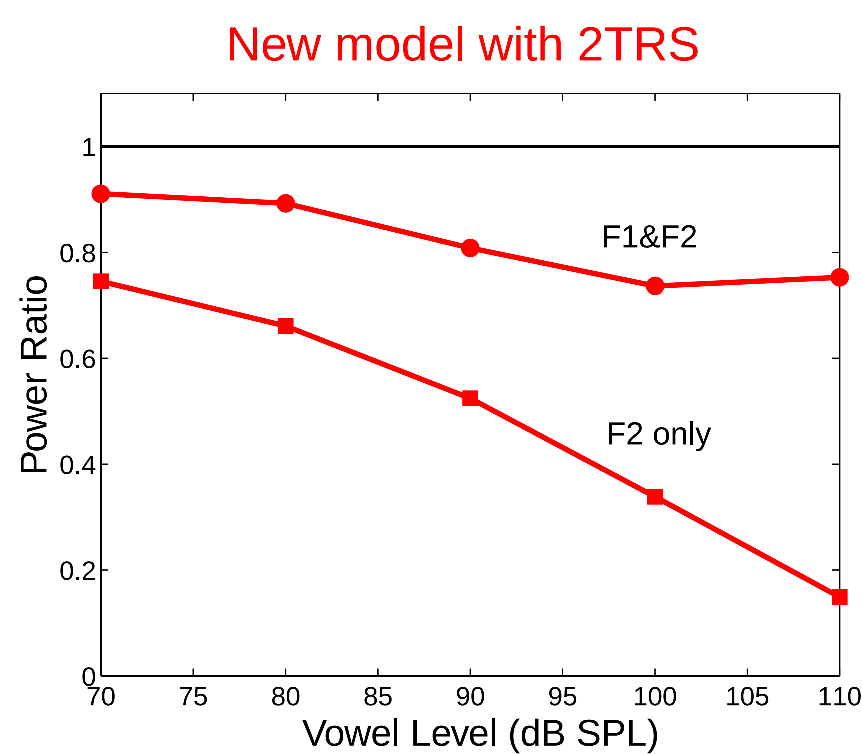


Power ratios can be used to quantify the synchronized response to vowel formants. The power ratio for F2, for example, is the fraction of the total synchronized rate that is contained in the F2-related harmonics.

Both the old and the new model correctly predict that most of the synchrony, even at very high intensities, is contained in the F1&F2-related harmonics combined. However, the old model is unable to predict the loss of synchrony capture to F2 at very high intensities. In contrast, the new is able to predict the reduction in the F2 power ratio to almost zero at very high intensities, although it predicts too much suppression of F2 by F1 at moderate sound levels.







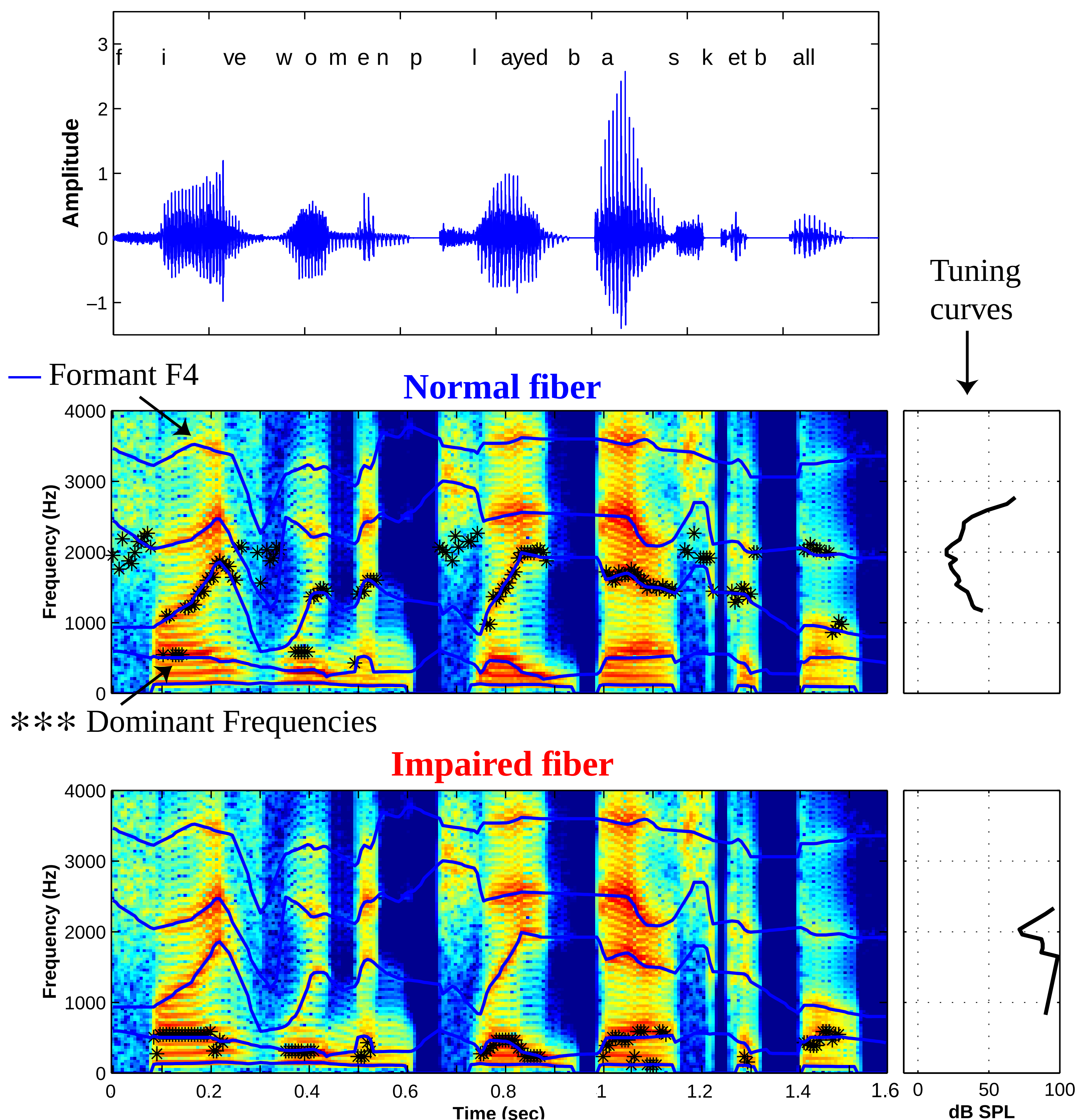
8 Dominant frequency analysis for running speech

Ji et al. (2000) have studied how the synchronized rate tracks formant trajectories in running speech. Fibers' responses to 50+ presentations of the sentence "Five women played basketball" were recorded in normal and impaired cats. The sentence was resynthesized from real speech so that formant trajectories were well-defined. Synchronized rates were computed for short segments (25.6 ms Hanning window) of the speech stream. Within each time window for which the spike rate was at least three standard deviations above the spontaneous rate, a dominant frequency was selected as the frequency component with the highest synchronized rate that was also two standard deviations above the mean synchronized rate across all frequencies (Delgutte and Kiang, 1984).

The sentence "Five women played basketball" was chosen for its large amount of vowel content and its commercial availability (Sensimetrics Corp., Somerville, MA). The figure to the right shows the waveform of the stimulus. The figures below show the spectrogram of the stimulus; the formants F0 to F4 are indicated by the blue lines.

The dominant frequency for each speech segment is plotted as an asterisk on top of the speech spectrogram. In normal fibers with BFs around 2 kHz, the dominant frequency components predominantly trace F2 and F3, depending on which formant contains higher energy and is closer to BF.

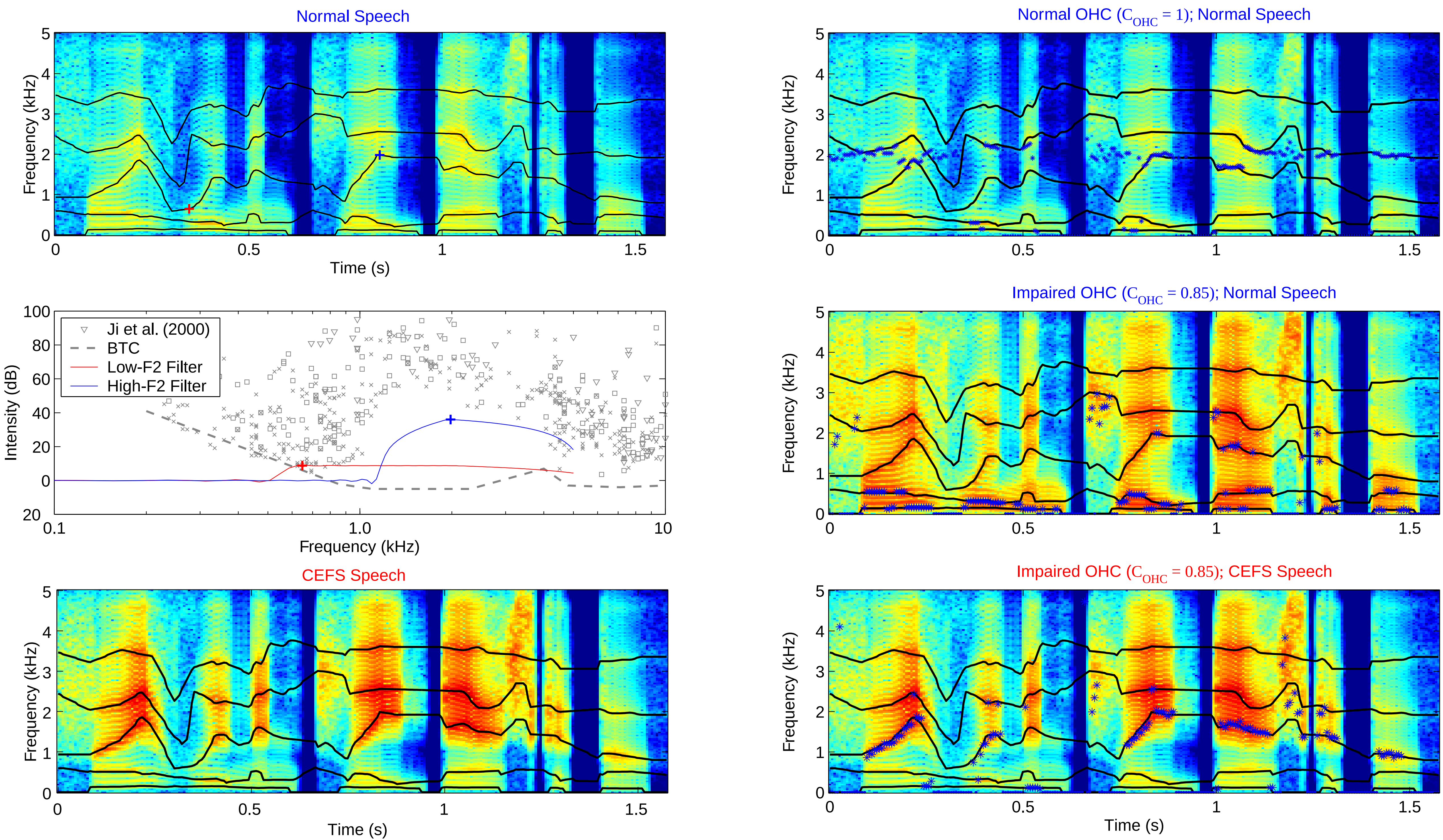
In impaired fibers with similar BFs, dominant frequencies do not exist for as many segments because of broadband synchrony. When they do occur, they trace multiple formants, often including F0 and F1 because of those formants' high energy content.



9 Contrast-enhancing frequency shaping of running speech

CEFS filtering can be generalized to running speech by using a time-varying FIR filter, where the frequency response (see the red and blue lines in the middle figure for two examples) is updated each sample as a function of F2 at that time (the red and blue pluses in the top figure) and half the hearing loss at the F2 frequency (red and blue pluses in the middle figure). The resulting speech spectrogram is show in the bottom figure.

Both the normal (top figure) and the impaired (middle figure) model fibers with BFs at 2 kHz show similar tracking behavior to the physiological data of Panel 8. The bottom panel illustrates that CEFS filtering of the speech is able to restore strong tracking of F2. As for the CEFS-modified vowel (see Panel 6), the representation of F3 may not be recovered with this amplification scheme.



10 Discussion

- The results of this poster show that a simple modification of the OHC controlled BM filter in a model of the auditory periphery is able to describe much of the effect of acoustic trauma on AN response to both vowels and running speech.
- A wide-band nonlinearity producing two tone suppression may also be an important feature of the model because of the high stimulus intensities used in the case of hearing loss.
- Unlike conventional amplification schemes, contrast-enhancing frequency shaping restores a strong representation of F2 in the AN response for both vowels and running speech, without producing anomalous responses at non-formant frequencies.
- The CEFS filtering scheme for running speech requires a knowledge of how F2 changes over time. We are currently developing and testing an algorithm for reliably tracking the formants of a speech signal, even in a moderately noisy environment.
- The physiological and modeling studies to date have concentrated on responses to static formants (i.e., vowels) and formant transitions in running speech. We are now investigating how consonants and consonant-vowel transitions are coded in the AN and how that coding is disrupted by hearing loss. The results of this study should suggest an amplification scheme for restoring correct AN responses to consonants, which in combination with the CEFS algorithm will provide a physiolgically based processing scheme for all segments of a speech stimulus that could be implemented in a digital hearing aid.
- The perceptual effects of these amplification schemes remain to be investigated. Such testing should be done in long-term studies to allow the central nervous system to adjust to the restored peripheral representation.

11 Acknowledgements

The authors would like to thank Laurel Carney for providing code for her 1993 auditory-periphery model and for collaboration (along with Xuedong Zhang and Michael Heinz) on the new model; Teng Ji for providing the running speech physiological data; and Phyllis Taylor for technical assistance. This research was supported by NIDCD grants DC00109 and DC00023.

12 References

Bruce IC, Young ED, Sachs MB (1999) Modification of an auditory-periphery model to describe the effects of acoustic trauma on auditory nerve response. Abstr. 22nd ARO Midwinter Meeting.

Carney LH (1993) A model for the responses of low-frequency auditory-nerve fibers in cat. J. Acoust. Soc. Am. 93:401–417.

Delgutte B, Kiang NYS (1984) Speech coding in the auditory-nerve: I. Vowel-like sounds. J. Acoust. Soc. Am. 75(3):866–878.

Harrison RV (1981) Rate-versus-intensity functions and related AP responses in normal and pathological guinea pig and human cochleas. J. Acoust. Soc. Am. 70:1036–1044.

Ji T, Bruce IC, Ngan E, May BJ, Sachs MB, Young ED (2000) Representation of a dynamic speech stream in cats with sensorineural hearing loss. Abstr. 23rd ARO Midwinter Meeting.

Liberman MC, Dodds LW (1984) Single-neuron labeling and chronic cochlear pathology: III. Stereocilia damage and alterations of threshold tuning curves. Hear. Res. 16:55–74.

Miller RL, Schilling JR, Franck KR, Young ED (1997) Effects of acoustic trauma on the representation of the vowel /ε/ in cat auditory nerve fibers. J. Acoust. Soc. Am. 101:3602–3616.

Miller RL, Calhoun BM, Young ED (1999) Contrast enhancement improves the representation of /ε/-like vowels in the hearing-impaired auditory nerve. J.Acoust. Soc. Am. 106(5):2693–2708.

Patuzzi RB, Yates GK, Johnstone BM (1989) Outer hair cell receptor current and sensorineural hearing loss. Hear. Res. 42:47–72.

Sachs MB, Abbas PJ (1974) Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli. J. Acoust. Soc. Am. 56:1835–1847.

Schilling JR, Miller RL, Sachs MB, Young ED (1998) Frequency-shaped amplification changes the neural representation of speech with noise-induced hearing loss. Hear. Res. 117, 57–70.

Wong JC, Miller RL, Calhoun BM, Sachs MB, Young ED (1998) Effects of high sound levels on responses to the vowel /eh/ in cat auditory nerve. Hear. Res. 123:61–77

Zhang X, Heinz MG, Bruce IC, Carney LH (2000) A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. Submitted to J. Acoust. Soc. Am.