Improvements of an Auditory-periphery Model to Describe Auditory Nerve Response to Speech Stimuli

Ian Bruce, Eric Young and Murray Sachs

Center for Hearing Sciences and Department of Biomedical Engineering
The Johns Hopkins University School of Medicine, Baltimore, MD 21205

960

0 Abstract

Previously we investigated modifications of the auditory-periphery model of Carney (1993) to describe auditory nerve responses to the vowel /eh/ in acoustically traumatized cats. The Carney model has a number of limitations for normal fibers, restricting our assessment of these modifications. First, published model parameters for basilar membrane tuning are only suitable for fibers with best frequencies (BFs) below 2 kHz. Second, the synapse section is only designed to describe high spontaneous rate fibers. Third, fiber thresholds do not change systematically with BF, as indicated by best threshold curves (BTCs). Fourth, the model does not describe wide-band nonlinearities such as two-tone suppression. In this presentation, we describe progress made in addressing these four issues. First, we have modified the basilar membrane tuning parameters to produce the correct change in Q10 with BF at least up to 15 kHz. Second, we have tested various nonlinear neurotransmitter release rate functions to model fibers with arbitrary spontaneous rate. Third, the BTC may be largely attributed to middle-ear filtering, and we have therefore added a model of the cat middle ear to the front of the auditory-periphery model. Fourth, we have begun investigating vowel responses obtained with a new model by Carney et al. designed to describe two-tone suppression effects. Preliminary simulation results show that in general these four modifications improve the model's predictions of normal fibers' response characteristics for a range of BFs and intensities and for different spontaneous rate classes. In contrast to the old Carney model but in agreement with experimental data, model fibers with BFs near the third format (2.5 kHz) synchronize well to that frequency and model fibers with BFs near the second format lose synchrony capture to that frequency at high stimulus intensities because of the wide-band nonlinear effects. These enhancements will improve our assessment of the model adjustments to describe acoustic trauma and may suggest further modifications.

1 Improved basilar membrane tuning parameters produce better prediction of synchrony to F3 of the vowel /eh/

Published model parameters in Carney (1993) for basilar membrane tuning are only suitable for fibers with best frequencies (BFs) below 2 kHz. We have modified the basilar membrane tuning parameters to produce the correct change in Q10 with BF at least up to 15 kHz. In contrast to the old Carney model but in agreement with experimental data, model fibers with BFs near the third format (2.5 kHz) of the synthetic vowel /eh/ synchronize well to that frequency.
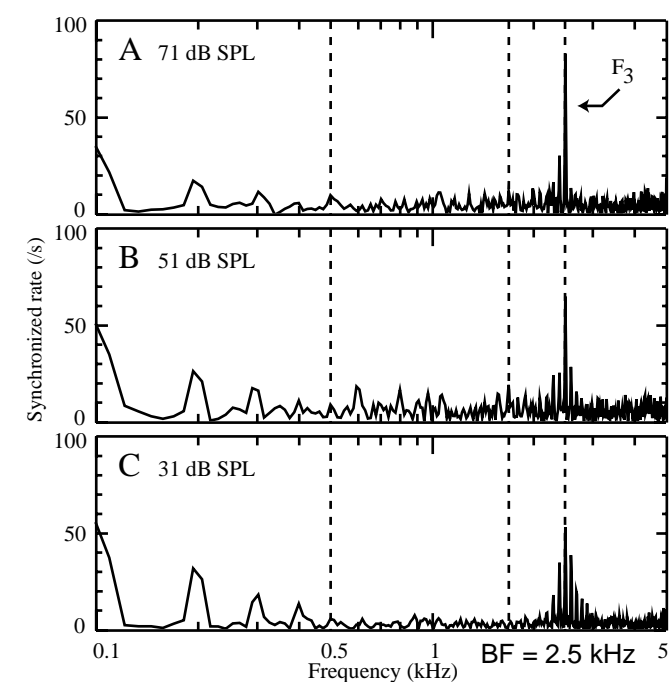
Model parameters and predictions
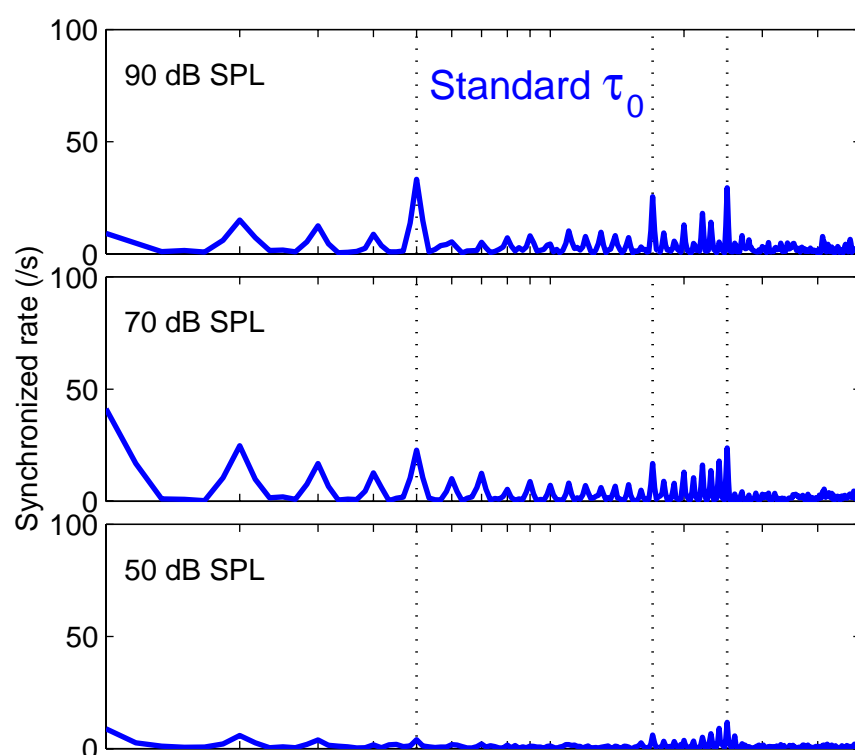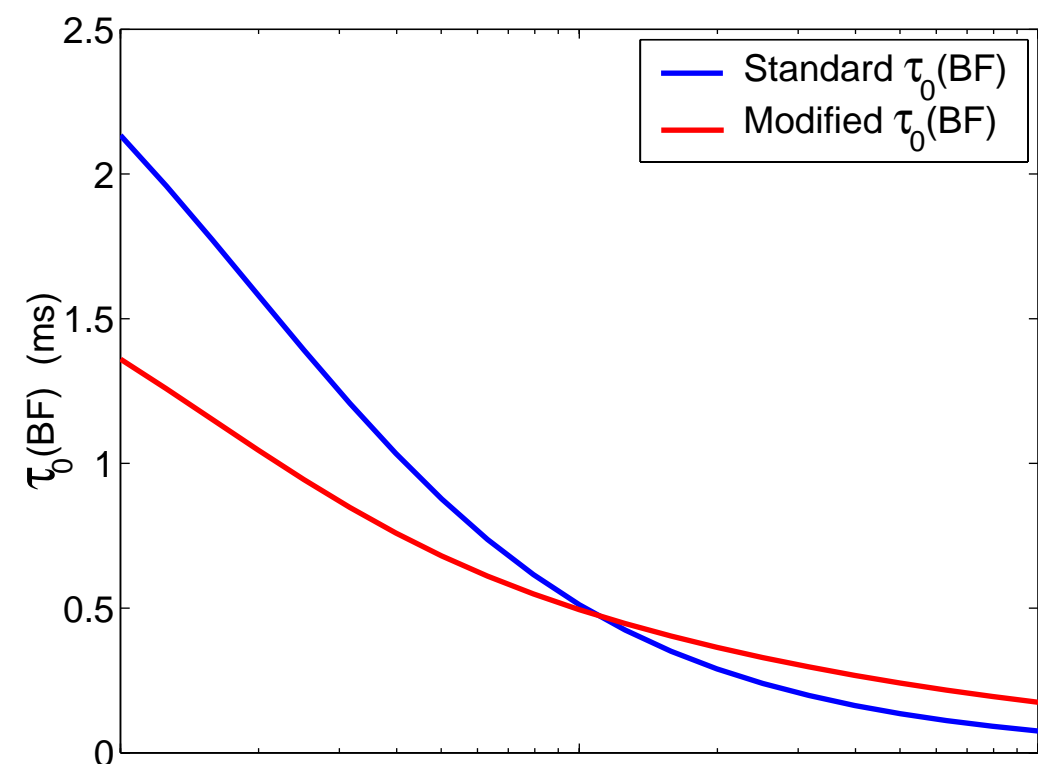
Power spectrum of the synthetic vowel /eh/
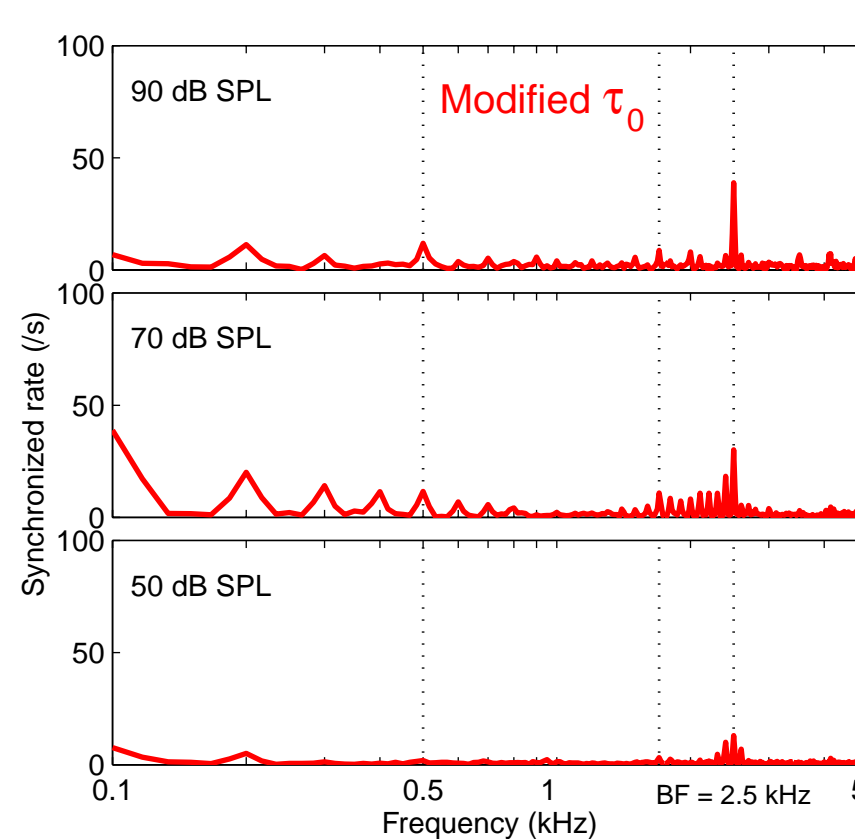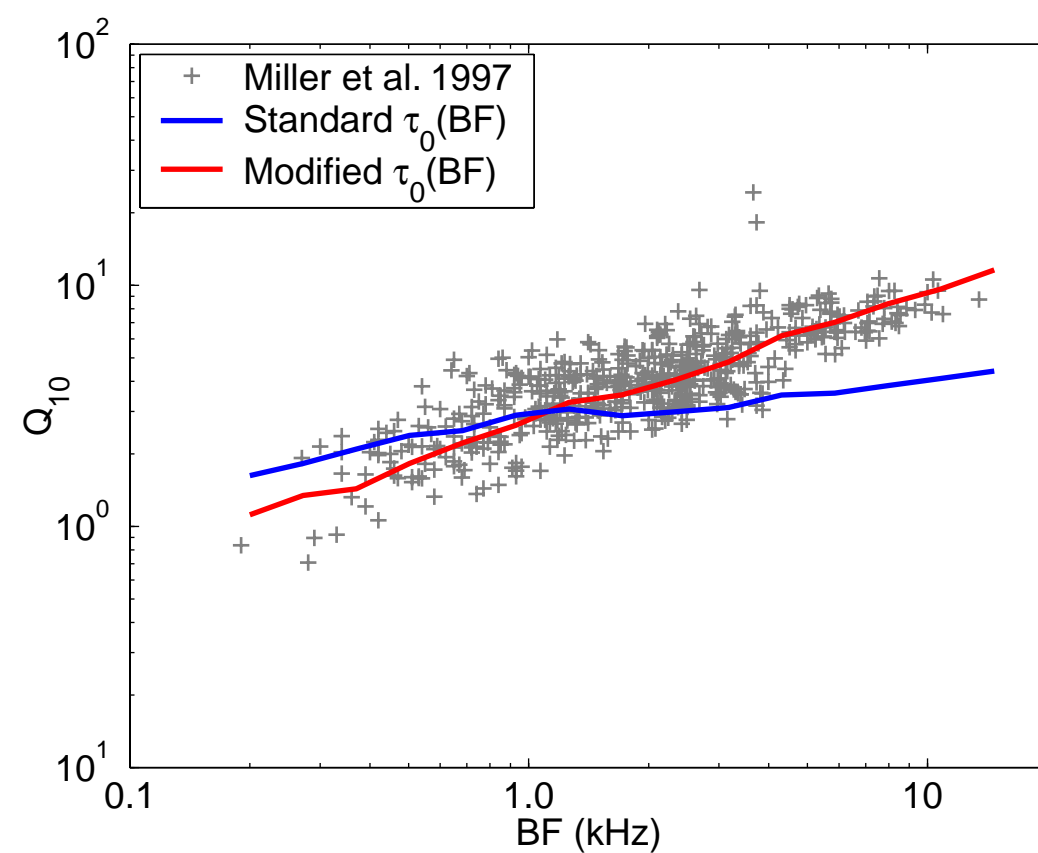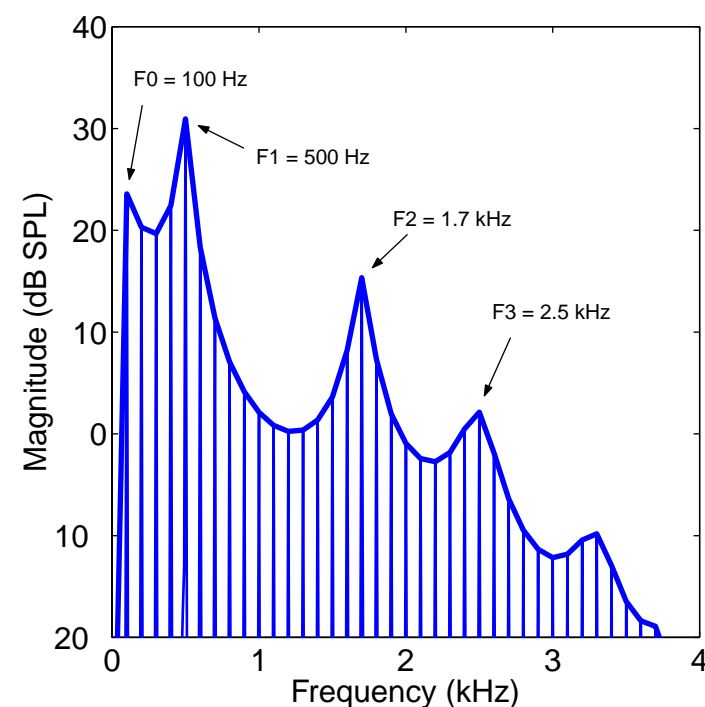
Single fiber data from Miller et al. 1997

2 Modification of the synapse model for an arbitrary spontaneous rate and the correlated threshold

The synapse model parameters given in Carney (1993) only describe high spontaneous rate fibers. The synapse model used by Carney is that of Westerman and Smith (1988), which is a phenomenological description involving three pools of "neurotransmitter." Diffusion of the neurotransmitter between the pools is mediated by three permeabilities, the final of which is dependent on the inner hair cell (IHC) potential. This system can be described by a pair of coupled linear equations. The utility of the Westerman and Smith model in comparison to other synapses models is that these equations have an analytical solution for a step response, and it is therefore possible to calculate model parameters to achieve a desired arbitrary spontaneous rate, as well as maximum discharge rate, adaptation time constants and ratio of onset to adapted amplitude.
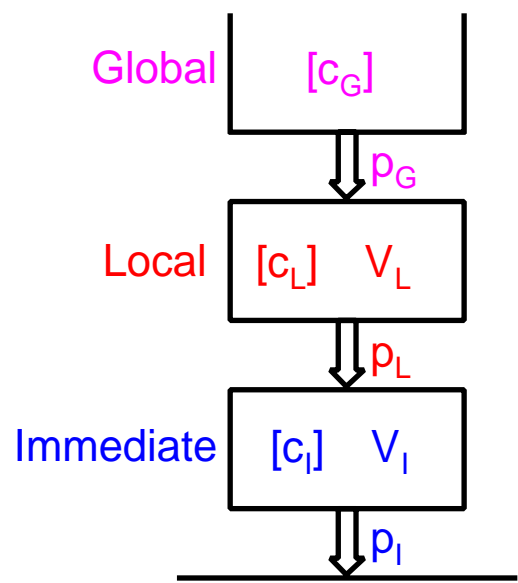
$$V_I \frac{dc_I}{dt} = -p_I(t) c_I(t) + p_L [c_L(t) - c_I(t)], \quad V_L \frac{dc_L}{dt} = -p_L [c_L(t) - c_I(t)] + p_G [c_G - c_L(t)]$$

IHC drive to the synapse is produced by varying $p_I(t)$ as some function of the IHC potential $V_{\text{IHC}}(t)$, for example, a rectified-linear function:

$$p_I(t) = p_{I\text{spont}} + \max[0, g_s V_{\text{IHC}}(t)]$$

To derive model parameters, the response to a tone is approximated by a step in $p_I(t)$ such that the synapse output is a double-exponential that can be fitted to AN data:

$$S(t) \overset{\Delta}{=} p_I(t) c_I(t) = A_r e^{-t/t_r} + A_{st} e^{-t/t_{st}} + A_{ss}$$

Model parameters can be derived from the amplitude and time-constant parameters of a fitted double-exponential. However, Westerman and Smith pooled fibers with different BFs and spontaneous rates. In collaboration with Zhang, Heinz and Carney, we have developed a method of deriving model parameters that give the correct (i) adaptation, (ii) threshold and (iii) rate and synchrony versus level behavior as a function of spontaneous rate and BF. These are each described in the next three panels respectively.

3 Altering the synapse model parameters allows both low and high spontaneous rate fiber adaptation behavior

Physiological data (e.g., Rhode and Smith, 1985) show that low spontaneous rate fibers show no adaptation even at high stimulus level, whereas high spontaneous rate fibers show increasing adaptation as the stimulus level is increased. The parameters from the analytical solution are set such that the ratio of onset amplitude to adapted amplitude, at an arbitrary reference level of the immediate permeability p_{Iref}, increases from 1 for a spontaneous rate of 0.1 spikes/s to 6 for a spontaneous rate of 100 spikes/s. Given the desired adapted discharge rate at that reference level, the adaptation time constants and the spontaneous rate, it is possible to derive required model volumes, concentrations and permeabilities.

For the low spontaneous rate fibers, the lack of adaptation is achieved by having small concentrations but relatively large permeabilities and volumes, with the permeabilities decreasing in magnitude from the global pool to the immediate pool, such that the pools are not depleted even at high intensities.

In contrast, for the high spontaneous rate fibers the concentrations are high and the permeabilities and volumes relatively low, with the permeabilities increasing in magnitude from the global pool to the immediate pool. Consequently, the immediate and local pools are depleted at higher intensities, giving rise to the marked adaptation behavior.
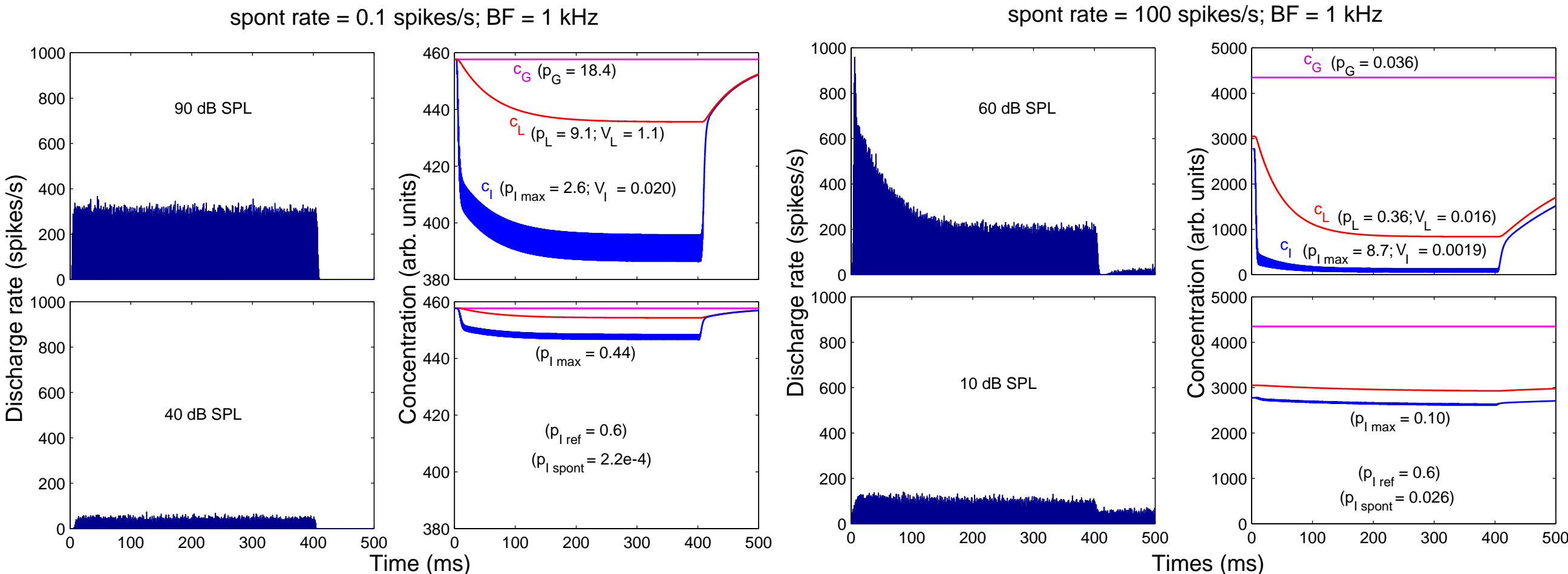

While this model is only a phenomenological description of the hair cell/auditory nerve synapse, the model parameters needed to model the different spontaneous rate classes may shed light on the correspondingly different anatomy seen in the synapse (e.g., Merchan-Perez and Liberman, 1996). For example, both the size and complexity of the synaptic body in the IHC and the number of presynaptic vesicles is much greater for low spontaneous rate fibers. From our modeling results, we hypothesize that this specialization may be in part to prevent neurotransmitter depletion at high intensities.

4 Modifying the synapse gain gives the correct behavior of threshold as a function of spontaneous rate and BF

Previous attempts to model the correct relationship between threshold and spontaneous rate have used a nonlinear function to relate the IHC potential to the synapse's immediate permeability, such as saturating-exponential (Schoonhoven et al., 1997) or quadratic (e.g., Geisler, 1990; Robert and Eriksson, 1999) functions. However, only the effects on threshold and rate-level functions, but not on synchrony, were investigated in these studies, and the effects of the IHC lowpass filter were not considered. Our simulations have shown that saturating functions produce clipping of the IHC potential at high intensities such that the synchronization behavior is dramatically degraded, an effect not seen in the physiological data. Furthermore, (i) the adaptation behavior described in the previous panel is a more accurate model of saturation effects, and (ii) the nonlinear relationship between the immediate permeability and the discharge rate in the Westerman and Smith model may be sufficient to explain the nonlinear effects in rate-level functions described by others. Therefore, we are presently utilizing a rectified-linear function of the form described in the previous panel to relate the IHC potential to the immediate permeability. We may add a second-order term if required, however this must be done with caution: when combined with our method for setting parameters described in the previous panel, overly high immediate permeabilities produce an instability in the high spontaneous rate synapse, even at very high sampling rates, due to over-depletion of the immediate pool.

We need then to know how this rectified-linear function should change with spontaneous rate. The constant term $p_{I\text{spont}}$ sets the spontaneous rate, and the slope g_S controls the "synapse gain" and consequently threshold. Our analysis of data collected by Liberman (1978) shows that thresholds should increase by approximately 10 dB per decade reduction in the spontaneous rate.

Additionally, the combination of the IHC nonlinearity and lowpass filter produces a drive to the synapse that has a dominant AC component below 1 kHz and a dominant DC component above approximately 6 kHz, with a transition region in between. With an IHC nonlinearity like the Boltzman function, the maximum DC component is smaller in magnitude than the maximum AC component, such that thresholds increase with BF (blue line on the plot to the right). This is not seen in the physiological data, and therefore it must be compensated for in some way. We are investigating if this compensation may be achieved by increasing the synapse gain g_S with increasing BF. Plotted to the left are curves for the synapse gain versus BF and spontaneous rate that produce no variation in threshold with BF, thresholds for a spontaneous rate of 100 spikes/s at 0 dB SPL and the desired increase in threshold with reducing spontaneous rate (red lines on the plot to the right). The divergence of the 100 spikes/s threshold curve from the BTC is explored on panel 6.

MCL92 from Liberman 1978

thrsh = 10.2*log10(spont)+15.5
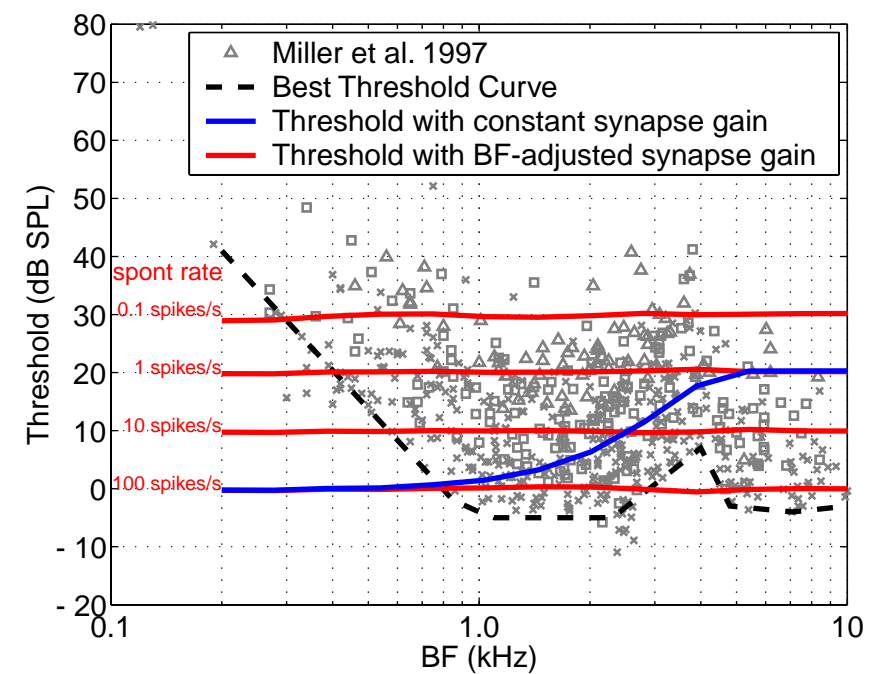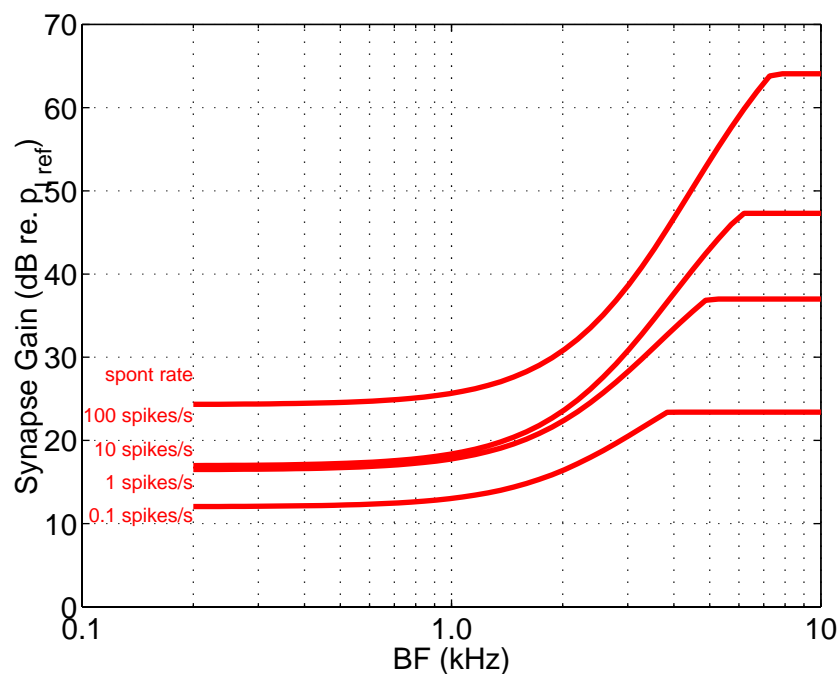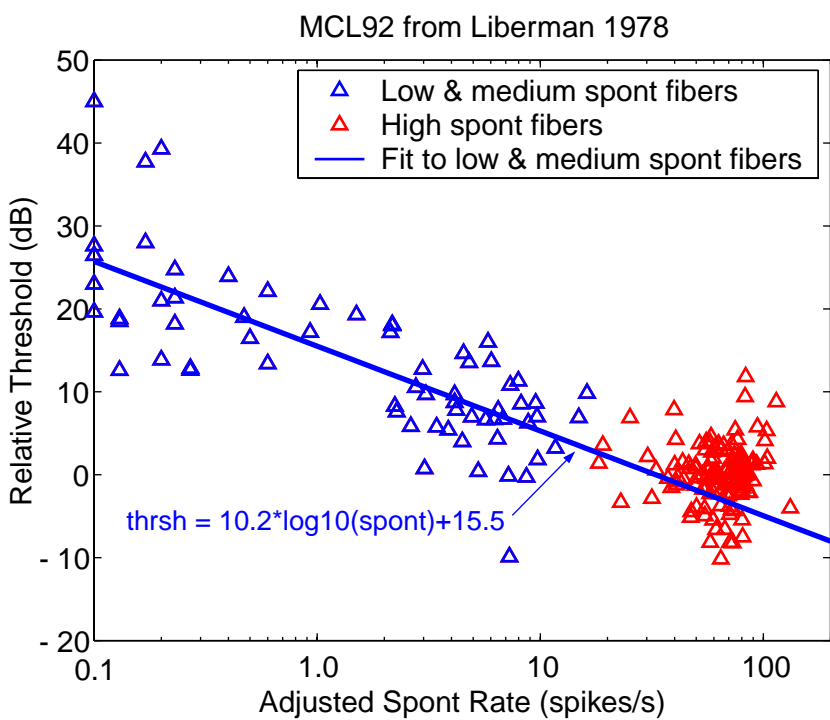
5 How do the new synapse parameters affect discharge rate and synchrony versus stimulus level?

The plot below shows rate-level functions for spontaneous rates from 0.032 spikes/s to 100 spikes/s. Consistent with physiological data (e.g., Sachs and Abbas, 1974), high spontaneous rate fibers saturate at high intensities, whereas low spontaneous rate fibers exhibit what has been termed "sloping saturation."

Note that the maximum discharge rate for the low spontaneous rate fibers may be adjusted to equal that of the high spontaneous rate fibers if desired.

Saturation discharge rates for high spontaneous rate fibers increase with increasing BF up to 1 kHz, in qualitative agreement with physiological data (Liberman, 1978), but plateau above 1 kHz instead of continuing to rise. This further dependence on BF may be included in the synapse parameters if desired.

Below is plotted synchronization index (SI) versus stimulus intensities for fibers with five different BFs. In all cases the SI increases with increasing intensity until it reaches a maximum and then falls to a plateau, as seen in the physiological data (e.g., Johnson, 1980). However, the plateaus for BFs above 2.0 kHz are much lower than those of the physiological data. This is due to transition of the IHC drive to the synapse from the AC component of the IHC potential to the DC component, as discussed in panel 4.
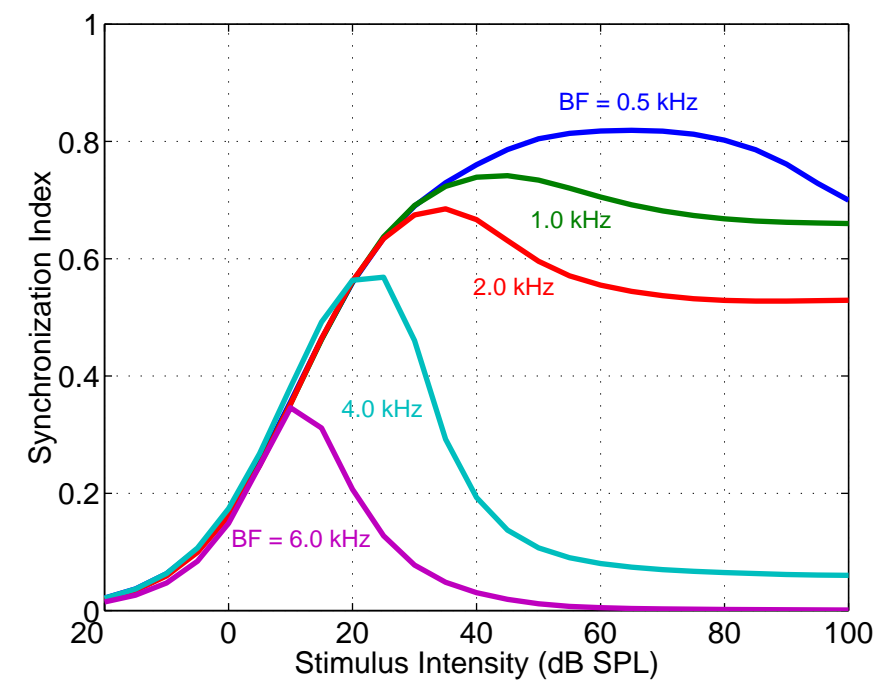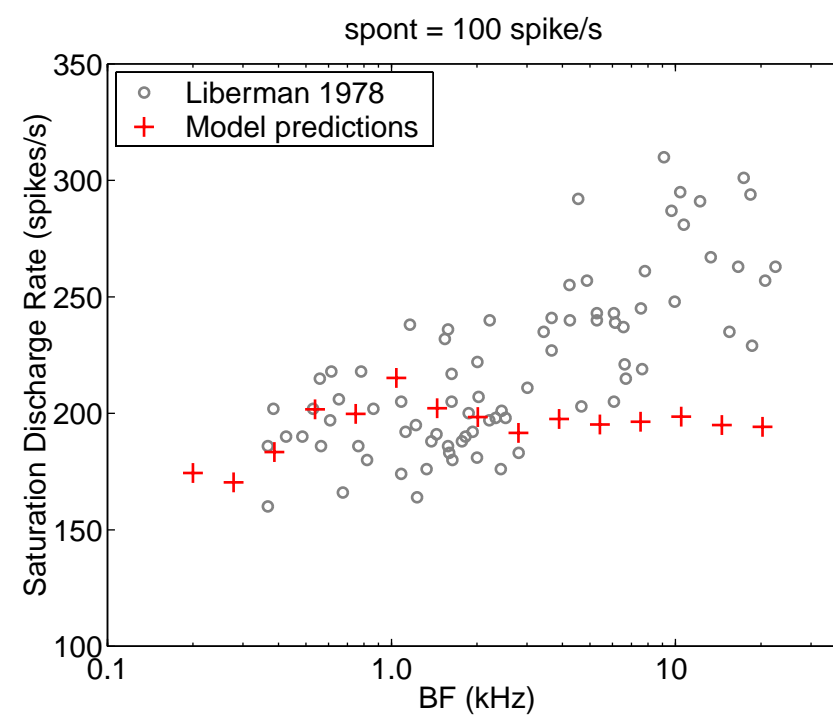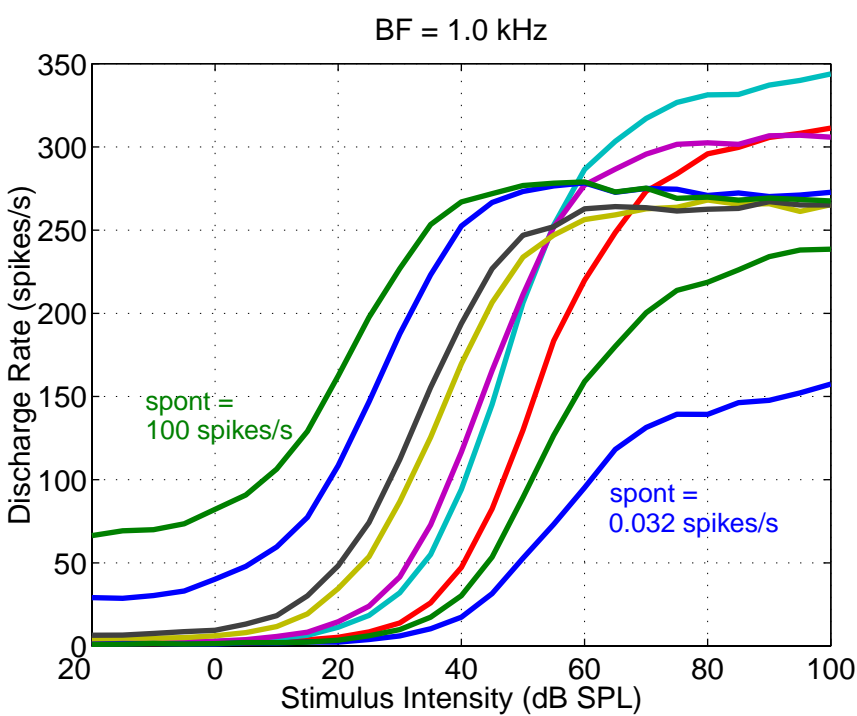
6 Including a middle ear model predicts much of the best threshold curve

Middle-ear filtering may have a significant effect on auditory nerve responses to broadband stimuli such as speech. However, the Carney model does not include a middle ear section. We have therefore incorporated a middle-ear model by combining the middle-ear cavities model of Guinan and Peake (1967) with the middle-ear model of Matthews (1983). We have found the transfer function representation of the model and implemented it as a 12th-order IIR digital filter. The effects of middle-ear filtering on threshold versus BF is shown below to the right. The middle ear model accurately produces the "4 kHz notch" as well as approximately half of the increase in threshold as BF decreases below 1 kHz. The remaining threshold shift at low BFs may be due to reduced basilar membrane, IHC and/or synapse gain in those fibers. Further investigation of this effect is needed to ascertain the cause.
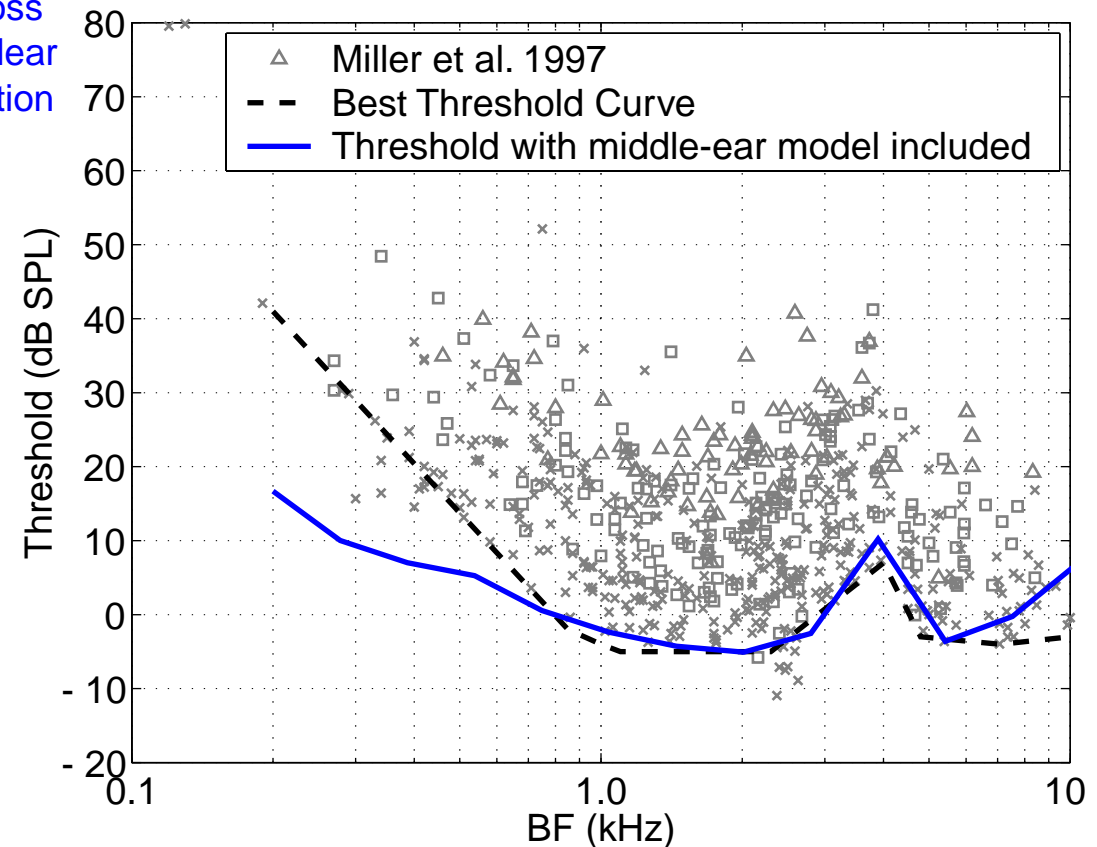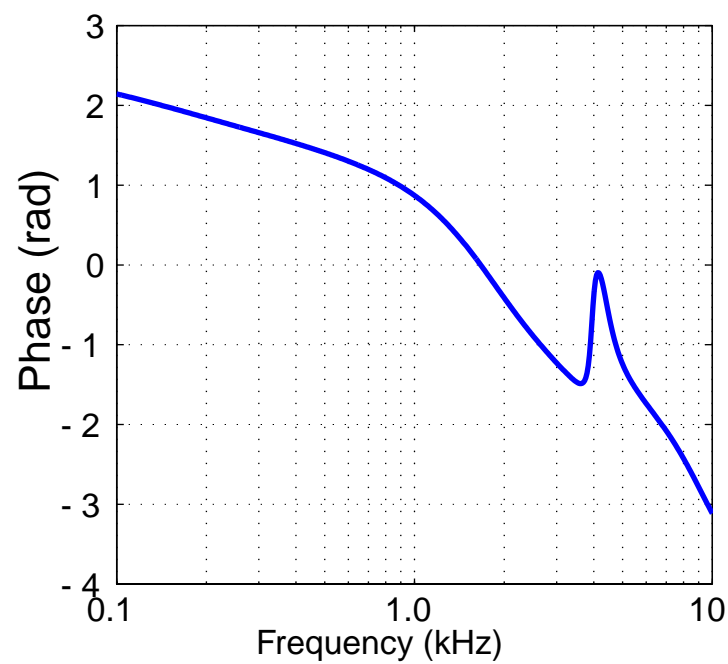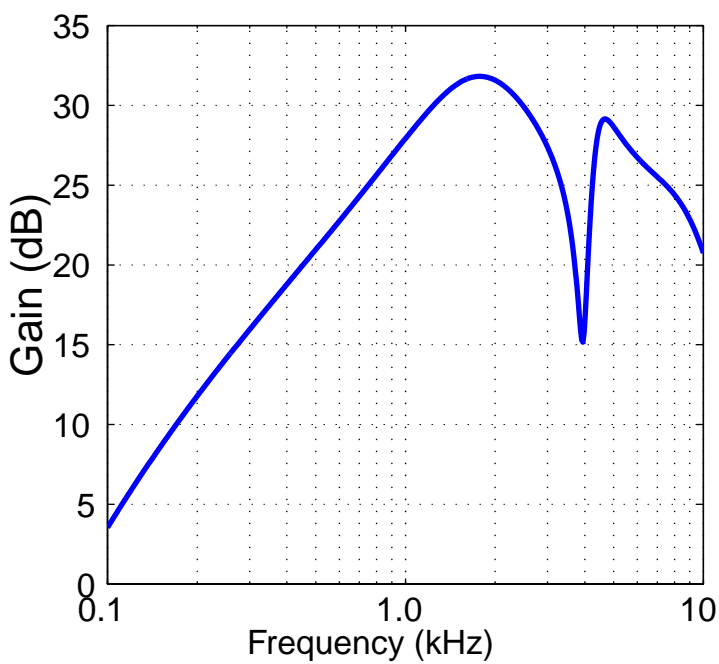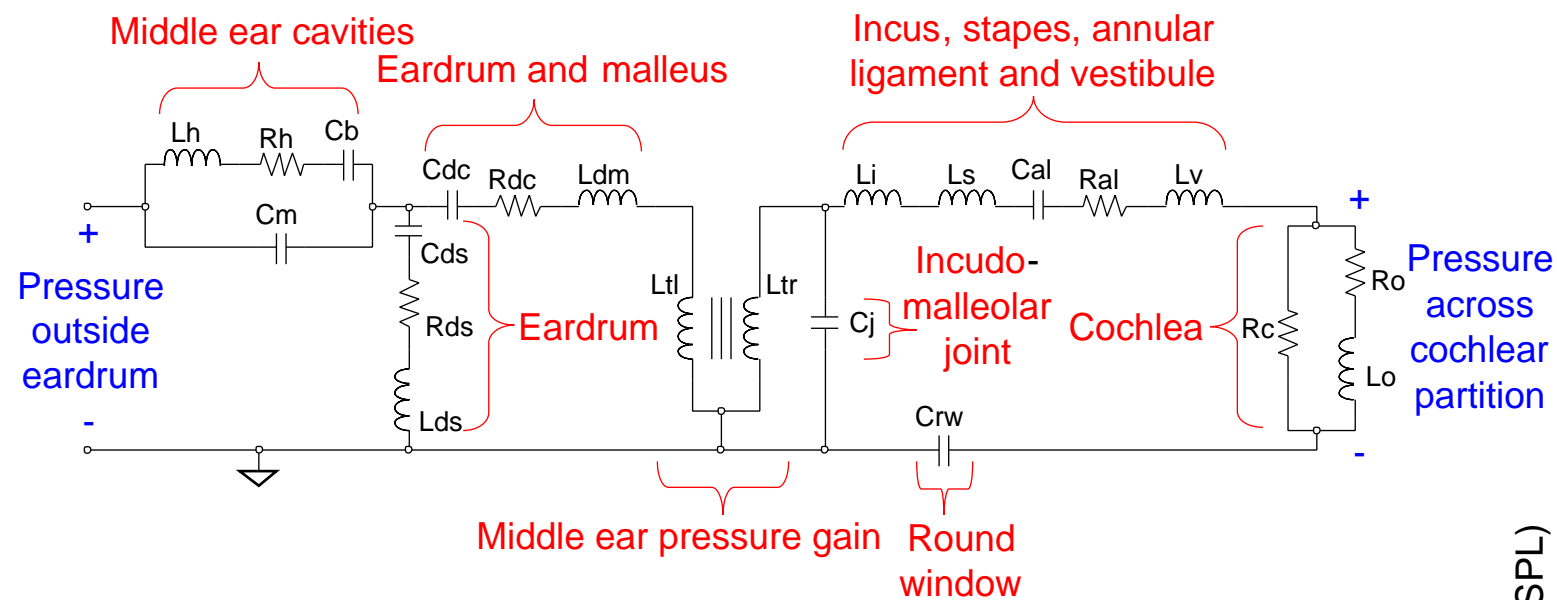
7 Including a basilar membrane wide-band nonlinearity better predicts synchrony in response to the vowel /eh/

The Carney (1993) model does not include any wide-band nonlinearities and is therefore unable to predict effects such as two tone rate suppression (2TRS). 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. Zhang, Heinz and Carney (1999) have developed a model that does include a wide-band nonlinearity and that does consequently predict 2TRS.
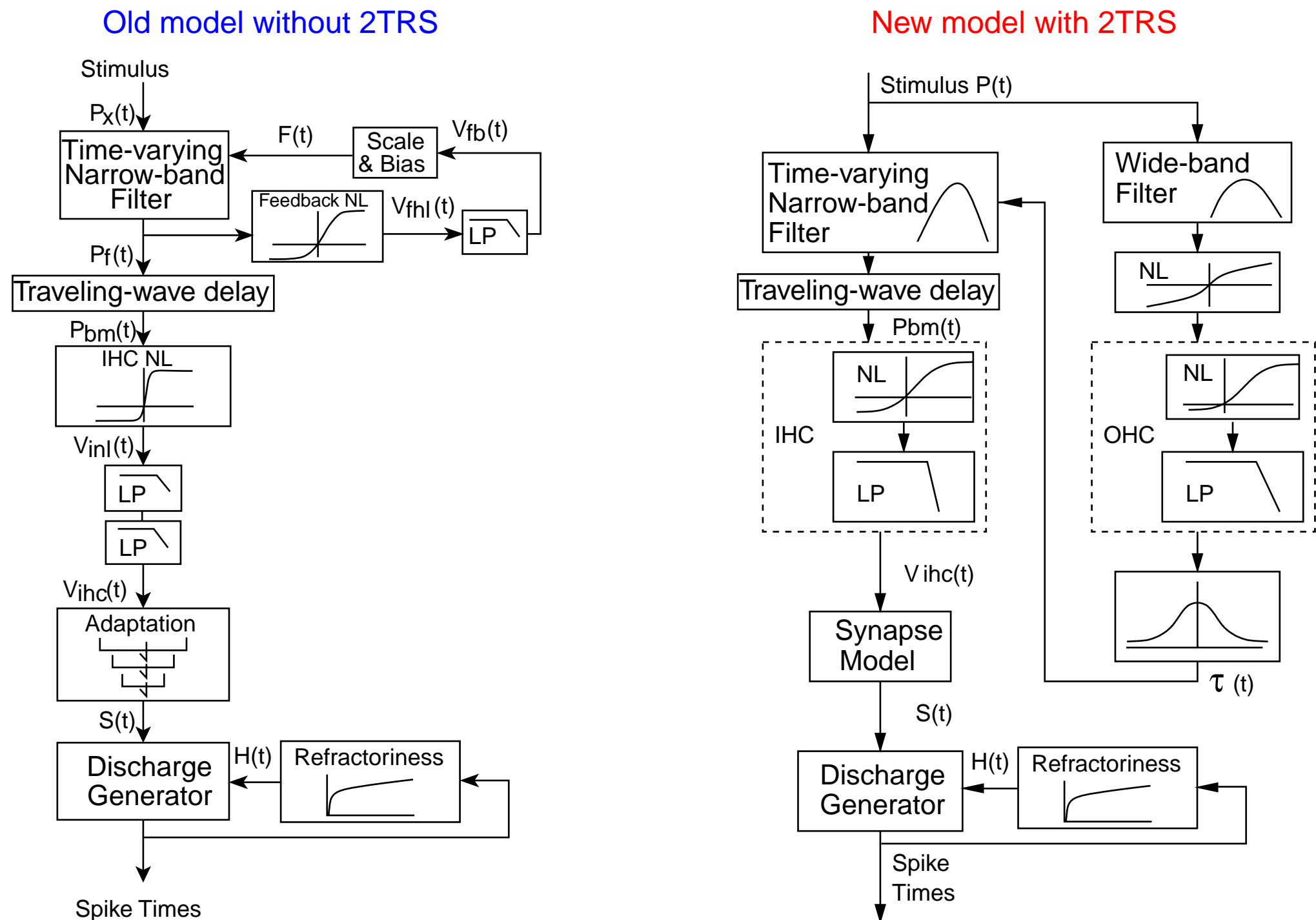
Power ratios can be used to quantify the synchronized response to vowel formants. The power ratio for F2, for example, is the fraction of the total synchronized rate that is contained in the F2-related harmonics.

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.
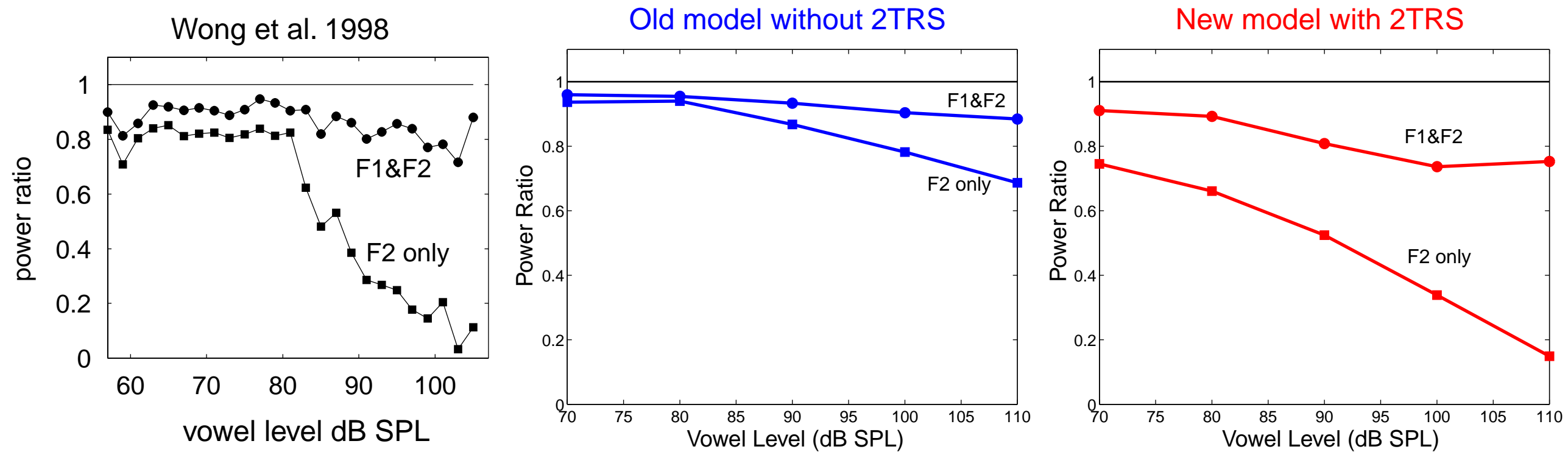
8 Discussion

The modifications to the Carney (1993) model described in this poster extend its utility to BFs above 2 kHz and to arbitrary spontaneous rates. They also include improved descriptions of synapse adaptation, discharge rate saturation, middle-ear filtering and wide-band nonlinearities in the basilar membrane. In all the examples shown, these modifications produce improved predictions of auditory nerve response. Future research directions include the following:

- In collaboration with Zhang, Heinz and Carney, we are investigating further refinements of the IHC nonlinearity and the "synapse gain" functions that may help improve the model's prediction of rate and synchrony in high BF fibers in response to high intensity pure-tones.

- We have begun to study more rate-level and synchrony-level characteristics for vowel stimuli. This will provide an important test for the model features that are primarily developed to give correct responses to non-speech stimuli such as pure tones, tone complexes, noise and clicks.

- Raised thresholds at BFs below 1 kHz are only partially explained by middle-ear filtering. Further investigation is required to explain the remaining elevation of thresholds.

- The basilar membrane wide-band nonlinearity developed by Zhang, Heinz and Carney (1999) should be compared with alternative approaches such as that of Robert and Eriksson (1999), i.e., lateral feedback control in a bank of nonlinear BM filters.

- We are presently adapting the method of Bruce, Young and Sachs (1999) used to describe the effects of acoustic trauma on the Carney model to produce similar effects with the new basilar membrane model. With our improved description of normal fibers, we will then be able to better assess how well we are able to model the effects of acoustic trauma on auditory nerve response to speech stimuli.

9 Acknowledgements

10 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.

Geisler CD (1990) Evidence for expansive power functions in the generation of the discharges of 'low- and medium-spontaneous' auditory nerve fibers. Hear. Res. 44:1–12.

Guinan JJ, Peake WT (1967) A circuit model for the cat's middle ear. Res. Lab. Elect., Q. Prog. Rep. 84:320–326.

Johnson DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J. Acoust. Soc. Am. 68:1115–1122.

Liberman MC (1978) Auditory-nerve response from cats raised in a low-noise chamber. J. Acoust. Soc. Am. 63:442–455.

Matthews JW (1983) Modeling reverse middle ear transmission of acoustic distortion signals. In: de Boer E and Viergever MA (eds) Mechanics of Hearing: Proceedings of the IUTAM/ICA Symposium. Delft, Delft University Press, pp. 11–18.

Merchan-Perez A, Liberman MC (1996) Ultrastructural differences among afferent synapses on cochlear hair cells: Correlations with spontaneous discharge rate. J. Comp. Neurol. 371:208–221.

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

Rhode WS, Smith PH (1985) Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory nerve fibers. Hear. Res. 18:159–168.

Robert A, Eriksson JL (1999) A composite model of the auditory periphery for simulating responses to complex sounds. J. Acoust. Soc. Am. 106:1852–1864.

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.

Schoonhoven R, Prijs VF, Frijns JHM (1997) Transmitter release in inner hair cell synapses: a model analysis of spontaneous and driven rate properties of cochlear nerve fibres. Hear. Res. 113:247–260.

Westerman LA, Smith RL (1988) A diffusion model of the transient response of the cochlear inner hair cell synapse. J. Acoust. Soc. Am. 83:2266–2276.

Zhang X, Heinz MG, Carney LH (1999) Nonlinear compression in an auditory-nerve model. Abstr. EMBS-BMES Joint Conference, Atlanta, GA.