19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

COMPUTATIONAL MODELLING OF THE CAT AUDITORY PERIPHERY: RECENT DEVELOPMENTS AND FUTURE DIRECTIONS

PACS: 43.64.Bt

Bruce, Ian C.[1]; Zilany, Muhammad S. A.[2]
[1]McMaster University; 1280 Main St W, Hamilton, ON, L8S 4K1, Canada; ibruce@ieee.org
[2]McMaster Univ.; 1280 Main St. W., Hamilton, ON, L8S 4K1, Canada; msazilany@gmail.com

ABSTRACT
After more than twenty years of development by many different research groups, much progress has been made in capturing the fundamental properties of cochlear processing in composite computational models of the mammalian auditory periphery to explain auditory nerve responses to a range of acoustic stimuli. In this paper we review recent developments in modelling the cat auditory periphery in particular. Several important cochlear nonlinearities, such as compression and suppression, the shift in tuning with sound pressure level, and the component-1/component-2 transition at very high sound pressure levels, have been incorporated into the latest models. Examination of the latter two properties in the model provides some interesting insights into how cochlear filtering and transduction may be functioning—as well as raising many more questions. In addition, we discuss the remaining inaccuracies of these models and possible approaches to correcting these problems.

INTRODUCTION
Since physiological investigations of auditory-nerve (AN) response properties began more than 40 years ago, efforts quickly followed to create mathematical and computational models to explain the experimental results. While some of these models were designed to explain a specific set of data for a particular stimulus paradigm, interesting physiological data for speech and noise stimuli motivated the development in the 1980s of computational models that could be applied to arbitrary stimuli. Many of these early models utilised linear basilar membrane (BM) models, but it was soon realised that cochlear nonlinearities were important for AN responses, especially to broadband stimuli such as speech and noise stimuli [1]. A recent review of different approaches to capturing these nonlinearities is provided by Lopez-Poveda [2].

One of the first models of the complete auditory-periphery to tackle incorporating BM nonlinearity was the 1993 model of Carney [3]. A series of models of the cat auditory periphery has subsequently been developed by Laurel H. Carney, students and collaborators over the past 14 years, which has seen continued improvement in describing physiological data recorded from the cat AN. Table I provides a comparison of some of the features of the successive models. The original model of [3] is a phenomenological model that incorporates a feedback signal to control the gain and bandwidth of a gammatone BM filter, such that BM compression and broadening of the BM filter is observed with increasing presentation level of a broadband noise, consistent with the physiological data. However, because of the use of a feedback mechanism to control the filter, changes to the filter gain and bandwidth are only observed for stimulus components falling within the excitatory tuning curve of a model AN fibre. Consequently, Zhang et al. [4] replaced the feedback control path with a feed-forward control path with a wider filter than the BM signal path, such that it is able to produce wide-band suppression effects that are observed for stimuli such as multi-tone complexes and vowels.

Table I.- Comparison of features of some models of the cat auditory periphery.

	Carney (1993)	Zhang et al. (2001)	Bruce et al. (2003)	Tan & Carney (2003)	Zilany & Bruce (2006, 2007)
Middle-ear filtering	no	no	yes	yes	yes
Signal-path filter	gammatone 4^{th}-order	gammatone 4^{th}-order	gammatone 4^{th}-order	chirp filter 20^{th}-order	chirp filter 10^{th}-order
Moderate-level effects:					
i) Compression	yes	yes	yes	yes	yes
ii) Suppression	no	yes	yes	yes	yes
High-level effects:					
i) Broadened tuning	yes	yes	yes	yes	yes
ii) BF-shift with level	no	no	no	yes	yes
iii) C1/C2 transition	no	no	no	no	yes
iv) Peak-splitting	no	no	no	no	yes

Bruce et al. [5] made several further improvements to this model to enable it to more accurately describe the responses of normal and impaired AN fibres to vowel stimuli. One addition is a middle-ear filter, which is important because of its effect on the relative amplitudes of the formants of a vowel. It was also necessary to modify somewhat the architecture of the control path to avoid some undesirable distortion products. This model allows for impairment of outer hair cells (OHCs) and inner hair cells (IHCs), such that it can describe data for low and moderate stimulus sound pressure levels from cats with noise-induced hearing loss. However, the gammatone BM filter used in these three models is unable to describe the frequency glide in the impulse response of the BM that has been observed in BM vibration data and AN recordings. This level-independent frequency glide or "chirp" combines with the level-dependent change in the BM impulse response envelope to produce a shift in the BM filter's peak or "best frequency" (BF) with increasing stimulus level. Tan and Carney [6] replaced the gammatone BM filter with a chirping filter; with careful and systematic placement of the filter poles and zeros, the filter's impulse response has a frequency glide in its instantaneous frequency and a gamma-shaped envelope, consistent with the physiological data. However, this model retained the control-path architecture of Zhang et al. [4]. On the other hand, the model of Bruce et al. [5] is not able to accurately describe high-level effects such as the BF-shift with level, a phase shift in responses at high levels (referred to as the C1/C2 transition) and "peak splitting" in the phase-locked response. This prompted the development of the most recent version of the model by Zilany and Bruce [7, 8].

THE MODEL OF ZILANY & BRUCE (2006, 2007)

A schematic of the model of Zilany and Bruce [7, 8] is given in Fig. 1. The input of the model is the stimulus time-domain waveform and the output is the resulting spike times for an AN fibre with a particular characteristic frequency (CF). After middle-ear filtering, the stimulus passes through three parallel filter paths. The C1 filter path describes the filtering properties of the primary mode of BM vibration, which dominates at low and moderate stimulus intensities. This filter is based on the chirping filter of Tan and Carney [6] and consequently exhibits a frequency shift with stimulus level. The compressive and suppressive BM nonlinearities are generated in this filter by moving its poles and zeros according to the wide-band control path output. The control path models the effects of OHC function, and OHC impairment can be realised by varying a parameter C_{OHC}. The C2 filter describe a passive (linear) mode of BM vibration; this filter is the same as the C1 filter with complete OHC dysfunction ($C_{OHC} = 0.0$). The C1 and C2 filter paths each have their own separate IHC transduction function. The C1 transduction function is a rectifying Boltzmann-like nonlinearity (NL) that can be impaired via adjustment of the parameter C_{IHC}. In contrast, the C2 transduction function is an inverting (INV), non-rectifying

function that is shallower than the C1 function for low to medium sound pressure levels, then steeper for high sound pressure levels, and eventually saturating for extremely high sound pressure levels. Consistent with physiological data, the C2 transduction function is robust to cochlear damage. The outputs of the two IHC transduction functions are summed and low-pass (LP) filtered to produce a model IHC receptor potential. This potential drives an IHC-AN synapse model, which subsequently leads to the generation of spikes according to an inhomogeneous Poisson process with refractoriness.

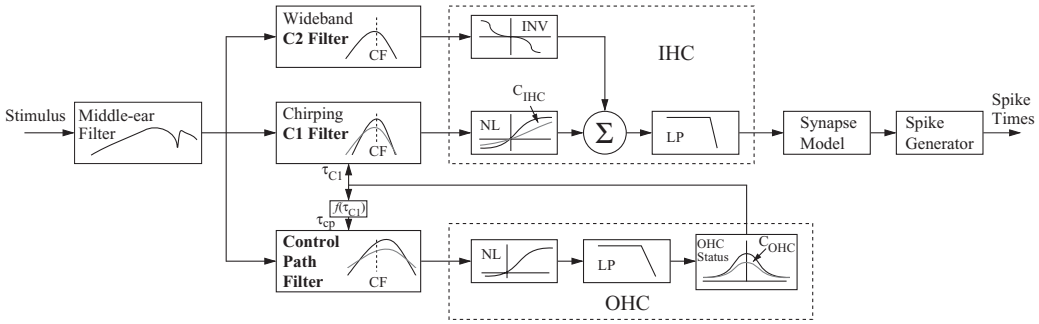

Figure 1.- Schematic of the auditory periphery model of Zilany and Bruce [7, 8]. Reprinted with permission of the Acoustical Society of America © 2006.

Model threshold tuning curves for three fibres with different CFs are shown in Fig. 2. In all cases, impairment of the OHCs leads to elevation and broadening of the tuning curve. For fibres with normal CFs < 750 Hz an upwards shift in CF is observed with increasing impairment, while a downward shift is exhibited for fibres with normal CFs > 1.5 kHz; no shift occurs for fibres with normal CFs between 750 Hz and 1.5 kHz. These different directions of the CF shift are produced by the differing directions of the BM filter chirp in these different CF regions. These changes in the AN fibre tuning are also created dynamically as a function of the stimulus (via the control-path output) for normal model fibres, such that broadened, elevated and frequency-shifted tuning arises for moderate to high stimulus levels. Note that the threshold tuning properties (in both the tip and the tail of the tuning curve) are determined only by the C1 filter; the C2 filter contributes to responses only at much higher stimulus levels.

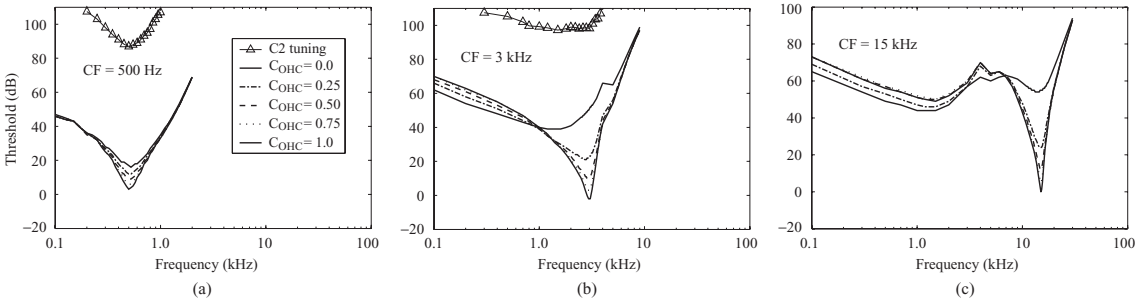

Figure 2.- Tuning curves for model fibres with CFs of (a) 500 Hz, (b) 3 kHz and (c) 15 kHz, for different levels of OHC impairment, ranging from normal function (C_{OHC} = 1.0) to complete dysfunction (C_{OHC} = 0.0). Also shown are the respective C1/C2 transition-threshold tuning curves. The transition threshold corresponds to a phase shift of 90°. Modified from Fig. 5 of [7] with permission of the Acoustical Society of America © 2006.

Also plotted in Fig. 2 are the C1/C2 transition-threshold tuning curves for the two fibres with lower CFs; fibres with CFs > 5 kHz do not exhibit sufficient phase locking to a tone around CF for a reliable estimate of the response phase to be estimated. The C1/C2 transition occurs when two conditions are met: i) the stimulus level needs to be high enough to drive the C1 filter into its broadest and most attenuating state, such that its frequency response matches that of the C2 filter, and ii) the stimulus level must also be high enough for the output of the C2 transduction function to grow as large as the output of the C1 transduction function. Some example rate-level and corresponding phase-level plots are given in Fig. 3. In most cases, the C1/C2 interaction produces only a phase shift without any noticeable reduction in the mean spike rate, although a

"notch" in the rate-level function is observed for one case, as is also sometimes found in the physiological data [9].

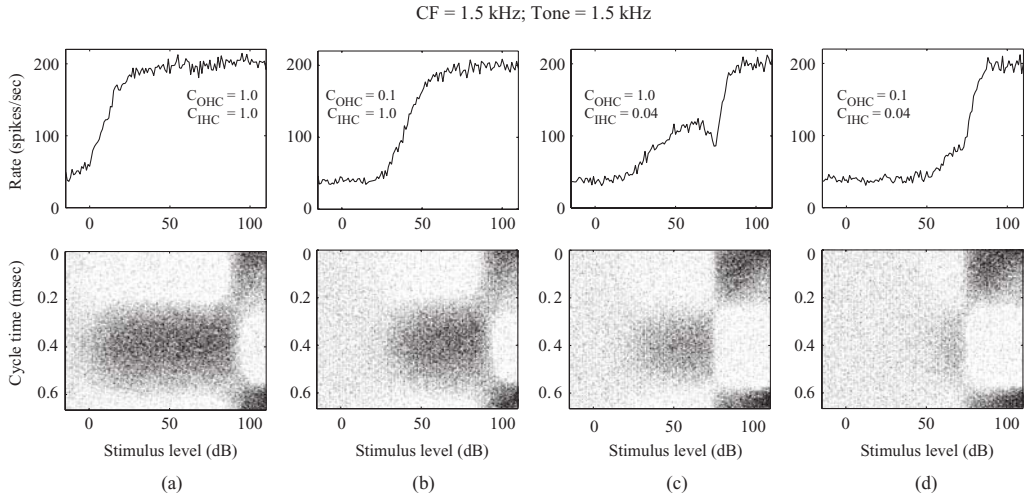

Figure 3.- Spike rate (top row) and phase (bottom) versus stimulus level for different degrees of OHC and IHC impairment. Reprinted from [7] with permission of the Acoustical Society of America © 2006.

The reason for the absence of a notch for many AN fibres is that peak splitting can occur in the phase-locked response in the C1/C2 transition region, consistent with the two-factor cancellation hypothesis of Kiang [9]. How this arises in the model of Zilany and Bruce [7, 8] is illustrated in Fig. 4. At low to moderate levels the C1 response dominates the IHC receptor potential, while at high levels the C2 response takes over. In the transition region, the fundamental frequency component is equal in amplitude and opposite in phase in the C1 and C2 transduction function outputs, leading to cancellation of the fundamental. However, rectification in the C1 transduction function produces a 2nd-harmonic that is not cancelled by the non-rectified C2 response, leading to a residual 2nd-harmonic component in the phase-locked response. For stimulus frequencies much above 1 kHz, this 2nd-harmonic will be greatly attenuated by the IHC LP filter, so peak splitting will be absent and a notch in the rate-level function will appear instead. The presence of a notch and the absence of peak splitting will also arise if the C1 transduction function is only weakly rectifying, as illustrated in Fig. 3(c), the case of substantial IHC impairment without OHC impairment.

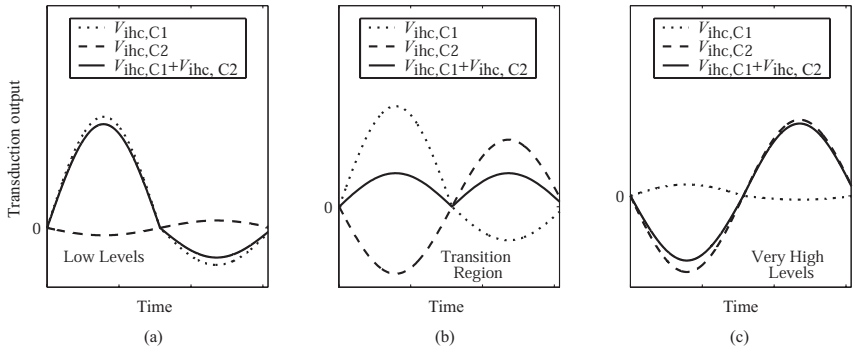

Figure 4.- Schematic illustrating how "peak splitting" occurs in the C1/C2 transition region. Reprinted from [7] with permission of the Acoustical Society of America © 2006.

An important property of the C1/C2 transition behaviour is that it is also observed in responses to broadband stimuli such as vowels [10]. Fig. 5 shows model responses to the vowel /ɛ/ for increasing stimulus levels. These model predictions are consistent with the physiological data [8]. For a fibre with a CF at the vowel's 2nd formant (F2), at low to moderate levels the F2 component of the vowel "captures" the phase-locked response of the fibre because of the narrow C1 tuning and suppressive C1 nonlinearity. At higher stimulus levels, the C1 tuning

broadens and other frequency components of the vowel begin to appear in the AN fibre response, particularly the higher-intensity 1st formant (F1) of the vowel. At the C1/C2 transition threshold (around 104 dB SPL in this case), the F1 and F2 components are cancelled by the anti-phase C1 and C2 responses, leaving the other harmonic components of the vowel to dominate the response. Above the transition threshold, F1 dominates the response because of the broad and linear tuning of the C2 filter.

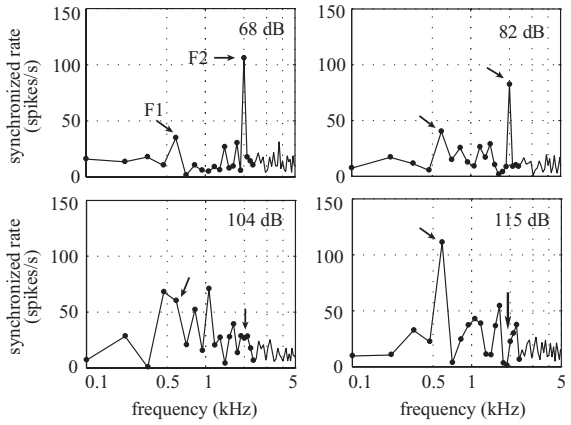

Figure 5.- The frequency spectra of phase-locked responses to the vowel /ε/ for a model AN fibre with a CF at F2 (2.0 kHz). Modified from Fig. 7 of [8] with permission of the Acoustical Society of America © 2007.

An important feature of the C1/C2 transition for vowels is that the formants undergo the transition at the same vowel level, rather than at the level at which a tone with the formant frequency undergoes the C1/C2 transition [10]. This rules out separate processing of the different formants, as might occur in other models incorporating the C1/C2 transition [11]. In the model of Zilany and Bruce [7, 8], the overall stimulus level determines the tuning of the C1 filter (via the broadband control path). As described above, when the stimulus level is sufficiently high, the C1 filter tuning will almost exactly match that of the C2 filter, such that all frequency components will be out of phase in the C1 and C2 filter outputs. Consequently, all the stimulus components undergo the C1/C2 transition simultaneously [8]. For individual tones, a somewhat higher stimulus level may be required to drive the C1 filter frequency response to match the C2 frequency response, and consequently the transition threshold will be higher for a single tone at a formant frequency than for the formant component of a vowel [8].

REMAINING ISSUES

Despite the greatly improved predictive accuracy that has been achieved over the past 14 years, a number of issues remain unresolved that could lead to further improvements in the accuracy of the model in predicting some sets of data.

The IHC-AN synapse section of the model has remained relatively unchanged in the various versions of the model, and the model is still unable to accurately describe adaptation to increments and decrements in an ongoing stimulus, and predictions of responses to amplitude-modulated signal are also somewhat inaccurate. In addition to an improved model of the synapse, it may be necessary to incorporate adaptation into the IHC and discharge generator sections of the model. Another unresolved issue is how exactly IHC impairment should be implemented. Zilany and Bruce [7, 8] used the method proposed by Bruce et al. [5], which ensured that maximum spike rates for that model were not reduced for impaired fibres, consistent with the physiological data. However, because of the inclusion of the C2 response in Zilany and Bruce [7, 8], the maximum spike rate is unaffected by impairment of the C1 transduction function, and thus different approaches to applying IHC impairment may be valid.

Cochlear amplifier (CA) gain is the difference in the C1 filter gain for normal OHC function and for complete OHC dysfunction. Direct measurement of BM vibrations is impractical in the cat, so indirect methods have been used for estimating the magnitude of the CA gain as a function of CF. In the model of Zhang et al. [4], values for low and high CFs were estimated from guinea

pig and chinchilla BM data, and linear interpolation was applied in between. The resulting function is shown by the dashed line in Fig. 6.

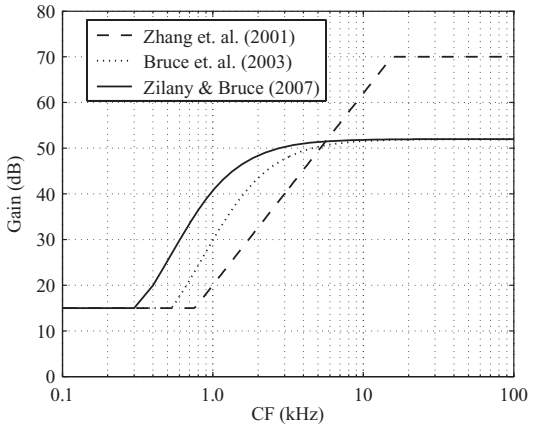

Figure 6.- Cochlear amplifier gain versus CF for three different auditory-periphery models. Modified from Fig. 3 of [8] with permission of the Acoustical Society of America © 2007.

Bruce et al. [5] found that predictions of a number of different sets of physiological data could be improved if the CA gain was increased at mid frequencies and reduced at high frequencies, producing the dotted curve in Fig. 6. However, Zilany and Bruce [8] found that the model's accuracy in predicting C1/C2 transition thresholds and vowel responses could be improved more by further increasing the CA gain at low to mid frequencies, as shown by the solid line in Fig. 6. However, this is still an indirect estimate of CA gain versus CF, and the correct CA gain at low frequencies has not been determined for cats. One method for directly determining the CA gain as a function of frequency in cats could be to apply sufficient doses of cisplatin to fully impair OHCs without damaging IHCs and to measure the resulting changes in AN fibre tuning curves.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Laurel H. Carney & her many students and Drs. Eric D. Young & Murray B. Sachs for their foundational work and continuing contributions to the model of the cat auditory periphery.

References: [1] L. Deng, C. D. Geisler: A composite auditory model for processing speech sounds. Journal of the Acoustical Society of America **82** (1987) 2001–2012

[2] E. A. Lopez-Poveda: Spectral processing by the peripheral auditory system: Facts and models. International Review of Neurobiology **70** (2005) 7–48

[3] L. H. Carney: A model for the responses of low-frequency auditory-nerve fibers in cat. Journal of the Acoustical Society of America **93** (1993) 401–417

[4] X. Zhang, M. G. Heinz, I. C. Bruce, L. H. Carney: A phenomenological model for the responses of auditory-nerve fibers. I. Nonlinear tuning with compression and suppression. Journal of the Acoustical Society of America **109** (2001) 648–670

[5] I. C. Bruce, M. B. Sachs, E. D. Young: An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. Journal of the Acoustical Society of America **113** (2003) 369–388

[6] Q. Tan, L. H. Carney: A phenomenological model for the responses of the auditory-nerve fibers. II. Nonlinear tuning with a frequency glide. Journal of the Acoustical Society of America **114** (2003) 2007–2020

[7] M. S. A. Zilany, I. C. Bruce: Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery. Journal of the Acoustical Society of America **120** (2006) 1446–1466

[8] M. S. A. Zilany, I. C. Bruce: Representation of the vowel /ε/ in normal and impaired auditory nerve fibers: Model predictions of responses in cats. Journal of the Acoustical Society of America **122** (2007) – In Press

[9] N. Y.-S. Kiang: Curious oddments of auditory-nerve studies. Hearing Research **49** (1990) 1–16

[10] J. C. Wong, R. L. Miller, B. M. Calhoun, M. B. Sachs, E. D. Young: Effects of high sound levels on responses to the vowel /ε/ in cat auditory nerve. Hearing Research **123** (1998) 61–77

[11] T. Lin, J. L. Goldstein: Quantifying 2-factor phase relations in non-linear responses from low characteristic-frequency auditory-nerve fibers. Hearing Research **90** (1995) 126–138