Description of a human-based computational auditory-nerve model with linear tuning

August 16, 1999

Michael G. Heinz, H. Steven Colburn and Laurel H. Carney

Excerpt from:
Performance limits for frequency and level discrimination based on a computational auditory-nerve model

Michael G. Heinz

Speech and Hearing Sciences Program
Division of Health Sciences and Technology
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, Massachusetts 02139

Hearing Research Center [a)]
Biomedical Engineering Department
Boston University
44 Cummington Street
Boston, Massachusetts 02215

H. Steven Colburn and Laurel H. Carney

Hearing Research Center
Biomedical Engineering Department
Boston University
44 Cummington Street
Boston, Massachusetts 02215

[a)]Address and author to whom correspondence should be addressed.
Electronic mail: *mgheinz@mit.edu*

Running Title: Performance limits for auditory discrimination

(Received:

II. GENERAL METHODS

A. Computational auditory-nerve model

The type of AN model for which the current study was designed is implemented on a computer and can process an arbitrary stimulus and produce a time-varying discharge rate that can be used with a nonstationary point process (e.g., Poisson) to simulate AN discharge times. Examples of this type of model were described by Payton (1988) and Carney (1993). A simplified version of the Carney (1993) model was used in the present study to provide a better match to the AN model used by Siebert (1970), and thus to simplify comparisons between the two studies.

Figure 1 shows pure-tone response properties of the computational AN model that are important for the present study. Table I provides details of the implementation of the model. A linear fourth-order gamma-tone filter bank was used to represent the frequency selectivity of AN fibers [see population response in Fig. 1(a)], and was implemented similarly to Carney (1993), except there was no variation in tuning with level. Model filter bandwidths were based on estimates of human bandwidths from the psychophysical notched-noise method for estimating auditory filter shapes [Glasberg and Moore (1990), see Table I]. Psychophysically measured filters have been shown to match AN frequency tuning when both were measured in guinea pig (Evans *et al.*, 1992). The bandpass filter was followed by a memoryless, asymmetric, saturating nonlinearity (implemented as an arctan with a 3:1 asymmetry), which represents the mechano-electric transduction of the inner hair cell (IHC). The saturating nonlinearity contributes to the limited dynamic range of AN fibers [Fig. 1(b)]. [Note that the dynamic range is also affected by the IHC-AN synapse, see below and Patuzzi and Robertson (1988).] All AN model fibers had a rate threshold of roughly 0 dB SPL, a spontaneous rate of 50 spikes/s, and a maximum sustained rate of roughly 200 spikes/s. The model dynamic range for sustained rate was roughly 20-30 dB, while the dynamic range for onset rate was much larger. The synchrony-level curves showed a threshold that was roughly 20 dB below rate threshold, a maximum at a level that was just above rate threshold, and a slight decrease in synchrony as level was increased further [comparable to Johnson (1980) and Joris *et al.* (1994)].

An important property for the present study is the rolloff in phase-locking as frequency increases above 2-3 kHz [Fig. 1(c); Johnson, 1980; Joris *et al.*, 1994]. Weiss and Rose (1988) compared synchrony versus frequency in five species on a log-log scale and reported that the data from all species were well described by a lowpass filter with roughly 100 dB/decade rolloff (the only difference across species was the 3-dB cutoff frequency, e.g., $f_c = 2.5$ kHz for cat, and $f_c = 1.1$ kHz for guinea pig). To achieve the proper rolloff in synchrony (for all species) and cutoff frequency (for cat), seven first-order lowpass filters were used, each with a first-order cutoff frequency of 4800 Hz. The resulting filter had a 3-dB cutoff frequency near 2500 Hz and $\sim$ 100 dB/decade rolloff in the frequency range 4-6 kHz [Fig. 1(c)]. The model synchrony coefficients above 5 kHz are a simple extrapolation of the physiological data, which appears to be reasonable given the consistent lowpass shape and slope across many species reported by Weiss and Rose (1988).

Neural adaptation was introduced through a simple three-stage diffusion model based on data from Westerman and Smith (1988) for the IHC-AN synapse. The continuous-time

version of this adaptation model used by Carney (1993) was simplified by using fixed values for the immediate and local volumes and the local and global permeabilities, consistent with a similar implementation used by Lin and Goldstein (1995), and with Fig. 3 of Westerman and Smith (1988). The immediate permeability was a function of the input amplitude, where the relation was essentially linear above the resting permeability, and exponentially decayed to zero for negative inputs (see Table I). The model's time constants for rapid adaptation (1.3 ms at high levels) and short-term adaptation (63 ms at high levels) are consistent with those reported for AN fibers (e.g., Westerman and Smith, 1988). The output of the AN model [Fig. 1(d)] represents the instantaneous discharge rate $r(t, f, L)$ of an individual high-spontaneous-rate, low-threshold AN fiber.

The computational AN model contains several assumptions, and does not include several known AN response properties. It is assumed that the basic filter shape is the same for low and high CF's, with the only difference being filter bandwidth. The significance of this assumption, which ignores the tails of AN tuning curves observed for high-CF AN fibers (Kiang and Moxon, 1974), is minimized in the present study for which predictions are limited to low- and mid-level stimuli. Level-dependent variation in tuning (e.g., Ruggero *et al.*, 1997) is omitted from this model for simplicity. The potential effects of the variation in behavioral threshold as a function of frequency due to filtering by the external and middle ear are ignored. The model does not include the distribution of spontaneous rate and threshold across AN fibers (Liberman, 1978), or the increases in AN fiber innervation density from apex to base (Keithley and Schriber, 1987; Liberman *et al.*, 1990). The AN model does not include the olivocochlear efferent system or the middle-ear-muscle reflex system. The use of a nonstationary Poisson process to model the randomness of AN discharges ignores the absolute and relative refractory periods that are observed in AN interval histograms (Kiang *et al.*, 1965), and the related observation of lower variance in discharge counts than that predicted by a Poisson model (Teich and Khanna, 1985; Young and Barta, 1986; Delgutte, 1987; 1996).

REFERENCES

Carney, L.H. (**1993**). "A model for the responses of low-frequency auditory-nerve fibers in cat," J. Acoust. Soc. Am. **93**, 401-417.

Delgutte, B. (**1987**). "Peripheral auditory processing of speech information: implications from a physiological study of intensity discrimination," in *The Psychophysics of Speech Perception*, edited by M.E.H. Schouten (Nijhoff, Dordrecht, the Netherlands), pp. 333-353.

Delgutte, B. (**1996**). "Physiological models for basic auditory percepts," in *Auditory Computation*, edited by H.L. Hawkins, T.A. McMullen, A.N. Popper, and R.R. Fay (Springer-Verlag, New York), pp. 157-220.

Evans, E.F., Pratt, S.R., Spenner, H., and Cooper, N.P. (**1992**). "Comparisons of physiological and behavioral properties: Auditory frequency selectivity," in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, New York), pp. 159-169.

Glasberg, B.R., and Moore, B.C.J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103-138.

Greenwood, D.D. (**1990**). "A cochlear frequency-position function for several species - 29 years later," J. Acoust. Soc. Am. **87**, 2592-2605.

Johnson, D.H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. **68**, 1115-1122.

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

Keithley, E.M., and Schreiber, R.C. (**1987**). "Frequency map of the spiral ganglion in the cat," J. Acoust. Soc. Am. **81**, 1036-1042.

Kiang, N.Y.S., Watanabe, T., Thomas, E.C., and Clark, L.F. (**1965**). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge, MA).

Kiang, N.Y.S., and Moxon, E.C. (**1974**). "Tails of tuning curves of auditory-nerve fibers," J. Acoust. Soc. Am. **55**, 620-630.

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

Liberman, M.C., Dodds, L.W., and Pierce, S. (**1990**). "Afferent and efferent innervation of the cat cochlea: Quantitative analysis with light and electron microscopy," J. Comp. Neurol. **301**, 443-460.

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

Patterson, R., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (**1987**). "An efficient auditory filterbank based on the gammatone function," SVOS Final Report: The Auditory Filterbank.

Patuzzi, R., and Robertson, D. (**1988**). "Tuning in the mammalian cochlea," Physiol. Rev. **68**, 1009-1082.

Payton, K.L. (**1988**). "Vowel processing by a model of the auditory periphery: A comparison to eighth-nerve responses," J. Acoust. Soc. Am. **83**, 145-162.

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.

Siebert, W.M. (**1970**). "Frequency discrimination in the auditory system: place or periodicity mechanisms?," Proc. IEEE, **58**, 723-730.

Teich, M.C., and Khanna, S.M. (**1985**). "Pulse-number distribution for the neural spike train in the cat's auditory-nerve," J. Acoust. Soc. Am. **77**, 1110-1128.

Weiss, T.F., and Rose, C. (**1988**). "A comparison of synchronization filters in different auditory receptor organs," Hear. Res. **33**, 175-180.

Westerman, L.A., and Smith, R.L. (**1988**). "A diffusion model of the transient response of the cochlear inner hair cell synapse," J. Acoust. Soc. Am. **83**, 2266-2276.

Young, E.D., and Barta, P.E. (**1986**). "Rate responses of auditory-nerve fibers in noise near masked threshold," J. Acoust. Soc. Am. **79**, 426-442.

TABLE I. Equations and parameters needed to implement the computational AN model.

Symbol	Description (units)	Equations/values
Human cochlear map[a]		
x	distance from apex (mm)	
$f(x)$	frequency corresponding to a position x (Hz)	$= 165.4(10^{0.06x} - 0.88)$
Gamma-tone filters[b]		
CF	characteristic frequency (kHz)	
ERB	equivalent rectangular bandwidth (Hz)	$= 24.7(4.37CF + 1)$
τ	Time constant of gamma-tone filter (s)	$= [2\pi(1.019)ERB]^{-1}$
$gtf[k]$	output of gamma-tone filter	
Inner-hair-cell		
$ihc[k]$	output of saturating nonlinearity $ihc[k] = \frac{\arctan(K \cdot gtf[k]+\beta)-\arctan(\beta)}{\pi/2-\arctan(\beta)}$	
K	controls sensitivity	1225
β	sets 3:1 asymmetric bias	-1
$ihc_L[k]$	lowpass-filtered inner-hair-cell output (see text)	
Neural adaptation model[c]		
T_s	sampling period (s) (see text)	
r_o	spontaneous discharge rate (spikes/s)	50
V_I	immediate "volume"	0.0005
V_L	local "volume"	0.005
P_G	global permeability ("volume"/s)	0.03
P_L	local permeability ("volume"/s)	0.06
PI_{rest}	resting immediate permeability ("volume"/s)	0.012
PI_{max}	maximum immediate permeability ("volume"/s)	0.6
C_G	global concentration ("spikes/volume") $C_G = C_L[0](1 + P_L/P_G) - C_I[0]P_L/P_G =$	6666.67
$P_I[k]$	immediate permeability ("volume"/s) $P_I[k] = 0.0173\ln\{1 + \exp(34.657 \cdot ihc_L[k])\}$	
$C_I[k]$	immediate concentration ("spikes/volume") $C_I[k+1] = C_I[k] + (T_s/V_I)\{-P_I[k]C_I[k] + P_L[k](C_L[k] - C_I[k])\}$ $C_I[0] = r_o/PI_{rest} =$	4166.67
$C_L[k]$	local concentration ("spikes/volume") $C_L[k+1] = C_L[k] + (T_s/V_L)\{-P_L[k](C_L[k] - C_I[k]) + P_G(C_G - C_L[k])\}$ $C_L[0] = C_I[0](PI_{rest} + P_L)/P_L =$	5000.00
$r[k]$	instantaneous discharge rate (spikes/s)	$= P_I[k]C_I[k]$

[a] Greenwood (1990).

[b] see Patterson *et al.* (1987); Glasberg and Moore (1990); Carney (1993).

[c] see Westerman and Smith (1988); Carney (1993).

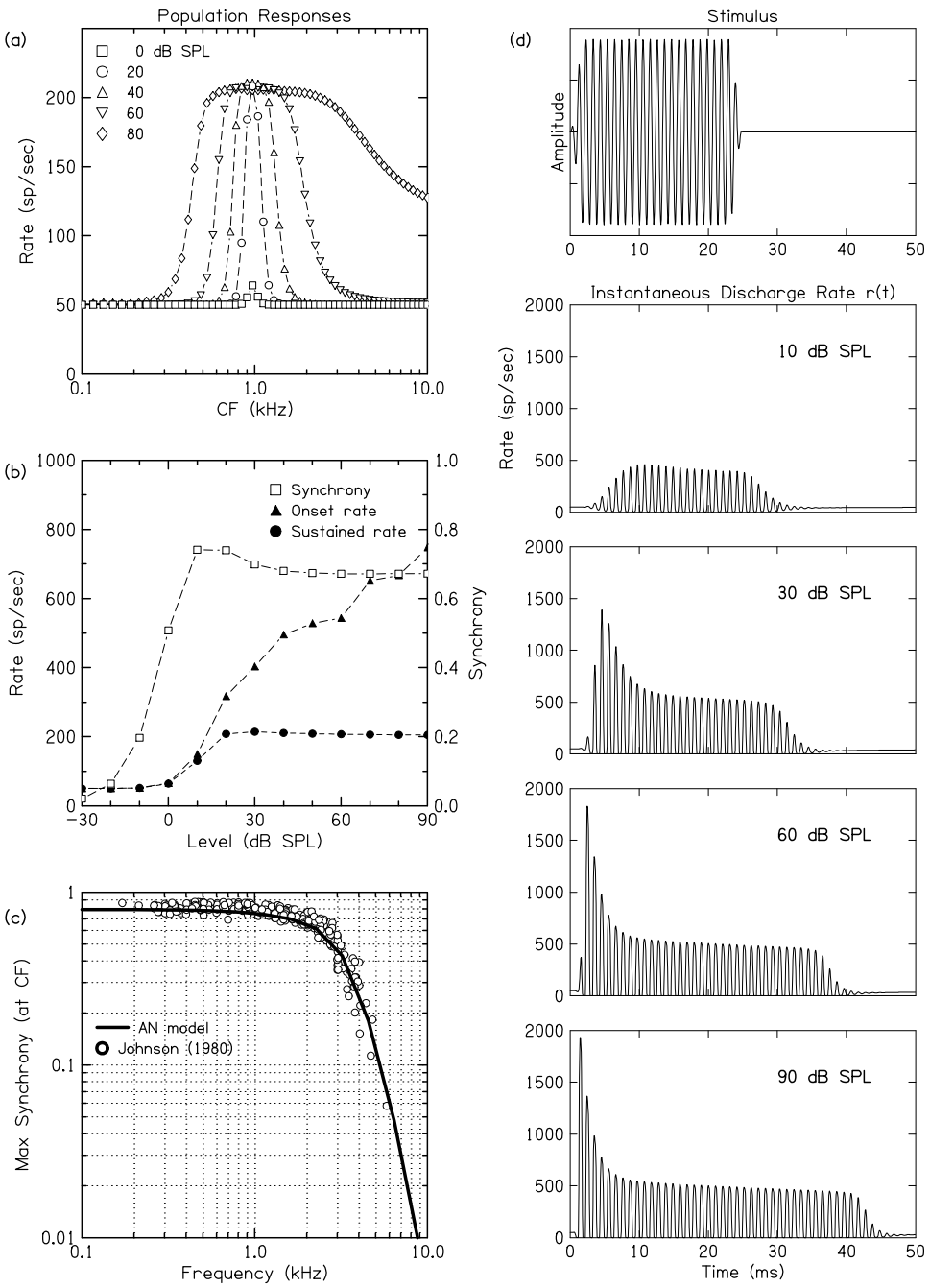

Figure 1: Basic pure-tone response properties of the computational AN model. (a) Sustained-rate population responses for the 60 AN-model CF's used in this study to pure tones at several levels. Tones were 970 Hz, 62 ms (10-ms rise/fall), with levels from 0 to 80 dB SPL. Sustained rate (calculated as the average rate over all full stimulus cycles within a temporal window from 10 to 52 ms) is plotted versus characteristic frequency (CF) of each model fiber. (b) Onset rate, sustained rate, and synchrony for a 970-Hz model fiber responding to a 970-Hz, 62-ms tone as a function of level. Onset rate was calculated as the maximum average rate over one stimulus cycle. The synchrony coefficient (units on right axis), represents the vector strength (Johnson, 1980) of the model response calculated over one cycle beginning at 40 ms. (c) Maximum synchrony coefficient over level for tones at CF as a function of frequency. Circles are data from cat (Johnson, 1980). (d) Stimulus waveform and instantaneous discharge rates $r(t, f, L)$ for a f = 970-Hz fiber in response to 970-Hz, 25-ms (2-ms rise/fall) tones over a range of levels (L).