

A two-stage nonlinear cochlear model possesses automatic gain control

C. Daniel Geisler

Department of Neurophysiology and Department of Electrical and Computer Engineering, University of Wisconsin—Madison, Madison, Wisconsin 53706

Steven Greenberg

Department of Neurophysiology, University of Wisconsin—Madison, Madison, Wisconsin 53706

(Received 27 February 1986; accepted for publication 28 May 1986)

A model of the cochlea is explored using as stimuli two simultaneously presented sinusoids of equal amplitude. The model consists of two stages: a linear bandpass filter, followed by a reservoir-type representation of the hair-cell/nerve-fiber complex. Fast Fourier transforms of the model's output were computed. While the amplitudes of the individual response components were strongly nonlinear functions of intensity, the *ratio* of the magnitudes of the response components at the frequencies of the two stimulating sinusoids was found to be nearly equal, over a wide intensity range, to the *ratio* of the amplitudes which those stimulating sinusoids possessed at the output of the filter. Thus the reservoir stage exerts "automatic gain control."

PACS numbers: 43.63.Bq, 43.63.Nc

INTRODUCTION

There is a curious paradox in the temporal patterning of responses of auditory-nerve fibers. On the one hand, when a single tone is used, the synchrony evoked by that tone (as measured by the unnormalized Fourier transform coefficient at that frequency, frequently called the "synchronized rate") is a strongly nonlinear function of stimulus amplitude. On the other hand, the shape of the period histogram remains that of a sinusoid (typically half-wave rectified, as no negative values can occur) over a wide intensity range (Rose *et al.*, 1967). Likewise, when complex sounds are used, there are many conditions in which the response patterns behave like those of a linear system. Fourier transforms of the reverse correlograms obtained using white noise stimuli, for one example, are close approximations to the observed frequency selectivity obtained using pure tones (Evans, 1977, 1985). For another example, we demonstrated recently that the single-tone frequency selectivity of a cochlear-nerve fiber could be determined from the relative degree of synchronization it displayed to the components of a two-tone signal, and that this measure changed relatively little over a moderate range of sound-pressure levels (Greenberg *et al.*, 1986).

It is the purpose of this paper to show that a simple model of the auditory periphery that incorporates a single reservoir of hair-cell neurotransmitter shares these same paradoxical characteristics. We will show that this reservoir acts like an "automatic gain control" device, providing good reproduction of the short-term waveforms while adjusting the overall amplitude envelope of the response according to stimulus strength.

I. THE MODEL

A schematic diagram of the model is shown in Fig. 1. It consists of two stages: a linear bandpass filter, followed by a

reservoir-type representation of the hair-cell/nerve-fiber complex. The linear filter is of the type described by Duifhuis (1973). It is a minimum-phase filter which is specified by the resonant frequency and by the sum of the absolute values of the low- and high-frequency slopes. For the purposes of this paper, the resonant (characteristic) frequency was set at 978 Hz and the combined slopes at 81 dB/oct. With these parameters, the filter's frequency-response characteristic, shown in Fig. 5, is a close approximation in shape to that of fiber A33 of the Carney-Geisler study (1986).

The reservoir stage of our model is a modified version of the Oono-Sujaku reservoir-type representation of the hair-cell/nerve-fiber complex (1975a,b). In this representation the instantaneous discharge of the afferent fiber is proportional both to the contents, $c(t)$, of the single neurotransmitter reservoir and to a permeability function, $z(t)$, which is related to the stimulus waveform. Specifically, the probability, $p(t)$, of a nerve-fiber discharge within a small period of time Δt is, approximately,

$$p(t) = c(t)z(t)r(t)\Delta t, \quad (1)$$

where $r(t)$ is a refractory function. The permeability function, $z(t)$, is assumed in this representation to be proportional to a half-wave rectified version of the sum of the input waveform and a dc bias (required to simulate a fiber's spontaneous discharge rate). For the refractory function, we assumed an absolute refractory period of 1 ms, followed (our modification) by a 4-ms period of exponentially decreasing relative refractoriness. Parameters were adjusted for a spontaneous rate of 60 discharges/s.

Waveforms generated by the model, when tuned to a resonant frequency of 100 Hz and presented with a tone burst of that frequency at 40 dB SL¹ amplitude, are shown in Fig. 2. The output of the model, probability of discharge, is shown as a function of time in Fig. 2(a), while the permeability function and reservoir contents are shown in Fig.

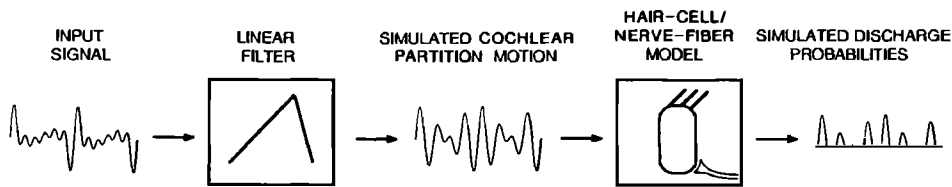


FIG. 1. Schematic diagram of the model. The linear filter is from Duifhuis (1973), and the reservoir model modified from Oono and Sujaku (1975a,b).

2(b) and (c), respectively. The envelope of the discharge probabilities has a realistic exponential type of decay, and then shows a slow recovery to the spontaneous level. A skewing of the response probabilities is seen during single response cycles, as has previously been demonstrated by this class of models (Schwid and Geisler, 1982). Such skewing also occurs in primary-fiber response data obtained with low-frequency stimulation (Johnson, 1980). The reservoir's contents [Fig. 2(c)], which clearly show the successive epochs of depletion and replenishment, drop to an average value of about 0.15 by the end of the 100-ms stimulus presentation period.

Responses of the model to sinusoidal stimuli of 1000 Hz presented at several different amplitudes are shown in Fig. 3. As seen in the responses of actual primary fibers with high spontaneous rates, a large initial transient is followed by a

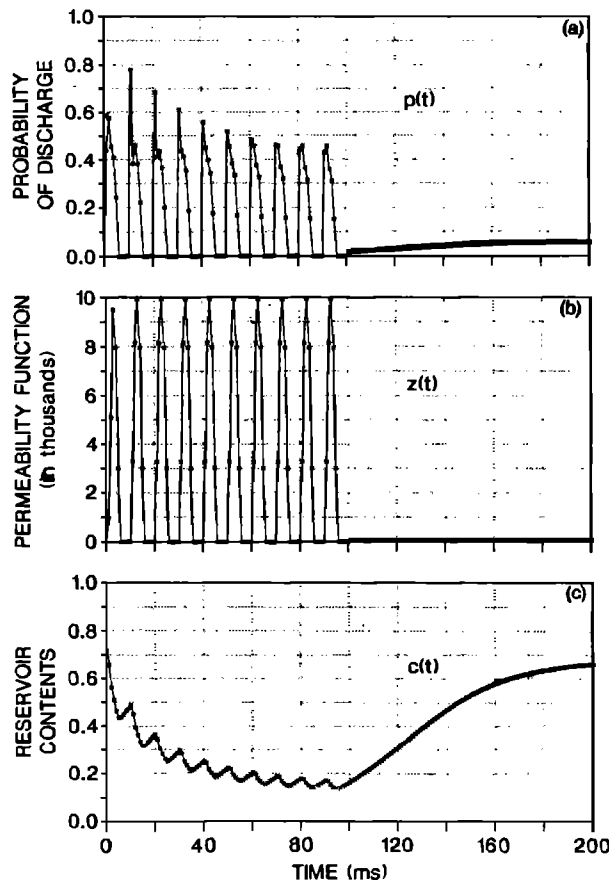


FIG. 2. (a) Discharge probabilities of the model, $p(t)$ (for 1-ms time bins) for a 40-dB SL¹ CF sinusoid of 100 Hz lasting 100 ms. (b) Permeability function $z(t)$, produced by the 40-dB stimulus. (c) Reservoir contents $c(t)$, produced by the 40-dB stimulus. Temporal increments used in the computer simulation were 50 μ s. A stimulus rise time of 2.5 ms and a fall time of zero were used.

gradual adaptation to a steady-state level. The higher the stimulus level the greater is the initial transient. The rate of the "short-term" adaptation, that which extends from the initial transient for 100 ms or so, while somewhat variable with intensity level, is approximately the same for all the curves, as seen in neural data (cf. Westerman and Smith, 1984). To be sure, the model's adaptation behavior at 70 dB SL, the highest intensity used, is not exactly exponentially decaying. This shortcoming was considered acceptable, however, because a single exponential curve fitted to that response would have the same order of magnitude as the others, and the exact rate of adaptation seems to have a negligible effect upon the primary conclusions to be drawn. Recovery of the model's discharge rate following stimulation depends on stimulus level, as is also observed physiologically (Harris and Dallos, 1979).

The amplitudes of the steady-state portions of the responses to sinusoidal stimuli of various levels are shown in Fig. 4. A rate threshold of about 5 dB SL is seen along with a dynamic range of about 45 dB. This latter value is somewhat larger than the 20–30-dB dynamic range generally seen with high-spontaneous-rate primary fibers (cf. Sachs and Abbas, 1974), but this difference does not affect the basic conclusions we arrived at.

II. RESPONSES OF THE MODEL TO TWO-TONE SIGNALS

Signals comparable to those presented in our previous study (Greenberg *et al.*, 1986) were used to stimulate the two-stage model. These stimuli were the sum of two sinusoids of equal amplitude and starting phase. The frequency of one of the components was always the same as the filter's characteristic frequency (f_{CF}), while the frequency of the second component (f_V) was varied across a range of an oc-

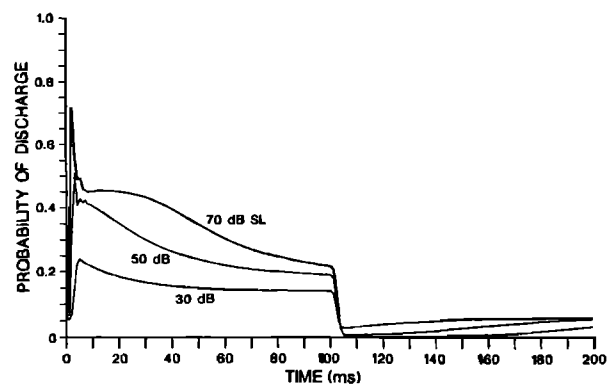


FIG. 3. Discharge probabilities of the model (for 1-ms time bins) for three intensities of a 1000-Hz CF tone lasting 100 ms. A rise time of 2.5 ms and a fall time of zero were used.

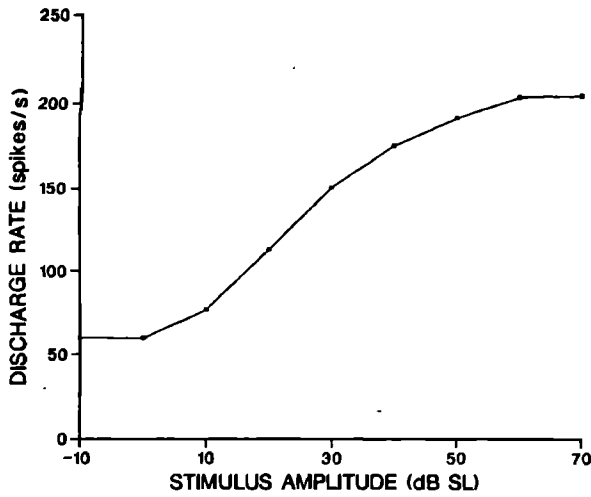


FIG. 4. Model discharge rate, calculated over the last 1.0 ms of a 100-ms CF stimulus, for various input amplitudes.

tave and a half. Signals of 100-ms duration were used in this computer (DEC VAX/750) simulation. However, only the last 51.2 ms of the model's responses were analyzed in order to minimize the adaptation effects associated with stimulus onset. Fast Fourier transforms (FFTs) of these segments were computed to determine the magnitude of the response components synchronized to the characteristic frequency and variable-frequency tones.

As in our previous report, the frequency selectivity was determined by computing the ratio of the magnitude of the Fourier component at f_V over that at f_{CF} . This ratio, expressed in dB, is shown by the asterisks in Fig. 5 for various pairs of sinusoids. Of course, when both sinusoids had the same frequency ($f_V = 978$ Hz), the ratio of the response amplitudes is unity (0 dB). As the frequency of the variable-frequency sinusoid was gradually moved away from CF, the response to the variable-tone component dropped, while the CF-tone response component stayed almost constant. Thus

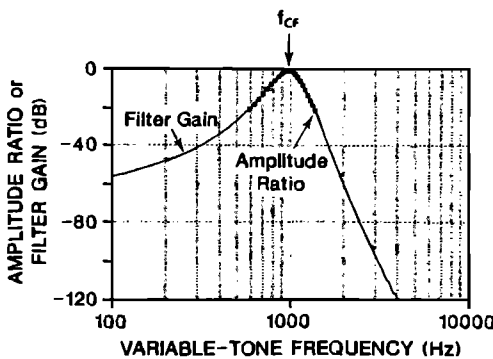


FIG. 5. The amplitude (in dB) versus frequency curve for the linear filter used in this study, with maximum gain set to zero dB (solid line). Also shown (asterisks) are the amplitude ratios, A_V/A_{CF} , derived from the model's output in response to two equal-amplitude sinusoids presented at 70 dB SL, for different pairs of frequencies. Only the last 51.2 ms of the 100-ms stimulus periods were used in the analyses (time bins of 50 μ s were used in the simulation). For purposes of preventing artifacts due to the finite time window used, f_V and f_{CF} were set at integer multiples of 19.53 Hz (1/0.0512 s).

the ratio of their amplitudes decreases as the difference between the two frequencies grows larger. Notice that the function traced out by these response ratios is very similar in shape to that of the filter's frequency characteristic, shown as the solid curve. That is, the ratio of response amplitudes for any pair of input sinusoids is very similar to the ratio of the amplitudes that those sinusoids have at the output of the filter.

Figure 6 shows the response ratios produced by two-tone pairs of equal amplitude presented over an 80-dB range of sound-pressure levels. The shapes of the functions obtained at the various SPLs, ranging from near threshold to suprasaturation levels, are remarkably similar to each other (and thus to the filter's frequency characteristic). Regardless of presentation level, the linear filter's frequency-response curve is effective in predicting the relative amplitudes of the principal response components. For comparison purposes, the response-ratio function obtained from the model with a different filter function (the same characteristic frequency but with slopes of only half the magnitude) is also shown. The resulting ratio function is rather flat (as is the filter's frequency-selectivity characteristic in this case) and bears no resemblance to those of the previous response-ratio functions. It is therefore clear that the filter's bandpass characteristic, and not a particular property of the reservoir stage determines the response ratios.

There are some variations of phases with varying intensity conditions, but these are not large. For example, when stimulated with the two-component complex of 781- and 978-Hz sinusoids over a wide variety of intensity combinations varying in level from 0–70 dB SL, the range of phases which occurred over all of these combinations was only 16° for the 781-Hz response component, and 12.5° for the response component at 978 Hz. Thus, aside from the skewing of the waveform which occurs at low frequencies (cf. Fig. 2),

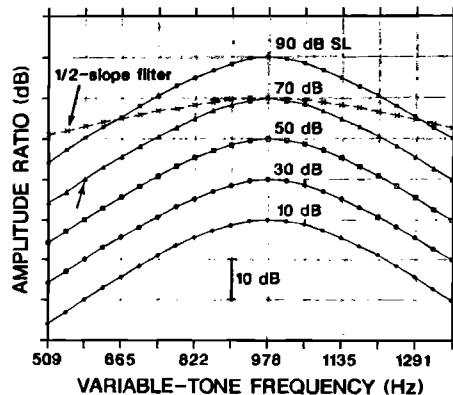


FIG. 6. Amplitude ratios, A_V/A_{CF} derived from the model using two equal-amplitude sinusoids, plotted as a function of f_V for five different intensities (solid lines). As an example, for 70-dB SL stimulation with 586- and 978-Hz sinusoids, A_V , the response amplitude to the 586-Hz component was 16 spikes/s (24.1 dB re: 1.0) while A_{CF} was 159 spikes/s (44.0 dB re: 1.0). Thus A_V/A_{CF} in that case was 16/159 or -19.9 dB (indicated by the upward-pointing arrow). Data points connected by the dotted line were obtained with 70-dB SL sinusoidal pairs when the model's high- and low-frequency slopes were nominally halved. The maximum amplitude ratio in all cases was 0 dB. Curves for the different intensities were spaced 10 dB apart to facilitate comparisons.

the waveforms which emerge from the filter stage are largely preserved in the output signals.

III. DISCUSSION

The results presented here can be understood by referring to the formulation of the Oono-Sujaku model. From Eq. (1), it is easily seen that for an instantly replenished reservoir [constant $c(t)$], the output of the model is approximately proportional to $z(t)$, a half-wave-rectified version of the input signal, plus a dc bias. Because the rectification will affect the amplitudes of the two components approximately equally, the ratio of the response components at the two principal frequencies remains relatively constant in this case.

However, the contents of the reservoir do not remain unchanged as a function of signal input level. At low sound-pressure levels the reservoir is relatively full, while at the highest intensities, it is almost completely depleted. Thus reservoir contents $c(t)$, act in the capacity of a volume control. For small inputs, its relatively large value enhances responses, while for large inputs, its relatively small value limits response amplitudes. For example, the average reservoir contents dropped to about 0.15 at the end of the stimulus duration period when 40-dB SL stimuli were used [Fig. 2(c)]; however, when 60-dB SL sinusoids of 100 Hz were used, the average value of reservoir contents dropped much further, to approximately 0.04. By contrast, 20-dB SL stimulation at 100 Hz caused much smaller rates of transmitter release and average reservoir contents in that case only dropped to 0.45 by the termination of the stimulus. The consequence of this continual adjustment of average reservoir contents, when coupled with the small amounts of phase shift which occur, is to preserve the principal features of the filtered (and rectified) stimulus waveform, as originally pointed out by Oono and Sujaku (1975a,b). The filtered spectra of some other complex stimuli, such as random noise or tone complexes, would also be largely preserved across a wide intensity range, as has been demonstrated physiologically (Evans, 1977, 1985). This automatic gain control was previously recognized by Evans, who incorporated circuits for it into his analog model of the cochlea (1980).

Looked at in another way, the model exhibits "two-tone synchrony suppression." Consider, for example, the case when the model, with the filter stage tuned to 978 Hz, was stimulated with a single sinewave of 586 Hz at 70 dB SL. This single tone was attenuated by 20.6 dB in the filter (cf. Fig. 5), yet still caused a nearly saturated synchronized-rate response of 146 spikes/s (43.3 dB relative to unity). When a 70-dB sinusoid at the characteristic frequency of the filter (978 Hz) was added to the first tone, the model responded as an automatic-gain-control device. That is, the synchronized response to the major component, the characteristic-frequency sinusoid, became 159 spikes/s (44.0 dB), while that to the 586-Hz tone dropped to 16 spikes/s (24.1 dB). This 19.9-dB ratio of response amplitudes is nearly the same as that of the two stimulating sinusoids as they left the filter (20.6 dB). The model therefore preserves the filtered waveform, but in doing so has reduced the 586-Hz response component by 19.2 dB (43.3–24.1 dB). This certainly is synchrony suppression, as commonly defined, but we feel that

viewing the process instead as automatic gain control provides the more fundamental insight.

The present model does not, however, apply in all cases. It is well known that stimulus components placed on the outskirts of a fiber's response area can, under certain conditions, either reduce the total discharge rate (e.g., Sachs and Kiang, 1968) or dominate the timing of discharges with a strength greater than that indicated in the fiber's single-tone responses (e.g., Sinex and Geisler, 1984). Neither of these phenomena² are exhibited by the model. As they tend to occur *in vivo* when the suppressing or capturing component is at a large amplitude relative to that of the CF tone, we believe that the model presented here generally applies only in those situations where all of the stimulus components have the same order of magnitude. In situations where large spectral peaks occur, other processes may also have to be invoked (e.g., lateral coupling, Jau and Geisler, 1983).

The similarity of the results obtained using the present model with those observed in cat auditory-nerve-fiber responses (Greenberg *et al.*, 1986) suggests an interesting interpretation of the physiological data. With the present model, the function composed of the output amplitude ratios evoked by a two-tone stimulus traces out the filter's frequency characteristic (cf. Figs. 5 and 6). According to the model, if the shapes of those amplitude-ratio functions flattened out with stimulus level in the physiological data, it would imply that the frequency-selective properties of the basilar-membrane/organ-of-Corti complex also decreased with input level. This is the result expected from high-frequency basilar-membrane measurements (cf. Rhode, 1978), and is indeed seen in our physiological amplitude-ratio data for those fibers whose Q_{10} s were greater than two (generally the ones with characteristic frequencies above 1 kHz). For fibers with low Q_{10} s (usually the ones with CFs below 1 kHz), however, the frequency selectivity either was relatively invariant with increasing intensity level or even improved slightly. This implies that cochlear-partition frequency selectivity in the apex is not a strong function of stimulus intensity. Thus it will be interesting to see if frequency-response measurements of basilar-membrane motion in this area prove to be invariant with stimulus intensity.

ACKNOWLEDGMENTS

This research was sponsored by NIH Program Project Grant NS-12732. We thank Phillip Rossman for his help in running the computer model.

¹A 0 dB SL (stimulus level) was assigned to that stimulus amplitude which was just below discharge-rate threshold (cf. Fig. 4).

²Several different phenomena appear to contribute to the reduction of the amplitude of the synchronized response to one tone when other components are added to the stimulus. These effects have usually been lumped together under the single term "suppression." We feel that new terminology is needed which will distinguish between the different effects. We suggest, therefore, that the phenomena described in this paper be denoted "automatic gain control," those which actually reduce rate as "rate suppression," and those where one tone takes more than its share of synchrony (as estimated from the filter function) be termed "strong-signal capture." If these latter two phenomena are shown to have the same source (cf. Jau and Geisler, 1983), then a more compact terminology would be possible.

- Carney, L. H., and Geisler, C. D. (1986). "A temporal analysis of auditory-nerve fiber responses to spoken stop consonant-vowel syllables," *J. Acoust. Soc. Am.* **79**, 1896-1914.
- Duifhuis, H. (1973). "Consequences of peripheral frequency selectivity for nonsimultaneous masking," *J. Acoust. Soc. Am.* **54**, 1471-1488.
- Evans, E. F. (1977). "Frequency selectivity at high signal levels of single units in cochlear nerve and nucleus," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 185-192.
- Evans, E. F. (1980). "An electronic analogue of single unit recording from the cochlear nerve for teaching and research," *J. Physiol.* **298**, 6-7.
- Evans, E. F. (1985). "Aspects of the neural coding of time in the mammalian peripheral auditory system relevant to temporal resolution," in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer, New York), pp. 74-95.
- Greenberg, S., Geisler, C. D., and Deng, L. (1986). "Frequency selectivity of single cochlear nerve fibers based on the temporal response pattern to two-tone signals," *J. Acoust. Soc. Am.* **79**, 1010-1019.
- Harris, D. M., and Dallos, P. (1979). "Forward masking of auditory nerve fiber responses," *J. Neurophysiol.* **42**, 1083-1107.
- Jau, Y. C., and Geisler, C. D. (1983). "Results from a cochlear model utilizing longitudinal coupling," in *Mechanics of Hearing*, edited by E. de Boer and M. Viergever (Delft U.P., Delft), pp. 169-176.
- 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.
- Oono, Y., and Sujaku, Y. (1975a). "A model for automatic gain control observed in the firing of primary auditory neurons," *Abstracts Trans. IECE (Japan)* **58**, 61-62 (in English).
- Oono, Y., and Sujaku, Y. (1975b). "A model for automatic gain control observed in the firings of primary auditory neurons," *Trans. IECE (Japan)* **58**, 352-358 (in Japanese with English figure captions).
- Rhode, W. S. (1978). "Some observations on cochlear mechanics," *J. Acoust. Soc. Am.* **64**, 158-176.
- Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (1967). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," *J. Neurophysiol.* **30**, 769-793.
- Sachs, M. B., and Abbas, P. H. (1974). "Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli," *J. Acoust. Soc. Am.* **56**, 1835-1847.
- Sachs, M. B., and Kiang, N. Y. S. (1968). "Two-tone inhibition in auditory-nerve fibers," *J. Acoust. Soc. Am.* **43**, 1120-1128.
- Sinex, D. G., and Geisler, C. D. (1984). "Comparison of the responses of auditory nerve fibers to consonant-vowel syllables with predictions from linear models," *J. Acoust. Soc. Am.* **76**, 116-121.
- Schwid, H. A., and Geisler, C. D. (1982). "Multiple reservoir model of neurotransmitter release by a cochlear inner hair cell," *J. Acoust. Soc. Am.* **72**, 1435-1440.
- Westerman, L. A., and Smith, R. L. (1984). "Rapid and short-term adaptation in auditory nerve responses," *Hear. Res.* **15**, 249-260.