LETTER Communicated by John Lazzaro

Evaluating Auditory Performance Limits: I. One-Parameter Discrimination Using a Computational Model for the Auditory Nerve

Michael G. Heinz
Speech and Hearing Sciences Program, Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A., and Hearing Research Center, Biomedical Engineering Department, Boston University, Boston, MA 02215, U.S.A.

H. Steven Colburn
Laurel H. Carney
Hearing Research Center, Biomedical Engineering Department, Boston University, Boston, MA 02215, U.S.A.

A method for calculating psychophysical performance limits based on stochastic neural responses is introduced and compared to previous analytical methods for evaluating auditory discrimination of tone frequency and level. The method uses signal detection theory and a computational model for a population of auditory nerve (AN) fiber responses. The use of computational models allows predictions to be made over a wider parameter range and with more complete descriptions of AN responses than in analytical models. Performance based on AN discharge times (*all-information*) is compared to performance based only on discharge counts (*rate-place*). After the method is verified over the range of parameters for which previous analytical models are applicable, the parameter space is then extended. For example, a computational model of AN activity that extends to high frequencies is used to explore the common belief that rate-place information is responsible for frequency encoding at high frequencies due to the rolloff in AN phase locking above 2 kHz. This rolloff is thought to eliminate temporal information at high frequencies. Contrary to this belief, results of this analysis show that rate-place predictions for frequency discrimination are inconsistent with human performance in the dependence on frequency for high frequencies and that there is significant temporal information in the AN up to at least 10 kHz. In fact, the all-information predictions match the functional dependence of human performance on frequency, although optimal performance is much better than human performance. The use of computational AN models in this study provides new constraints on hypotheses of neural encoding of frequency in the auditory system; however, the method is limited to simple tasks with deterministic stimuli. A companion article in this issue

("Evaluating Auditory Performance Limits: II") describes an extension of this approach to more complex tasks that include random variation of one parameter, for example, random-level variation, which is often used in psychophysics to test neural encoding hypotheses.

1 Introduction

One of the challenges in understanding sensory perception is to test, quantitatively, hypotheses for the physiological basis of psychophysical performance. The random nature of neural responses (i.e., that two presentations of an identical stimulus produce different discharge patterns) imposes limitations on performance. The fundamental requirement for relating neural encoding to psychophysical performance is a statistical description of neural discharge patterns in response to the relevant stimuli. In general, if the statistics of the response change significantly as a stimulus parameter is varied, then accurate discrimination of that parameter is possible.

Attempts to relate neural responses to psychophysical performance often involve the question of whether a small set of individual neurons can statistically account for performance, but can also involve studies of encoding by an entire population of neurons (see Parker & Newsome, 1998). Implicit in any hypothesis focused on a small set of neurons is the assumption that the brain ignores statistically significant information in other neurons within the population. To avoid this assumption, the total amount of information in a neural population can be quantified. The amount of total information, and the trends in information as stimulus parameters are varied, can be used to test hypotheses of neural encoding and suggest neural information-processing mechanisms. These studies, which require modeling approaches due to the inability to record from all neurons within a population, have typically been limited to simple psychophysical tasks with simple stimuli that can be studied with analytical models. This article describes a general method based on detection and estimation theory that allows computational models for stochastic neural responses to be used in studies that relate physiological properties to psychophysical performance. Computational models allow one to calculate the response pattern of neural activity for an arbitrary stimulus, whereas analytical models are typically functional descriptions of activity to a well-defined class of stimuli. A companion article in this issue ("Evaluating Auditory Performance Levels: II") describes a further generalization of this approach to include more complex psychophysical tasks in which one stimulus parameter is randomly varied in order to limit the cues available to a subject.

1.1 Relating Physiology to Psychophysics in the Auditory System. The study reported here uses the auditory system as an example of the application of signal detection theory (SDT) with computational neural models to quantify the total information in a neural population; however, the gen-

eral types of questions that this method is able to address quantitatively are relevant to neural encoding in any sensory system. The auditory nerve (AN) is an obligatory pathway between the cochlea (inner ear) and the central nervous system (Ryugo, 1992), and thus contains all of the information about an auditory stimulus that a listener can process. Siebert (1965, 1968, 1970) was the first to evaluate limits on psychophysical performance in auditory discrimination tasks by combining methods from SDT with models of peripheral auditory signal processing.[1] (See Delgutte, 1996, for a review of Siebert's approach and many subsequent studies.) While Siebert (1965, 1968, 1970) applied his approach to the auditory system, the general questions addressed with his method are fundamental to theoretical neuroscience, such as whether physiological noise inherent in neural encoding can account for human performance limits and whether temporal information in neural responses is used to encode sensory stimuli (Abbott & Sejnowski, 1999). In addressing these questions, Siebert (1965, 1968, 1970) used simple analytical models that included only those AN response properties considered to be of primary importance: bandpass frequency tuning, saturating rate-level curves, phase locking to tonal stimuli, a population of 30,000 AN fibers with logarithmic characteristic frequency (CF, the frequency of best response) spacing, and nonstationary Poisson statistics of AN discharges (for reviews, see Kiang, Watanabe, Thomas, & Clark, 1965; Ruggero, 1992; and Delgutte, 1996). These studies were limited to simple psychophysical tasks, such as pure-tone level and frequency discrimination, because of their relatively simple interpretations and the limits of the analytical AN models. Valuable insight into the encoding of frequency and level in the peripheral auditory system was garnered from these previous studies; however, fundamental questions are still debated, such as the frequency ranges over which average rate and temporal information are used to encode frequency and how a system in which the majority of AN fibers have a small dynamic range can encode sound level over a broad dynamic range.

1.2 Auditory Encoding of Frequency. The physiological mechanism of auditory frequency encoding has been a disputed issue since (at least) the 1940s, when spectral (place) cues (von Helmholtz, 1863) were challenged by temporal cues (Schouten, 1940; Wever, 1949). The task of pure-tone frequency discrimination has been widely used to test rate-place and temporal hypotheses for frequency encoding by measuring the just-noticeable-difference (JND) in frequency between two tones (see Moore, 1989, and Houtsma, 1995, for reviews of human performance; also see Figure 4). Rate-place schemes are based on the frequency selectivity (or tuning) of AN fibers (Kiang et al., 1965; Patuzzi & Robertson, 1988). Two tones of differing

[1] Fitzhugh (1958) previously used this approach to investigate visual psychophysical performance based on the random nature of optic nerve responses.

frequency excite different AN fibers, and thus the listener can discriminate these tones by detecting differences in the sets of AN fibers that are activated. Temporal schemes are based on the ability of AN fibers to phase-lock in response to tones, that is, AN discharges tend to occur at a particular phase of the tone (Kiang et al., 1965; Johnson, 1980; Joris, Carney, Smith, & Yin, 1994).

It is generally believed that average-rate information must be used at high frequencies (at least above 4–5 kHz; Wever, 1949; Moore, 1973, 1989; Dye & Hafter, 1980; Wakefield & Nelson, 1985; Javel & Mott, 1988; Pickles, 1988; Moore & Glasberg, 1989), because of the rolloff in phase locking above 2–3 kHz (Johnson, 1980; Joris et al., 1994; or see Figure 1c). For low frequencies, there is continuing debate as to whether rate-place or temporal information is used to encode frequency (Pickles, 1988; Javel & Mott, 1988; Moore, 1989; Cariani & Delgutte, 1996a, 1996b; Kaernbach & Demany, 1998). The assumption that there is no temporal information above 4–5 kHz has been used to interpret numerous psychophysical experiments and develop theories for the encoding of sound (Moore, 1973; Viemeister, 1983). However, no study has accurately quantified the total amount of temporal information in the AN at high frequencies or compared the trends in rate-place and temporal information versus frequency at high frequencies using the same AN model. Siebert (1970) used an analytical AN model and SDT to evaluate the ability of temporal and rate-place information to account for human frequency discrimination. His model included a nonstationary Poisson process with a time-varying discharge rate that described the AN fibers' phase locking to low-frequency tonal stimuli; however, the model did not include rolloff in phase locking above 2–3 kHz, and thus Siebert's predictions were limited to low frequencies. Siebert (1970) found that temporal information significantly outperformed rate-place information at low frequencies, but that optimal performance based on rate-place is more similar to levels of human performance. Goldstein and Srulovicz (1977) found that when a simple description of the rolloff in phase locking at high frequencies was included in Siebert's (1970) AN model, predicted performance based on temporal information (in terms of $\Delta f/f$) became worse as frequency increased above 1 kHz, similar to human performance (Moore, 1973). However, the description of the lowpass rolloff in phase locking used by Goldstein and Srulovicz (1977) had a cutoff frequency that was too low (600 Hz rather than 2500 Hz) and a slope that was too shallow (40 dB/dec rather than 100 dB/dec) when compared to more recent data from cat (see Figure 1c and Johnson, 1980; Weiss & Rose, 1988). In addition, Goldstein and Srulovicz (1977) quantified performance based on only a small set of AN fibers with CF equal to the tone frequency, and thus it is not possible from their predictions to know whether there is significant temporal information at high frequencies or to calculate performance bounds.

Another significant result from Siebert (1970) was that optimal performance based on all information improved much too quickly as a function of duration compared to human performance, at least for long durations.

Siebert concluded that human listeners are severely limited in their ability to use temporal information in AN discharges and that there is sufficient rate-place information to account for human performance if this information were used efficiently. Goldstein and Srulovicz (1977) demonstrated that a restricted temporal scheme based on only the intervals between adjacent discharges (first-order intervals) corrected the duration dependence at long durations, and they concluded that temporal schemes should not be ruled out. However, both of these analytical models are based on steady-state responses and therefore are limited to long-duration stimuli due to the absence of onset and offset responses as well as neural adaptation.[2] This limitation is significant because the slope of the dependence of human frequency-discrimination performance on duration has been shown to differ for long- and short-duration stimuli (Moore, 1973; Freyman & Nelson, 1986).

1.3 Auditory Encoding of Level. The encoding of level in the auditory system is an interesting problem given the discrepancy between the wide dynamic range of human hearing (at least 120 dB; Viemeister and Bacon, 1988) and the limited dynamic range of the majority of AN fibers (less than 30 dB; May & Sachs, 1992) (see Viemeister, 1988a, 1988b, for reviews of this "dynamic-range problem"). For many types of stimuli, the JND in level is essentially constant across a wide range of levels. The observation of constant performance across level (i.e., constant JND in decibels or, equivalently, that the smallest detectable change in intensity is proportional to intensity) is referred to as Weber's law. Weber's law is observed for broadband noise (Miller, 1947) and for narrow-band signals in the presence of band-reject noise (Viemeister, 1974, 1983; Carlyon & Moore, 1984). However, for narrow-band signals in quiet where spread of excitation across the population of AN fibers is possible, performance improves slightly as a function of level, an effect referred to as the "near-miss" to Weber's law (McGill & Goldberg, 1968; Rabinowitz, Lim, Braida, & Durlach, 1976; Jesteadt, Wier, & Green, 1977; Florentine, Buus, & Mason, 1987). Siebert (1965, 1968) demonstrated that a rate-place encoding scheme based on a population of AN fibers with low thresholds and a limited dynamic range (i.e., the majority of AN fibers; Liberman, 1978) can produce Weber's law for pure-tone level discrimination through spread of excitation. However, Siebert's (1965, 1968) model did not predict the near-miss to Weber's law for pure tones and would not be expected to predict Weber's law for broadband noise or narrowband stimuli in band-reject noise. Other studies have implicated the role of a small population of AN fibers that have higher thresholds and broader dynamic ranges (Liberman, 1978) in producing Weber's law for conditions in which spread

[2] An analytical model that included transient responses was used by Duifhuis (1973) to evaluate the consequences of peripheral filtering on nonsimultaneous masking; however, his model did not include neural adaptation.

of excitation is assumed to be restricted (Colburn, 1981; Delgutte, 1987; Viemeister, 1988a, 1988b; Winslow & Sachs, 1988; Winter & Palmer, 1991). However, when models have been used to quantify the total information available in a restricted CF region (i.e., assuming no spread of excitation) with physiologically realistic distributions of threshold and dynamic range, performance degrades as level increases above 40 dB SPL (Colburn, 1981; Delgutte, 1987), inconsistent with Weber's law and with trends in human performance. In addition, no model has accurately quantified the effect of noise maskers on the spread of rate or temporal information across the population of AN fibers; instead simple assumptions for the influence of the noise maskers have been used.

1.4 Combining Computational Models with Signal Detection Theory. There are now computational AN models (Payton, 1988; Carney, 1993; Giguère & Woodland, 1994a, 1994b; Patterson, Allerhand, & Giguère, 1995; Robert & Eriksson, 1999; Zhang, Heinz, Bruce, & Carney, 2001) that can provide more accurate physiological responses over a wider range of stimulus parameters than the analytical models already described. In addition, computational models can simulate AN responses to arbitrary stimuli and thus can be used to study a wider range of psychophysical tasks, such as those involving noise maskers. This article focuses on extension of the SDT approach to allow the use of computational models to predict psychophysical performance limits for auditory discrimination based on information encoded in the stochastic AN discharge patterns. Several studies have used SDT to relate computational auditory models to psychophysical performance (Dau, Püschel, & Kohlrausch, 1996a, 1996b; Dau, Kollmeier, & Kohlrausch, 1997a, 1997b; Gresham & Collins, 1998; Huettel & Collins, 1999). Dau et al. (1996a, 1996b, 1997a, 1997b) developed computational models of effective auditory processing with the goal of matching predicted and human performance (i.e., in terms of both absolute values and trends for various stimulus parameters). Their auditory models are physiologically motivated but are not intended to describe the processing at specific locations in the auditory pathway, and therefore are not compared directly to physiological responses. Gresham and Collins (1998) and Huettel and Collins (1999) used SDT to evaluate information loss at different stages of several more physiologically based computational auditory models. Psychophysical performance was limited in their analysis only by the random variability associated with the noise stimulus (i.e., their analysis did not include any form of internal physiological noise). These previous studies using computational auditory models and SDT have not taken into consideration the fact that information is encoded in the population of AN fibers in terms of neural discharges, that is, stochastic all-or-none events that represent a severe transformation of information from the continuous waveforms analyzed in previous studies using computational models with SDT. In contrast, Siebert (1965, 1968, 1970) and Colburn's (1969, 1973) approach quantified the effect of neural

encoding on AN population information, but was limited to analytical AN models. This and the companion study describe a general method that combines and extends previous approaches by quantifying the information in the discharge patterns of a specific neural population of AN fibers using physiologically based computational models.

The potential of rate-place (based on AN discharge counts) and all-information (based on AN discharge times) encoding schemes to account for human performance is characterized for both frequency and level discrimination. The adjective *all-information* is used to emphasize the fact that this encoding scheme makes use of all the information in the distribution of all discharges times across the population of AN fibers and includes average-rate information. The contribution of temporal information in the responses can be inferred by comparing the predictions of the rate-place and all-information models. Note that the all-information model does not assume any specific form of temporal processing, such as calculating synchrony coefficients or creating interval histograms. All-information predictions thus provide an absolute limit on achievable performance given the total information available in AN discharge patterns.

It is critical, before proceeding to more complex computational AN models, to verify the computational method by comparing it to previous analytical model predictions (Siebert, 1965, 1968, 1970). Thus, optimal performance predictions were made for pure-tone frequency and level discrimination as a function of frequency, level, and duration based on a relatively simple computational AN model. This simple AN model does not include several complex physiological response properties that have been hypothesized to influence auditory frequency and level discrimination, and thus these limitations are discussed in terms of the conclusions about auditory encoding that can be made from this study. The general agreement between computational and analytical predictions provides confidence that the computational method is valid. Previous studies have typically focused on either frequency or level discrimination, have often examined only a limited parameter range, and have used different AN models. By using the same AN model to make both rate-place and all-information predictions for both frequency and level discrimination versus frequency, level, and duration, this study provides a unified description of previously suggested constraints on rate-place and temporal models. This study also quantifies that there is significant temporal information in the AN for frequency discrimination up to at least 10 kHz, and that all-information performance demonstrates the same trends versus frequency as human performance, whereas rate-place does not. These new predictions are important to theories of frequency encoding in the auditory system because they contradict the generally held belief that the rolloff in AN phase locking eliminates all temporal information at high frequencies.

Thus, this study and the companion study demonstrate several benefits of using computational neural models to predict psychophysical per-

formance limits and illustrate the types of general neural encoding questions that can be addressed by quantifying the total information (based on discharge times or counts) in a specific neural population. In addition to the general applicability of this quantitative approach to theoretical neuroscience, several engineering applications could potentially benefit from the use of this approach to study basic auditory encoding issues. Audio coding algorithms for speech or music are often based on auditory masking phenomena that can be studied using this general approach (Heinz, 2000). In addition, the discriminability of the original and encoded versions of an audio sample could be evaluated based on AN information with this method. Hearing aid algorithms could be developed and tested by quantifying the effect of the algorithm on AN information in the impaired auditory system. This method could be used to develop optimal processors for specific types of auditory information, which could be useful in algorithms for speech recognition or blind-source separation.

2 General Methods

2.1 Computational Auditory Nerve Model. The type of computational AN model for which this study was designed 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 that have been compared directly to extensive physiological AN response properties have been described by Payton (1988), Carney (1993), and Robert and Eriksson (1999). A simplified version of the Carney (1993) model was used in this study to provide a better comparison to the analytical AN model used by Siebert (1970) and to provide a simple comparison for future predictions with more complex AN models.[3]

Figure 1 shows pure-tone response properties of the computational AN model that are important for this study (see Ruggero, 1992, for a review of basic AN response properties). Table 1 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 the population response in Figure 1a). Model filter bandwidths were based on estimates of human bandwidths from the psychophysical notched-noise method for estimating auditory filter shapes (Glasberg & Moore, 1990; see Table 1). Psychophysically measured filters have been shown to match AN frequency tuning when both were measured in guinea pig (Evans, Pratt, Spenner, & Cooper, 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; see

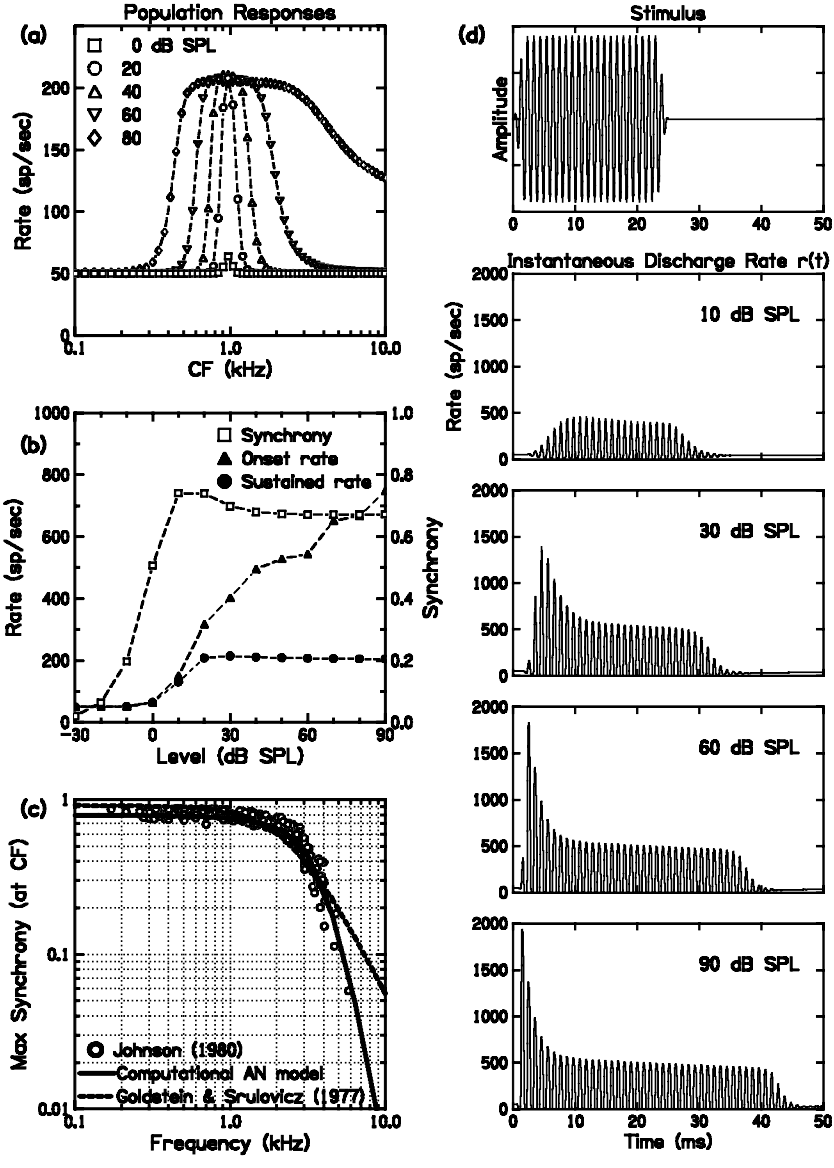
[3] Code for the AN model used in this study is available online at http://earlab.bu.edu/.

Mountain & Hubbard, 1996, for a review). The saturating nonlinearity contributes to the limited dynamic range of AN fibers (see Figure 1b). (Note that the dynamic range is also affected by the IHC-AN synapse; see below and Patuzzi & Robertson, 1988.) All AN model fibers had a rate threshold of roughly 0 dB SPL, a spontaneous rate of 50 spikes per second, and a maximum sustained rate of roughly 200 spikes per second. The model dynamic range for the sustained rate of these high-spontaneous-rate (HSR) fibers was roughly 20–30 dB, while the dynamic range for onset rate was much larger (see Figure 1b). The HSR fibers described by the AN model represent the majority (61%) of the total AN population (Liberman, 1978). Medium- (23%) and low-SR (16%) fibers, which have higher thresholds and larger dynamic ranges, are not described by our model. The synchrony-level curves had a threshold 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 this study is the rolloff in phase locking as frequency increases above 2–3 kHz (see Figure 1c; 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 per 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 rolloff in phase locking had a 3 dB cutoff frequency near 2500 Hz and $\sim$ 100 dB/decade rolloff in the frequency range 4–6 kHz (see Figure 1c). The model synchrony coefficients above 5 kHz are a simple extrapolation of the physiological data, consistent with the lowpass shape and slope across many species reported by Weiss and Rose (1988). For comparison, a typical description of synchrony rolloff used in analytical AN models is also shown (Goldstein & Srulovicz, 1977) and is seen to underestimate the slope at high frequencies.

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 (Lin & Goldstein, 1995; and see Figure 3 of Westerman & 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 1). 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 (Westerman & Smith, 1988). The output of the model (see Figure 1d) represents the instantaneous discharge rate $r(t, f, L)$ of an individual high-spontaneous-rate, low-threshold AN fiber.

2.2 Signal Detection Theory. The application of SDT to evaluate auditory discrimination performance limits in this study is an extension of Siebert's (1970) study of frequency discrimination (see Figure 2). A model of the auditory periphery describes the time-varying discharge rate $r_i(t, f, L)$ of the ith AN fiber in response to a tonal stimulus of level L, frequency f,



and duration T. In order to calculate performance limits based on the activity in the AN, the information in all fibers in the AN population must be considered, that is, 30,000 total AN fibers (Rasmussen, 1940) spanning the entire CF range (20 Hz–20 kHz; Greenwood, 1990). The total information in the AN population depends on how the stochastic activity in each AN fiber varies with the stimulus and on the statistical relations between AN fibers. A Poisson process provides a good description of the random nature of AN discharges (e.g., exponential interval distribution, and independent successive intervals; Colburn, 1969, 1973; Siebert, 1970). All AN fibers in the population were assumed in this study to have independent discharge-generating mechanisms, including the 10–30 AN fibers that innervate each IHC (Spoendlin, 1969; Liberman, 1980; Ryugo, 1992). This assumption implies that the stochastic discharge activity of any two AN fibers will be independent when the effects due to a common stimulus drive have been removed and is consistent with evidence to date from studies that have investigated correlated AN activity (Johnson & Kiang, 1976; Young & Sachs, 1989; Kiang, 1990). For the deterministic tonal stimuli used in this study, the independence assumption appears reasonable.

Two types of analyses are considered in this study. *Rate-place* predictions are based on the assumption that the set of AN discharge counts over the duration of the stimulus, $\{K_i\}$, is the only information used by the listener. Thus, the rate-place analysis (left branch of Figure 2) is based on a set of independent, stationary Poisson processes with rates equal to the average discharge rates $r_i(f, L)$ of the AN-model fibers (Siebert, 1968). In contrast, the *all-information* predictions (right branch of Figure 2) are based on the assumption that the set of observations are the population of discharge times and counts $\{t^i_1, \ldots, t^i_{K_i}\}$ produced by a set of independent, nonstation-

Figure 1: *Facing page*. Basic pure-tone response properties of the computational AN model. (a) Sustained-rate population responses for the 60 AN-model CFs 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). Dashed line is from analytical model used by Goldstein and Srulovicz (1977). (d) Stimulus waveform and instantaneous discharge rates $r(t, f, L)$ for a 970 Hz fiber in response to $f = 970$ Hz, 25 ms (2 ms rise/fall) tones over a range of levels (L).

Table 1: Equations and Parameters Used to Implement 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	
	$ihc[k] = \{\arctan(K \cdot gtf[k] + \beta) - \arctan(\beta)\}/\{\pi/2 - \arctan(\beta)\}$	
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)	$= P_I[k]C_I[k]$

Notes: [a] Greenwood (1990).
[b] See Patterson, Nimmo-Smith, Holdsworth, and Rice (1987); Glasberg & Moore (1990); Carney (1993).
[c] See Westerman and Smith (1988); Carney (1993).

ary Poisson processes with time-varying rates $r_i(t, f, L)$ provided by the AN model (Siebert, 1970). Descriptions of the probability density of the observations, $p(t_1^i, \ldots, t_{K_i}^i \mid r_i)$, are used to evaluate the optimal performance achievable by any decision process (see section 3). Optimal performance based on

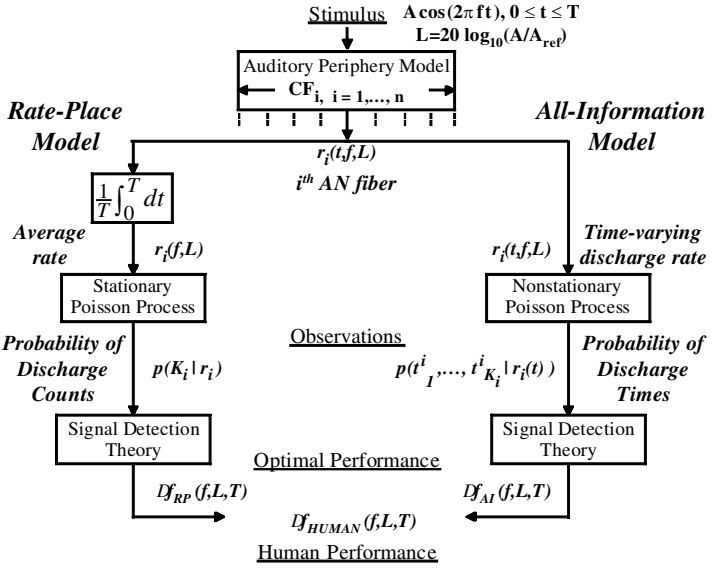


Figure 2: Overview of signal detection theory (SDT) used with stochastic models of the auditory periphery to evaluate fundamental limits of performance on a psychophysical task (frequency discrimination shown).

rate-place, $\Delta f_{RP}(f, L, T)$, and all-information, $\Delta f_{AI}(f, L, T)$, is compared directly to human JNDs measured psychophysically, $\Delta f_{HUMAN}(f, L, T)$.

It is worthwhile, as Siebert (1965, 1968, 1970) has done, to discuss the potential outcomes of a comparison between human and optimal performance based on a particular type of neural information (e.g., rate-place or all-information) in the AN. The best possible outcome is if human performance matches optimal performance in both absolute value and the trends versus stimulus parameters. This result would suggest that human performance is essentially determined by the peripheral transformation from acoustic waveform to action potentials in the AN and that the central nervous system can be viewed roughly as an optimal processor for this task. Another straightforward result is if human performance is superior to optimal performance. This result suggests that the information represented by the model of peripheral transformations is inadequate to account for human performance. A more likely result is that optimal performance is superior to human performance. In this case, the strongest conclusion that can be drawn is that there is sufficient information in the AN discharge patterns to account for human performance. However, if optimal performance is uniformly better than human performance (i.e., parallel to human

performance across the entire range of stimulus parameters), then a parsimonious hypothesis is that the AN information represented by the neural model determines human performance, but is used in an inefficient manner that uniformly degrades optimal performance. On the other hand, if optimal performance exceeds human performance in a nonuniform way and there is no realistic processor that would be nonuniformly inefficient in the required manner, then a fairly strong hypothesis can be made that although the information provided to the optimal detector is sufficient, this information is not likely to account solely for human performance.

3 General Analytical Results: Performance Limits for One-Parameter Discrimination

3.1 Overview of Basic Result. The results in this section are presented in terms of discrimination of a single stimulus parameter α, which is either frequency f (Hz) or level L (dB SPL) in this study. A measure $(\delta'_\alpha[CF])^2$, which is closely related to the sensitivity measure d' from SDT (Green & Swets, 1966), is used to represent the information from an individual AN fiber with characteristic frequency CF, where δ'_α represents the sensitivity per unit α (see section 3.2).[4] Based on the assumption of independent AN fibers (see section 2.2), the total information from the population is the sum of the information from individual AN fibers. The JND in α, $\Delta\alpha_{JND}$, is inversely related to the amount of information available from the AN population and can be calculated as

$$\Delta\alpha_{JND} = \left\{ \sum_i \left(\delta'_\alpha[CF_i]\right)^2 \right\}^{-\frac{1}{2}}$$

$$= \left\{ \sum_i \int_0^T \frac{1}{r_i(t,\alpha)} \left[\frac{\partial r_i(t,\alpha)}{\partial \alpha} \right]^2 dt \right\}^{-\frac{1}{2}}, \tag{3.1}$$

as described in section 3.2. Equation 3.1 describes the optimal performance possible from all information available in the AN in terms of the time-varying discharge rates $r_i(t, \alpha)$ of the population of AN fibers. Optimal rate-place performance is calculated by using the average discharge rate $r_i(\alpha)$ in equation 3.1 rather than $r_i(t, \alpha)$. Equation 3.1 is extremely general because it is applicable to any single-parameter discrimination experiment,

[4] Durlach and Braida (1969) used the metric δ' to represent sensitivity per bel for intensity discrimination [i.e., $\delta' = d'(I_1, I_2)/\log_{10}(I_2/I_1)$] (see also Braida & Durlach, 1988). In our study, δ'_α is used to represent sensitivity per unit α, where α can be level L in dB, or frequency f in Hz. Thus, the normalized sensitivity metric δ'_L differs from that used by Durlach and Braida (1969) only in the unit of level used (i.e., δ'_L represents sensitivity per decibel).

and it can be used with any neural model (e.g., analytical or computational) that describes the transformation from sensory stimulus to neural response and assumes statistically independent Poisson discharges.

3.2 Mathematical Analysis. A complete analysis of a general one-parameter discrimination experiment based on stochastic neural responses is described below using SDT and extends several previous results (e.g., Siebert, 1970; Colburn, 1981; Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997). Prediction of performance limits based on the Cramér-Rao bound is described first, because this was the method first described by Siebert (1970); however, this approach lacks an explicit description of the processing that is required to achieve these performance limits. The use of a likelihood ratio test, which is described second, provides insight into the meaning of the Cramér-Rao bound result, describes the form of an optimal processor, and demonstrates that the Cramér-Rao bound can be met with equality in this case (i.e., an efficient estimator exists; van Trees, 1968).

The assumption that a nonstationary Poisson process is a good model for the AN discharge patterns allows the random nature of the observations to be described (see Parzen, 1962; Snyder & Miller, 1991; Rieke et al., 1997). The joint probability density of the *unordered* discharge times $\{t^i_1, \ldots, t^i_{K_i}\}$ on the ith AN fiber is given by

$$p(t^i_1, \ldots, t^i_{K_i} \mid \alpha) = \frac{\prod_{n=1}^{K_i} r_i(t^i_n, \alpha)}{K_i!} \exp\left[-\int_0^T r_i(t, \alpha)\, dt\right], \tag{3.2}$$

where $r_i(t, \alpha)$ is the time-varying discharge rate, and α is the stimulus parameter to be discriminated. (Note that this formula applies to unordered discharge times; i.e., the times are *not* constrained such that $t^i_1 < t^i_2 < \cdots < t^i_{K_i}$.)

3.2.1 Cramér-Rao Bound. The Poisson description of the random nature of the observations can be used in the Cramér-Rao bound (Cramér, 1951; van Trees, 1968) to describe performance limits for estimating the value of α based on the response of a single AN fiber in terms of its time-varying discharge rate $r_i(t, \alpha)$. The variance $\sigma^2_{\hat{\alpha}}[i]$ of any unbiased estimate of the parameter α from the observed discharge times on the ith fiber is bounded by the Cramér-Rao inequality,

$$\frac{1}{\sigma^2_{\hat{\alpha}}[i]} \le E\left\{\left[\frac{\partial}{\partial \alpha} \ln p(t^i_1, \ldots, t^i_{K_i} \mid \alpha)\right]^2\right\}, \tag{3.3}$$

$$= \sum_{K_i=0}^{\infty} \int_0^T \cdots \int_0^T \left[\frac{\partial}{\partial \alpha} \ln p(t^i_1, \ldots, t^i_{K_i} \mid \alpha)\right]^2 \times p(t^i_1, \ldots, t^i_{K_i} \mid \alpha) dt^i_1 \cdots dt^i_{K_i}, \tag{3.4}$$

where the expectation in equation 3.3 is over the random observations $\{t_1^i, \ldots, t_{K_i}^i\}$ and represents the Fisher information in the random discharge times about α. This inequality provides a lower bound on the variance of any unbiased estimator by relating the accuracy of the estimate to the rate of change with α of the likelihood of the observations. The variance $\sigma_{\hat{\alpha}}^2[i]$ can be related to the time-varying discharge rate $r_i(t, \alpha)$ by substituting the joint probability density from equation 3.2 into equation 3.4 and performing some extensive simplifications to obtain

$$\frac{1}{\sigma_{\hat{\alpha}}^2[i]} \leq \int_0^T \frac{1}{r_i(t, \alpha)} \left[\frac{\partial r_i(t, \alpha)}{\partial \alpha}\right]^2 dt. \tag{3.5}$$

Equation 3.5 provides a lower bound on the variance of any unbiased estimate of α based on the observations from a single AN fiber in terms of the time-varying discharge rate $r_i(t, \alpha)$.

The JND $\Delta\alpha_{JND}$ of the ideal observer is equal to the minimum standard deviation of any estimator based on the observations, when threshold is defined as 75% correct in a two-interval, two-alternative, forced-choice task (Green & Swets, 1966; Siebert, 1968, 1970). This definition of threshold corresponds to a sensitivity $d' = 1$, where $d' = \Delta\alpha_{JND}/\sigma_{\hat{\alpha}}[i]$, and produces predicted JNDs that are directly comparable to JNDs measured psychophysically. The JND based on the population of AN fibers is given by equation 3.1, because the information from each AN fiber adds under the assumption that the discharge patterns of all fibers are statistically independent (see section 2.2).

3.2.2 Likelihood Ratio Test. This analysis extends previous results from Colburn (1981) and Rieke et al. (1997). It is shown that the performance of a sufficient statistic for a nonstationary Poisson process derived from a likelihood ratio test (LRT) meets the Cramér-Rao bound with equality (i.e., that an efficient estimator exists). This alternative approach provides more intuition in the present case and is therefore more accessible than the Cramér-Rao bound analysis; however, the LRT approach does not always lead to a form of an optimal processor that is simple enough to allow the evaluation of performance analytically, in which case the Cramér-Rao bound may be more useful (see the companion article in this issue). In the present case, the LRT analysis shows that the quantity

$$(\delta'_\alpha[CF_i])^2 \triangleq \int_0^T \frac{1}{r_i(t, \alpha)} \left[\frac{\partial r_i(t, \alpha)}{\partial \alpha}\right]^2 dt \tag{3.6}$$

represents the information available on the ith AN fiber for discriminating α, where $\delta'_\alpha[CF_i]$ represents the normalized sensitivity (see note 4) of the ith fiber, and is defined as the sensitivity d' per unit α (see Durlach & Braida, 1969; Braida & Durlach, 1988).

The form of the optimal processor for discriminating between α and $\alpha + \Delta\alpha$ can be derived from a log-likelihood test,

$$\ln \frac{p(t_1^i, \ldots, t_{K_i}^i | \alpha + \Delta\alpha)}{p(t_1^i, \ldots, t_{K_i}^i | \alpha)} \underset{\alpha}{\overset{\alpha+\Delta\alpha}{\gtrless}} 0, \tag{3.7}$$

where the threshold of 0 corresponds to the minimum-probability-of-error criterion with equal a priori probabilities (van Trees, 1968, pp. 23–30). Substituting the joint probability density of the discharge times, equation 3.2, into equation 3.7 and simplifying, one obtains the following form of an optimal test,

$$\begin{aligned} Y(t_1^i, \ldots, t_{K_i}^i) &\triangleq \sum_{n=1}^{K_i} \ln \frac{r_i(t_n^i, \alpha + \Delta\alpha)}{r_i(t_n^i, \alpha)} \\ &\underset{\alpha}{\overset{\alpha+\Delta\alpha}{\gtrless}} \int_0^T [r_i(t, \alpha + \Delta\alpha) - r_i(t, \alpha)]\, dt, \end{aligned} \tag{3.8}$$

where the decision variable Y depends on the observed data, and the right side of the comparison in equation 3.8 is a criterion based on a priori information.

The sensitivity metric d' can be used to characterize performance completely if the decision variable Y follows a gaussian distribution and has equal variance under both hypotheses, α and $\alpha + \Delta\alpha$ (Green & Swets, 1966; van Trees, 1968). These two conditions can be reasonably assumed to hold in this analysis. Since the decision variable Y is a sum of random variables, a gaussian assumption is reasonable based on the central limit theorem if K_i (the number of discharges on the ith fiber) is large. Even for conditions in which K_i is small (e.g., a 25 ms stimulus and an average discharge rate of 200 spikes per second), the gaussian assumption is reasonable for the total population decision variable (i.e., based on all 30,000 AN fibers; see below and section 4). The total population decision variable is a linear combination of the individual fiber decision variables based on the independence assumption. The variance of the decision variable Y is shown below to depend on α, and thus the variances under the two hypotheses are not strictly equal. However, for the small values of $\Delta\alpha$ associated with optimal JNDs, the equal-variance assumption is reasonable and makes the performance analysis significantly more simple. Thus, the d' metric will be used to characterize performance of the optimal detector $Y(t_1^i, \ldots, t_{K_i}^i)$, where

$$\left(d'[CF_i]\right)^2 = \frac{\{E[Y(t_1^i, \ldots, t_{K_i}^i) \mid \alpha + \Delta\alpha] - E[Y(t_1^i, \ldots, t_{K_i}^i) \mid \alpha]\}^2}{Var[Y(t_1^i, \ldots, t_{K_i}^i) \mid \alpha]}. \tag{3.9}$$

It can be shown (e.g., see Rieke et al., 1997) that for any decision variable Y of the form

$$Y(t_1, \ldots, t_K) = \sum_{n=1}^{K} g(t_n), \tag{3.10}$$

where $g(t_n)$ is any function of the Poisson discharge time t_n, the expected value and variance of Y are given by

$$E[Y(t_1, \ldots, t_K)|\alpha] = \int_0^T g(t) r_i(t, \alpha)\, dt, \tag{3.11}$$

$$Var[Y(t_1, \ldots, t_K)|\alpha] = \int_0^T g^2(t) r_i(t, \alpha)\, dt. \tag{3.12}$$

Given these relations with $g(t) = \ln[r_i(t, \alpha + \Delta\alpha)/r_i(t, \alpha)]$,

$$\left(d'[CF_i]\right)^2 = \left\{\int_0^T \ln\frac{r_i(t, \alpha + \Delta\alpha)}{r_i(t, \alpha)} [r_i(t, \alpha + \Delta\alpha) - r_i(t, \alpha)]\, dt\right\}^2 \Bigg/ \left\{\int_0^T \left[\ln\frac{r_i(t, \alpha + \Delta\alpha)}{r_i(t, \alpha)}\right]^2 r_i(t, \alpha)\, dt\right\}. \tag{3.13}$$

With the additional assumption that $r_i(t, \alpha)$ varies linearly and slowly over the incremental range from α to $\alpha + \Delta\alpha$ (Colburn, 1981), that is,

$$r_i(t, \alpha + \Delta\alpha) \simeq r_i(t, \alpha) + \dot{r}_i(t, \alpha)\Delta\alpha, \tag{3.14}$$

where $\dot{r}_i(t, \alpha) = \partial r_i(t, \alpha)/\partial\alpha$, one obtains

$$\left(d'[CF_i]\right)^2 \simeq \left\{\int_0^T \ln\left[1 + \frac{\dot{r}_i(t, \alpha)\Delta\alpha}{r_i(t, \alpha)}\right] [\dot{r}_i(t, \alpha)\Delta\alpha]\, dt\right\}^2 \Bigg/ \left(\int_0^T \left\{\ln\left[1 + \frac{\dot{r}_i(t, \alpha)\Delta\alpha}{r_i(t, \alpha)}\right]\right\}^2 r_i(t, \alpha)\, dt\right). \tag{3.15}$$

The approximation $\ln(1 + x) \simeq x$ for small x then leads to the equations[5]

$$\left(d'[CF_i]\right)^2 \simeq \left[(\Delta\alpha)^2 \int_0^T \frac{\dot{r}_i(t, \alpha)^2}{r_i(t, \alpha)}\, dt\right]^2 \Bigg/ \left[(\Delta\alpha)^2 \int_0^T \frac{\dot{r}_i(t, \alpha)^2}{r_i(t, \alpha)}\, dt\right]$$

[5] Note that the "small-x" approximation may not be appropriate for low discharge rates $r_i(t, \alpha)$. To avoid this potential problem (e.g., in the valleys of the phase-locked response), a constant was added to all discharge rates $r_i(t, \alpha)$ prior to the SDT analysis in this study. A value of $\simeq 7$ spikes per second was used, based on the minimum value of $r_i(t, f, L)$ in Siebert's (1970) model. The precise value of this constant is not significant, as long as the instantaneous discharge rate does not go to zero.

$$= (\Delta\alpha)^2 \int_0^T \frac{1}{r_i(t,\alpha)} \left[\frac{\partial r_i(t,\alpha)}{\partial \alpha}\right]^2 dt \tag{3.16}$$

$$= (\Delta\alpha)^2 \left(\delta'_\alpha[CF_i]\right)^2. \tag{3.17}$$

Finally, taking $d' = 1$ as the JND and combining optimally over independent fibers, one obtains equation 3.1. The summation over fibers in equation 3.1 results from the independence assumption (see section 2.2) and comes directly in the LRT analysis for the population response since the probability function in equation 3.2 is a product of the individual fiber probabilities. In this case, the total population decision variable is a weighted combination of the individual fiber decision variables given in equation 3.8. The equivalence of the equations derived using the Cramér-Rao bound and LRT analyses demonstrates that the Cramér-Rao bound provides a tight bound on achievable performance in this case (i.e., an efficient estimator exists) and thus describes optimal performance.

4 Computational Methods: Use of Auditory Nerve Models

The time-varying discharge rates $r_i(t, \alpha)$ from the computational AN model can be substituted into equation 3.1 to evaluate optimal performance for various stimulus conditions. However, the partial derivative of the rate waveform with respect to α must be calculated as a function of time. The partial derivative was approximated from the responses of the computational AN model, for example,

$$\frac{\partial r_i(t,\alpha)}{\partial \alpha} \simeq \frac{r_i(t,\alpha+\Delta\alpha) - r_i(t,\alpha)}{\Delta\alpha}. \tag{4.1}$$

An approximation of the partial derivative with respect to frequency using $\Delta f = 0.0001$ Hz is illustrated in Figure 3. The top panel shows the two stimuli (although only one waveform is apparent due to the small Δf), and the middle panel illustrates the two time-varying discharge rate waveforms from the AN model fiber. The bottom panel represents the approximation of the partial derivative squared as a function of time. The distinct growth of the derivative squared as time increases up to 30 ms (bottom panel of Figure 3) is a result of the increasing difference in the phase between the two tones as a function of time. Even if a random phase were used, as in some psychophysical experiments, the ability to distinguish between a difference in phase and a difference in frequency increases with duration. The bimodal shape within each stimulus period of the derivative squared arises because the two rate waveforms differ in both the upward- and downward-going portions of each response period. Values of $\Delta f = 0.0001$ Hz and $\Delta L = 0.0001$ dB were used in equation 4.1 to approximate the partial derivatives with respect to frequency and level. These values were chosen by verifying that

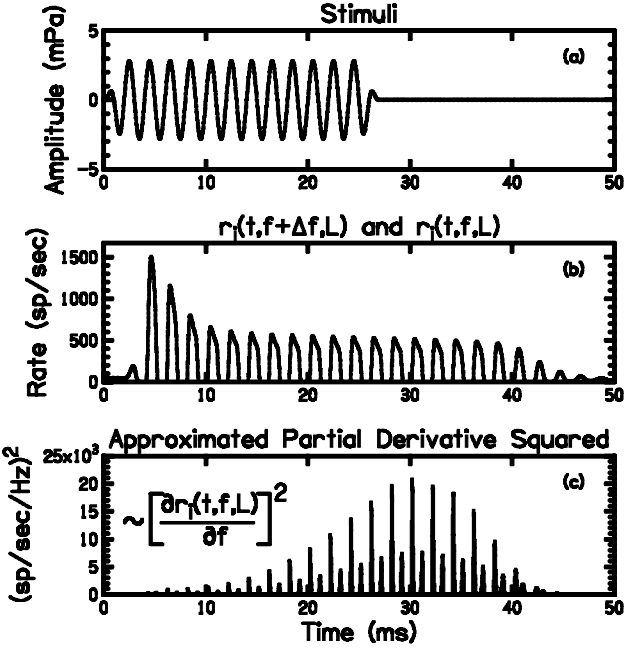


Figure 3: Approximation of the partial derivative in equation 3.1 for evaluating optimal frequency-discrimination performance for the stimulus condition: L = 40 dB SPL, f = 500 Hz, and T = 25 ms (2 ms rise/fall). Two tonal stimuli differing in frequency by $\Delta f = 0.0001$ Hz are shown in the top panel. The middle panel shows the two time-varying rate waveforms from the computational AN model (CF = 500 Hz) in response to the two stimuli. The two functions cannot be distinguished visually in the top and middle panels. The bottom panel illustrates the partial derivative squared, as approximated by equation 3.18.

the derivatives for a variety of conditions were unaffected by increases in Δf or ΔL by an order of magnitude.

All predictions of performance limits using the computational AN model (see section 2.1) are based on the total high-spontaneous-rate (HSR), low-threshold AN fiber population, which makes up 61% of the AN population in cat (Liberman, 1978). The total HSR fiber population response was simulated using 60 model CFs, which ranged from 100 Hz to 10 kHz and were uniformly spaced in location according to a human cochlear map (Greenwood, 1990; see Table 1).[6] The model CF range represents two orders of magnitude, while the commonly accepted human CF range (20 Hz–20 kHz;

[6] The choice of 60 model CFs represents a (somewhat arbitrary) compromise between a dense sampling of the frequency range and reduced computation time.

Greenwood, 1990) represents three orders of magnitude. Thus, roughly two-thirds of the 30,000 total AN fibers in human (Rasmussen, 1940) are represented by the model CF range. The total number of HSR fibers represented by the 60 model CFs was therefore 12,200, and thus roughly 200 independent AN fibers were assumed to be driven by each of the 60 model responses.

Stimulus conditions were chosen based on psychophysical experiments of interest,[7] and a sampling rate of 500 kHz was used.[8] Stimulus duration was defined to be the duration between half-amplitude points on the stimulus envelope. All rise-fall ramps were generated from raised cosine functions. The temporal analysis window (see equation 3.1) included the model response beginning at stimulus onset and ending 25 ms after stimulus offset, to allow for the response delay and the transient onset and offset responses associated with AN fibers (see Figures 1d and 3) over the range of CFs and stimulus parameters used in this study.

5 Computational Results

Human and predicted optimal performance on psychophysical discrimination tasks are compared in this section in terms of the JND in either tone frequency or level. Higher JNDs correspond to worse performance in the psychophysical task and in the optimal case indicate a reduction in the amount of sensory information about changes in the given stimulus parameter. Comparisons between optimal and human performance are made in terms of both absolute values and trends versus stimulus parameters, as discussed in section 2.2.

5.1 Frequency Discrimination in Quiet. Predicted optimal frequency JNDs for the population of HSR AN fibers were calculated as a function of frequency f, level L, and duration T for the all-information and rate-place encoding schemes using equation 3.1. Figure 4 shows the frequency discrimination predictions along with predictions from Siebert's (1970) study and human performance. Rate-place and all-information predictions from

[7] Two factors occasionally led to a slight mismatch between the stimulus parameters used in this study and those that were used in the psychophysical studies. (1) Stimulus conditions were chosen to be the same for frequency- and level-discrimination predictions in order to allow for direct comparisons of model results between frequency and level discrimination. (2) Tone frequencies were always chosen to be equal to one of the 60 model CFs in order to avoid artifacts related to the spatial undersampling of the human frequency range.

[8] A sampling frequency of 500 kHz was used for all predictions, except for three high-frequency conditions for which higher sampling rates were chosen to avoid subharmonic distortion generated by the saturating nonlinearity for specific relations between stimulus and sampling frequency. In general, this sampling artifact was most prominent for high levels and high frequencies, and thus for high frequencies the predictions were limited to low to medium levels.

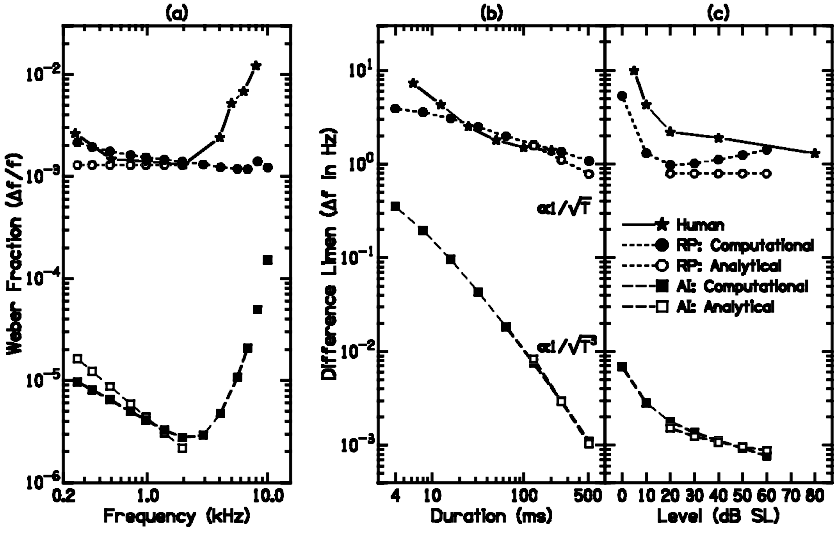


Figure 4: Optimal performance for pure-tone frequency discrimination. Predictions from the computational AN model are shown by filled circles for the rate-place (RP) scheme, and by filled squares for the all-information (AI) scheme. Predictions from Siebert's (1970) analytical model are shown for the two encoding schemes by corresponding open symbols. Typical human performance is illustrated by stars. (a) The Weber fraction $\Delta f/f$ is plotted as a function of frequency. All model predictions are for $L = 40$ dB SPL and $T = 200$ ms (20 ms rise/fall). Human data are from Moore (1973) for $T = 200$ ms and equal loudness across frequency ($L = 60$ dB SPL at 1 kHz). The two high-frequency points not connected by lines for each encoding scheme are conditions for which the upper limit of the computational model CF range influenced the results. (b) The difference limen Δf is plotted as a function of duration. Predictions are for $L = 40$ dB SPL, $f = 970$ Hz, and 4 ms rise/fall ramps. Predictions from Siebert (1970) follow $\Delta f \propto \sqrt{T^{-1}}$ for rate-place and $\Delta f \propto \sqrt{T^{-3}}$ for all-information. Human data are from Moore (1973) for $L = 60$ dB SPL, $f = 1000$ Hz and 2 ms rise/fall ramps. (c) The difference limen Δf is plotted as a function of level. Predictions were made for $f = 970$ Hz and $T = 500$ ms (20 ms rise/fall). Human data are from Wier et al. (1977) for $f = 1000$ Hz and $T = 500$ ms.

the computational AN model match predictions from Siebert well, both qualitatively and quantitatively, over the range of stimulus parameters for which Siebert's analytical model is applicable ($200 < f < 2000$ Hz, $T \geq 100$ ms, and $20 \leq L \leq 50$ dB SPL). This general agreement supports the validity of the computational approach.

The effect of frequency on pure tone frequency discrimination is shown in Figure 4a. Rate-place predictions are essentially flat as a function of fre-

quency due to the roughly constant filter shape on a logarithmic axis. The predictions from the computational AN model show a slight improvement in performance as frequency increases due to the small increase in the sharpness of tuning (quality factor $Q = CF/ERB$) as frequency increases (Table 1; Glasberg & Moore, 1990). Rate-place predictions match human performance very closely below 2 kHz but deviate significantly at high frequencies. As discussed in section 6.2.1, none of the limitations of the present simple AN model is likely to account for the large discrepancy between rate-place and human performance at high frequencies.

All-information performance is roughly two orders of magnitude better than rate-place and human performance, consistent with Siebert (1970). At low frequencies, predicted performance based on all-information improves as frequency increases, due to the increased number of stimulus cycles as frequency increases (i.e., the two tones become more out of phase by the end of the stimulus duration for higher frequencies; see the bottom panel of Figure 3). All-information predictions, when extended to high frequencies, demonstrate a sharp worsening in performance above 2–3 kHz due to the rolloff of phase locking at high frequencies included in the computational AN model (see Figure 1c). This worsening matches the general trend observed in human data (Moore, 1973; Wier, Jesteadt, & Green, 1977, Moore & Glasberg, 1989). The all-information predictions converge toward the rate-place predictions as frequency increases, as expected due to the loss of temporal information. However, all-information performance is still one order of magnitude better than rate-place performance at 10 kHz (see Figure 4a), despite low synchrony coefficients at high frequencies (see Figure 1c). The upper-CF limit in the model slightly influenced both rate-place and all-information predictions for high frequencies by underestimating the amount of information available (by up to a factor of 2), and thus overestimating the JND (by up to a factor of $\sqrt{2}$).

Performance as a function of duration is shown in Figure 4b. Computational predictions for long durations can generally be described by $\Delta f \propto \sqrt{T^{-1}}$ for rate-place and $\Delta f \propto \sqrt{T^{-3}}$ for all-information, consistent with the analytical predictions from Siebert (1970). Predicted performance from the all-information scheme improves much faster with duration than the rate-place predictions or the human data. The rapid improvement with increased duration is due to allowing the decision process to make all possible temporal comparisons throughout the entire time course of the stimulus (see the distinct growth of information with time in the bottom panel of Figure 3). Both the rate-place and all-information predictions have a slightly smaller slope at shorter durations than at longer durations, with the all-information slope matching the slope of human performance for $T \leq 16$ ms. The computational AN model included realistic onsets and offsets, as well as neural adaptation, which would be expected to be significant for shorter durations.

Frequency discrimination performance as a function of level is shown in Figure 4c. At low levels, both rate-place and all-information predictions from the computational AN model demonstrate the same sharp improvement in performance with increasing level, as seen in the human data. Above 20 dB SPL, rate-place performance is roughly flat, with the predicted JNDs from the computational AN model demonstrating a slight increase as level increases. Both all-information and human performance continue to improve slowly as level increases. The improvement with level in the all-information model is due to the increase in the number of excited fibers as level increases. The lack of improvement in rate-place performance with level results from the limited dynamic range and low thresholds of HSR fibers. Inclusion of AN fibers with higher thresholds and broader dynamic ranges may produce an improvement in rate-place performance with level, as discussed in section 6.2.1.

In order to understand the basis for predicted rate-place and all-information frequency discrimination performance, it is useful to examine the distribution of information across the population of AN fibers. Information profiles $(\delta'_f[CF])^2$ are plotted in Figure 5 for all 60 AN model CFs for low-, medium-, and high-frequency tones. The information profiles $(\delta'_f[CF])^2$ are plotted on a logarithmic scale to allow comparison across the three tone frequencies shown. The information profiles for the all-information and rate-place schemes are different in shape for the HSR fibers included in the computational AN model. In the all-information scheme, a contiguous range of CFs centered at the frequency of the tone provides information. Above 10 dB SPL, fibers near the frequency of the tone have similar high levels of synchrony, and thus contribute nearly equally to frequency discrimination. In contrast, in the rate-place encoding scheme, model AN fibers firing at the maximum average rate do not provide any information for estimating frequency. For HSR fibers, only AN fibers at the edge of the rate-response area provide useful information for estimating frequency based on rate-place (Siebert, 1968). The small populations of AN fibers with higher thresholds and broader dynamic ranges could provide useful rate-place information for CFs close to the tone frequency.

5.2 Level Discrimination in Quiet. Predictions of optimal performance for pure-tone level discrimination are compared with human performance in terms of ΔL as a function of level, duration, and frequency in Figure 6.[9] Rate-place predictions from an analytical AN model with HSR fibers are

[9] Level discrimination will be discussed in terms of the level difference between the two tones, defined as $\Delta L = 20\log_{10}[(p + \Delta p)/p]$ dB, where p is the sound pressure of the standard tone. While there has been much debate about the most appropriate metric for level discrimination (see Grantham & Yost, 1982; Green, 1988), ΔL has been reported to be proportional to sensitivity d' by several groups (e.g., Rabinowitz et al., 1976; Buus & Florentine, 1991), and thus this metric will be used in this study.

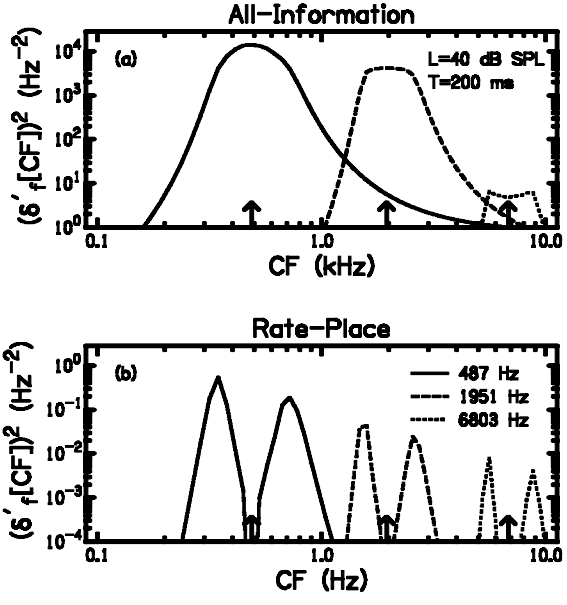


Figure 5: Information responsible for optimal frequency discrimination for low- (487 Hz), medium- (1951 Hz), and high-frequency (6803 Hz) conditions (indicated by arrows) with $L = 40$ dB SPL and $T = 200$ ms (20 ms rise/fall). The top and bottom panels show the information $(\delta'_f[CF])^2$ for each of the 60 AN model CFs (roughly 200 independent AN fibers per model CF) in the all-information and rate-place encoding schemes, respectively. The linear decline in the peaks of the rate-place information profiles as frequency increases contrasts the sharp drop in the all-information peaks at the highest frequency. These trends correspond directly to the flat rate-place $\Delta f/f$ and to the increasing all-information $\Delta f/f$ seen in Figure 4a as frequency increases.

also shown for comparison (Siebert, 1968). Optimal performance is roughly an order of magnitude better than human performance, and all-information JNDs are roughly a factor of two lower than rate-place predictions (due to the level information from AN phase locking; Colburn, 1981).

Both the human and computational model predictions demonstrate a decrease in ΔL as level is increased in Figure 6a, that is, the near-miss to Weber's law. The slope of ΔL versus level in the computational model predictions is similar to that in the human data for levels above 30 dB SPL. In contrast, Siebert's (1968) model demonstrates Weber's law above 40 dB SPL. This difference is a consequence of the nontriangular filters used in the computational model (see section 6.3). In addition, the inclusion of nonlinear tuning and distributions of AN-fiber threshold and dynamic range contributes additional information about changes in level that is not in-

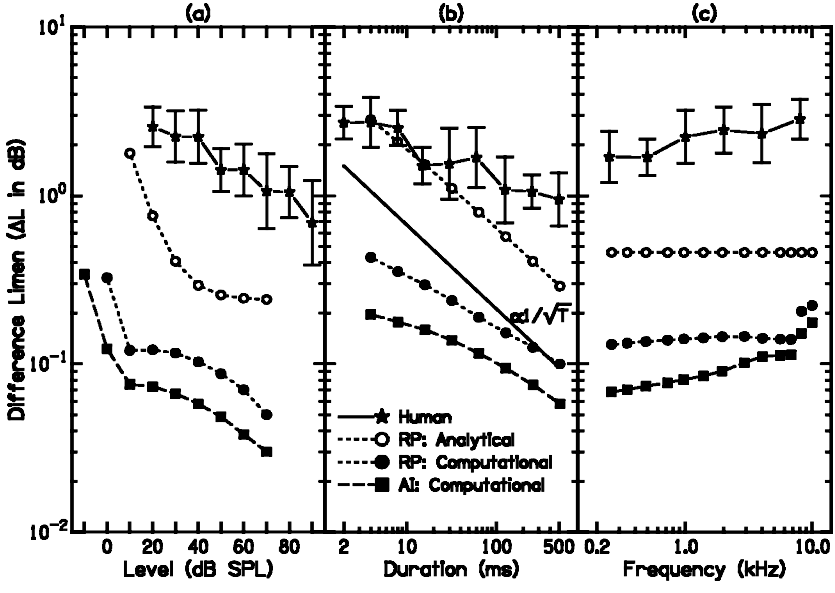


Figure 6: Optimal performance in terms of ΔL for pure-tone level discrimination. Predictions from the computational AN model are shown by filled circles for the rate-place scheme and filled squares for the all-information scheme. Rate-place predictions from an analytical AN model (Siebert, 1968) are shown as open circles. Typical human performance is illustrated by stars. (a) Effect of level. All model predictions are for $f = 970$ Hz and $T = 500$ ms (20 ms rise/fall). Human data are from Florentine et al. (1987) for $f = 1000$ Hz and $T = 500$ ms (20 ms rise/fall). (b) Effect of duration. Predictions are for $L = 40$ dB SPL, $f = 970$ Hz, and 4 ms rise/fall ramps. Human data are from Florentine (1986) for $L = 40$ dB SPL and $f = 1000$ Hz, and 1 ms rise/fall ramps. The solid line illustrates the slope associated with $\Delta L \propto T^{-1/2}$. (c) Effect of frequency. Predictions are for $L = 40$ dB SPL and $T = 200$ ms (20-ms rise/fall). Human data are from Florentine et al. (1987) for $L = 40$ dB SPL and $T = 500$ ms (20 ms rise/fall). The two high-frequency points not connected by lines for each encoding scheme are conditions for which the upper limit of the computational model CF range resulted in overestimates of the JNDs.

cluded in the model's HSR population (Heinz, 2000), as discussed in section 6.3.

As a function of duration, the computational model predictions for both all-information and rate-place illustrate the same rate of improvement in performance as is observed in the human data (see Figure 6b). The rate of improvement is shallower in slope than the predictions from the analytical AN model (open circles) and than the solid line representing $\Delta L \propto \sqrt{T^{-1}}$. The decay in discharge rate due to neural adaptation is likely to be respon-

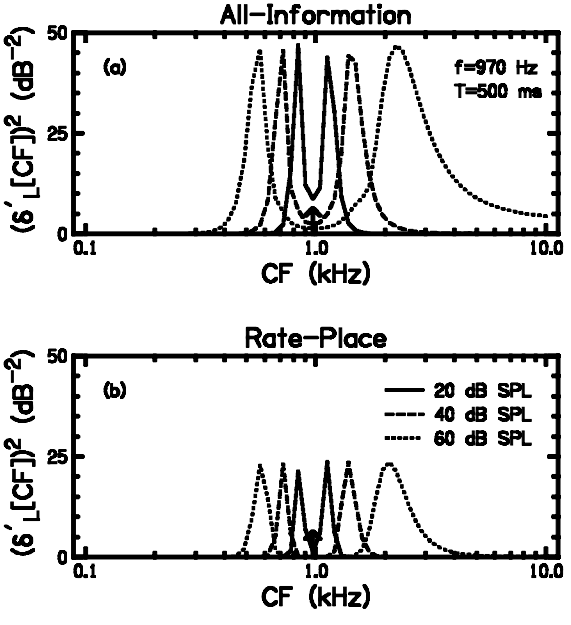


Figure 7: Information responsible for optimal level discrimination for three level conditions (20, 40, and 60 dB SPL) with $f = 970$ Hz (indicated by arrow)and $T = 500$ ms (20 ms rise/fall). The top and bottom panels show the information $(\delta'_L[CF])^2$ for each of the 60 AN model CFs (roughly 200 independent AN fibers per model CF) in the all-information and rate-place encoding schemes, respectively.

sible for the shallow slope in optimal performance demonstrated by the computational AN model.

As a function of frequency (see Figure 6c), the rate-place predictions are essentially flat except for the highest-frequency conditions for which the upper limit of the model CF range limited performance. The all-information predictions converge toward the rate-place predictions as frequency increases due to the loss of temporal phase-locking information at high frequencies. The shallow slope versus frequency in the all-information predictions is similar to that in the human JNDs for this stimulus level. A much sharper increase with frequency is observed in human data at high levels (Florentine et al., 1987), but this effect is not likely to be present in model predictions at higher levels.

Level-discrimination information profiles (see Figure 7) for rate-place and all-information have the same general shape, unlike the information profiles for frequency discrimination (see Figure 5). For level discrimination based on HSR fibers with a limited dynamic range, the edges of the

response area provide the only information in the model for both the all-information and rate-place encoding schemes due to the saturation of both rate and synchrony as level increases. Thus, spread of excitation across the population of AN fibers is the primary source of robust level encoding across a wide range of levels based on HSR fibers with a limited dynamic range (Siebert, 1968). Additional information (both rate and temporal) about changes in level is present in AN fibers with CFs near the tone frequency based on nonlinear cochlear tuning (Heinz, 2000), which is not included in the present AN model and is discussed further in section 6.3. The primary source of the near-miss in the model predictions is an increase in the number of fibers that contribute significantly as level increases, especially above the frequency of the tone.

6 Discussion

6.1 Analytical Versus Computational Approach. The limits on performance for level and frequency discrimination imposed by the random nature of AN responses were first examined roughly 30 years ago by Siebert (1965, 1968, 1970). The primary goal of this study is to extend the signal detection theory approach Siebert used to allow the use of more general (computational, rather than analytical) AN models. In order to verify the computational method, a simplified AN model was used in this study to permit direct comparisons to Siebert's predictions based on his analytical AN model.

The predictions from this study matched Siebert's (1965, 1968, 1970) predictions well in almost all aspects over the parameter range for which the analytical AN model was applicable:

1. There is significantly more temporal information than rate-place information in the AN for encoding changes in frequency.
2. Rate-place performance in terms of $\Delta f/f$ is roughly flat as a function of frequency.
3. All-information performance in terms of $\Delta f/f$ improves as frequency increases for low frequencies.
4. All-information performance improves as $\Delta f \propto \sqrt{T^{-3}}$, while rate-place improves as $\Delta f \propto \sqrt{T^{-1}}$.
5. All-information Δf continues to decrease as level increases, while rate-place Δf is roughly asymptotic above 20 dB SPL.
6. Spread of excitation can produce constant level-discrimination performance over a wide range of levels based on only low-threshold fibers with a limited dynamic range.
7. Level discrimination is essentially independent of frequency.

The predictions from the computational AN model differed from Siebert's (1965, 1968, 1970) predictions in two cases: (1) level discrimination in the computational model improved at a slower rate versus duration and similarly to human data, due to the effect of neural adaptation that was not included in Siebert's (1965, 1968) model, and (2) level-discrimination performance in our model demonstrated the near-miss to Weber's law that is observed in human performance, due to the nontriangular filters used in the computational model, whereas Siebert's (1965, 1968) model predicted Weber's law. Both of these discrepancies represent cases in which the computational model provides a better match to human performance. The general agreement between the predictions from the computational and analytical models establishes the validity of the computational approach.

The predictions from the computational AN model extended Siebert's (1965, 1968, 1970) predictions in two cases: (1) all-information predictions were able to be made at high frequencies due to the inclusion of the rolloff in phase locking, and (2) predictions were able to be made at short durations due to the inclusion of realistic onset and offset responses and neural adaptation. Both of these extensions raise important issues for auditory encoding of frequency and level, as discussed below.

6.2 Encoding of Frequency.

6.2.1 Rate-Place. The absolute JNDs for optimal use of rate-place information are close to human performance—much closer to human JNDs than are all-information predictions (see Figure 4). Further, rate-place predictions match the trends in human performance versus frequency below 2 kHz and versus duration above 20 ms. However, there are several inconsistencies between predicted rate-place and human performance that must be explained if a rate-place model is to account for human performance.

First, predicted frequency-discrimination performance based on rate-place information is inconsistent with the distinct worsening in human performance above 2 kHz (see Figure 4a). Our prediction of invariant performance with frequency is consistent with all other rate-place models for frequency discrimination that incorporate realistic filter slopes as a function of frequency (e.g., Siebert, 1968, 1970; Javel & Mott, 1988). Thus, in order for a rate-place model to account for human frequency discrimination, some alteration must be included to account for the order-of-magnitude increase in $\Delta f/f$ between 2 and 8 kHz. Several possible explanations are considered:

1. A decrease in the efficiency of processing rate-place information above 2 kHz could explain human performance; however, it is not clear why a mechanism for counting AN discharges would be optimal at low frequencies but an order of magnitude less efficient at high frequencies.
2. A reduction in the innervation density of independent AN fibers at high frequencies by a factor of 100 could explain the order of magni-

tude worsening in human performance. However, this is inconsistent with the increase in AN-fiber innervation density from apex to base of the cat cochlea (Keithley & Schreiber, 1987; Liberman, Dodds, & Pierce, 1990).

3. The maximum potential increase in $\Delta f/f$ that would result from the disappearance of the entire upper side of the information profile (see Figure 5) off the basal end of the cochlea as tone frequency increases is $\sqrt{2}$, which is too small to account for human performance.

4. A decrease in the sharpness of tuning at high frequencies is inconsistent with reports of human (Glasberg & Moore, 1990, used in this study) and cat tuning (Liberman, 1978; Miller, Schilling, Franck, & Young, 1997), which become sharper as frequency increases.

Thus, there appears to be no physiologically realistic explanation for the worsening of rate-place performance as frequency increases above 2 kHz, as observed in the human data.

The second possible explanation is that asymptotic rate-place performance as level increases (actually slight increase in computational-model JNDs) is inconsistent with the shallow improvement in human performance as level increases (see Figure 4c). The inclusion of high-threshold, low-spontaneous-rate AN fibers could be expected to lead to improved performance at high levels. However, Erell (1988) demonstrated that inclusion of medium- and low-spontaneous-rate fibers (with higher thresholds and larger dynamic ranges; Liberman, 1978) did not significantly extend the limited range of levels over which rate-place performance improves.

6.2.2 All-Information. In contrast to rate-place, all-information performance matches the trends in human performance across all frequencies, especially at high frequencies, and across level (see Figures 4a and 4c). The current predictions quantify that there is significant temporal information in the AN up to at least 10 kHz, despite very low synchrony coefficients for AN fibers at high frequencies (see Figure 1c). However, there are several inconsistencies between predicted all-information performance and human performance.

First, there is a two-orders-of-magnitude discrepancy between human performance and optimal all-information performance at both low and high frequencies. Second, all-information performance for frequency discrimination improves much too rapidly as duration increases ($\Delta f \propto \sqrt{T^{-3}}$ rather than $\Delta f \propto \sqrt{T^{-1}}$; see Figure 4b). This rapid improvement is due to allowing all possible comparisons between discharges to be used by the optimal processor (see Figure 3).

6.2.3 Potential Signal-Processing Mechanisms. It is of interest to hypothesize signal processing mechanisms that would incorporate the useful as-

pects of both rate-place and all-information encoding schemes for explaining human performance. A processor that makes use of the temporal discharge patterns appears to be required to produce the U-shaped dependence of $\Delta f/f$ on frequency (see Figure 4a); however, the temporal information must be used inefficiently to account for the duration effect (see Figure 4b). In a pilot modeling study, the computational AN model population was used to make predictions based on Goldstein and Srulovicz's (1977) analysis of first-order intervals.[10] While the first-order interval behavior demonstrated a $\Delta f \propto \sqrt{T^{-1}}$ relation at long durations ($T \geq 50$ ms), a $\Delta f \propto \sqrt{T^{-3}}$ relation was observed at short durations. The steep improvement at short durations is similar to Goldstein and Srulovicz's predictions, but is steeper than the improvement in human performance at short durations (see Figure 4b; Moore, 1973). The U-shaped dependence of performance versus frequency was maintained by the computational first-order-interval scheme, similar to Goldstein and Srulovicz, with a parallel upward shift of the entire curve by roughly one order of magnitude (i.e., roughly halfway closer to human performance).

A temporal processor that considered all discharge times (i.e., all-order intervals) within a limited temporal window would be expected to demonstrate the desired $\Delta f \propto \sqrt{T^{-1}}$ dependence for stimulus durations much longer than the length of the temporal window. For durations shorter than the temporal window, such a scheme would be expected to behave similarly to the all-information processor, which matches the rate of improvement in human performance with duration for $T \leq 16$ ms in this study (see Figure 4b). A restricted temporal processor of this type would demonstrate the same shallow improvement with level as the all-information scheme, consistent with human performance, and would be considerably closer to overall human performance levels due to the large amount of temporal information discarded (see Figures 3 and 4b). It is reasonable to assume that neurons in the central nervous system perform signal processing over restricted temporal windows, given that cell membranes have finite time constants; however, specific restricted temporal processors must be proposed and tested quantitatively to determine whether the above expectations hold.

6.3 Encoding of Level. Our predictions for pure-tone level discrimination did not demonstrate a large difference between performance based on

[10] There are several simplifying assumptions in Goldstein and Srulovicz's (1977) analysis of frequency-discrimination performance based on first-order intervals. Several of the simplifications that were made by assuming that RT was large (where R is average rate), were removed in the pilot study with the computational AN model and had little effect. However, the assumption that all intervals within the duration of the stimulus are independent, which is not true for a limited-duration stimulus or for a nonstationary Poisson process, was not able to be removed. The significance of this assumption for predicting optimal performance based on first-order intervals is unclear.

the all-information and rate-place encoding schemes. Performance based on optimal use of the information available in the AN was roughly one order of magnitude better than human performance. Optimal performance demonstrated the same general trends seen in human performance for the conditions studied (see Figure 6). This result is of interest because the AN model used in this study did not include many physiological properties that have been hypothesized to be important for level discrimination, as discussed below.

The computational AN model used in this study did not include several important aspects of the nonlinear tuning on the basilar membrane (BM): changes in gain, bandwidth, and phase with level (Rhode, 1971; Anderson, Rose, Hind, & Brugge, 1971; Geisler & Rhode, 1982; Ruggero, Rich, Recio, Narayan, & Robles, 1997; Recio, Rich, Narayan, & Ruggero, 1998), and suppression (Sachs & Kiang, 1968; Delgutte, 1990; Ruggero, Robles, & Rich, 1992). The perceptual significance of the compressive gain observed on the BM has been primarily examined psychophysically (see Moore, 1995, and Moore & Oxenham, 1998, for reviews). The compressive gain can extend the dynamic range over which changes in level can be detected in individual AN fibers (Heinz, 2000). The level-dependent phase response of AN fibers has a much larger dynamic range than rate information (Anderson et al., 1971) and thus can provide additional temporal information for level discrimination at high levels for which the majority of AN fibers are saturated in terms of average rate (Carney, 1994; Heinz, 2000). The absence of suppression in our model prohibits the accurate simulation of AN responses to complex stimuli, as discussed in section 6.5.

The near-miss to Weber's law in the computational model predictions matches human performance (see Figure 6a), but contrasts with Siebert's (1965, 1968) prediction of Weber's law. Both the computational AN model and Siebert's (1965, 1968) model have only high-spontaneous-rate (HSR), low-threshold fibers and rely strongly on spread of excitation across the AN population to encode changes in level for medium and high levels. In the computational model, an increase in the number of CFs contributing useful information as level increases, especially above the frequency of the tone, produces the near-miss (see Figure 7). Potentially significant differences between the computational AN model and Siebert's (1965, 1968) analytical model include (1) the use of gamma-tone filters rather than triangular filters, (2) the form of the variation in filter bandwidth with CF, and (3) the distribution of CF across the population of AN fibers. Our results, and those of Teich and Lachs (1979), demonstrate that the near-miss does not require Weber's law to hold in narrow frequency regions (an assumption that is often made in psychophysical models; e.g., Florentine & Buus, 1981). However, the simplified AN model used in this study has several significant limitations that preclude any definitive conclusions regarding level encoding.

Our model would not likely predict Weber's law for broadband noise (Miller, 1947) or for narrowband noise in band-reject noise (Viemeister,

1983), or the nonmonotonic dependence of ΔL on level at high frequencies (Carlyon & Moore, 1984; Florentine et al., 1987). A significant limitation of the AN model used in this study is the exclusion of LSR, high-threshold AN fibers (16% of the population; Liberman, 1978), which have been implicated in accounting for Weber's law in limited frequency regions (Colburn, 1981; Delgutte, 1987; Viemeister, 1988a, 1988b; Winslow & Sachs, 1988; Winter & Palmer, 1991). However, more efficient processing of the LSR fibers than the HSR fibers is required to achieve Weber's law in narrow frequency regions, and thus the near-miss when information is combined across the population of CFs (Delgutte, 1987). A more complex AN model that includes the nonlinear changes in gain and phase as well as the LSR population is needed to quantify the relative contributions of these physiological properties to the encoding of sound level at high levels (Heinz, 2000).

An interesting finding from this study, despite the limitations mentioned, is that the shallow improvement in ΔL with duration observed in the computational model predictions matches human performance (see Figure 6b). This shallow improvement is likely due to neural adaptation, as suggested by Buus and Florentine (1992), and indicates that a limitation in the temporal integration capabilities of a central processor is not needed to explain the shallow improvement in level-discrimination performance as duration increases.

6.4 Relation Between Frequency and Level Encoding. The bandpass tuning in the auditory system implies that a change in rate on an AN fiber can result from a change in either stimulus level or frequency. The implications of this property for random-level frequency discrimination are evaluated quantitatively in the companion study. Siebert (1968) used a simple AN model with triangular-shaped filters to demonstrate a fundamental relationship between level and frequency discrimination based on rate-place information. A shortcoming of rate-place models has been their inability to account for both frequency and level discrimination in terms of the ratio of Weber fractions W_A/W_F (Siebert, 1968; Erell, 1988). Siebert (1968) predicted that the ratio of the Weber fraction for amplitude, $W_A = \Delta A/A$, to the Weber fraction for frequency, $W_F = \Delta f/f$, should be roughly 15 for rate-place models based on the shapes of AN tuning curves. Values of 10 to 15 for this ratio are typical of most rate-place models; however, human psychophysical data typically yield a ratio closer to 50 (Erell, 1988). Erell (1988) suggested that this discrepancy could be resolved by using filter bandwidths that were four to five times narrower than those based on tuning in cat and that human tuning is sharper than in cat. Typical values of the ratio W_A/W_F from this study (based on human tuning estimated psychophysically) were 11 for rate-place and 710 for all-information (for $f = 970$ Hz, $L = 40$ dB SPL, $T = 64$ ms, and 4 ms rise/fall). Human behavior lies between the predicted behavior based on rate-place and all-information encoding schemes. This result supports the notion that a restricted tempo-

ral encoding scheme (one intermediate to rate-place and all-information) should be pursued.

6.5 Limitations of This Study. *6.5.1 Computational Auditory Nerve Model.* The computational AN model used in this study has a basic filter shape that is the same for low and high CFs, with the only difference being filter bandwidth. This assumption ignores the tails of AN tuning curves observed for high-CF AN fibers (Kiang & Moxon, 1974). Inclusion of tails would affect the details of the predicted performance as a function of level; however, the significance of this assumption is minimized for the predictions versus frequency and duration, which were limited to low- and mid-level stimuli. The omission of tuning curve tails and of nonlinear interaction between different CFs, for example, suppression (Sachs & Kiang, 1968; Delgutte, 1990), also limits the applicability of this AN model for complex (broadband) sounds.

The Poisson assumption used in our model ignores refractory effects and could affect some of the details of the performance limits predicted in this study. Siebert (1965, 1968, 1970) and Colburn (1969, 1973) have argued that this assumption is not significant because AN fibers never respond at a sustained rate for which the typical interspike interval is less than four to five times larger than the absolute refractory period; however, this argument does not rule out a significant influence of the relative refractory period for sustained responses and does not address short-duration stimuli. The standard deviation of AN discharge counts to long-duration stimuli (i.e., $\geq$ 50 ms) has been reported to be less than that predicted by the Poisson model (Teich & Khanna, 1985; Young & Barta, 1986; Delgutte, 1987, 1996; Winter & Palmer, 1991). The maximum deviation from the Poisson model occurs at high discharge rates, and the Poisson model becomes more accurate as discharge rate decreases. Young and Barta (1986) reported that the standard deviation of the highest discharge counts for 200-ms stimuli ranged from 55% to 80% of the Poisson value, quantitatively consistent with other reports that used 50-ms stimuli (e.g., Delgutte, 1987, and Winter & Palmer, 1991). Thus, the maximum discrepancy from using the Poisson model for long-duration stimuli ($\geq$ 50 ms) could be an overestimate of some rate-place JNDs by a factor of two; however, the discrepancy will typically be much less since fibers with high discharge rates do not contribute significant rate-place information (see Figures 5 and 7; Colburn, 1981; Winter & Palmer, 1991). Miller, Barta, and Sachs (1987) provided evidence that single-unit phase-locked responses are consistent with a nonstationary Poisson model. They showed that the mean and variance of period histograms (from sustained responses) had roughly the same pattern and absolute value. Thus, for stimulus durations greater than 50 ms, the errors from using a Poisson model without refractory effects are small compared with the important trends in our predictions, which often range over more than an order of magnitude.

For short-duration stimuli, refractory effects could be expected to have a larger influence than for long-duration stimuli; however, there are insufficient data to estimate quantitatively the effect of refractoriness on rate-place and all-information predictions for short-duration stimuli. The overestimation of rate-place JNDs may be more significant for short-duration stimuli because onset responses have higher discharge rates (and larger dynamic ranges; see Figure 1b) than sustained responses and therefore may deviate more from the Poisson model. The effect of refractoriness on the all-information model can be thought of as removing those discharges that fall within the refractory period of a previous discharge. Thus, all-information predictions underestimate JNDs because refractoriness results in less information (i.e., fewer observations) than the Poisson model. However, all-information predictions for frequency discrimination may not be significantly influenced by refractoriness, even for short-duration stimuli, because low-order intervals are removed, and all-information performance is dominated by information from large-order intervals (see Figure 3). Thus, refractoriness would generally act to worsen all-information performance and improve rate-place performance, and this prediction can be used to bound the potential effects of refractoriness in this study—for example, a small effect is predicted for level-discrimination (see Figure 6).

6.5.2 Signal Detection Theory. The Cramér-Rao bound is restricted in that it only provides a limit on the performance of all possible decision processes for a Poisson process, and thus should be viewed primarily as a method for quantifying an upper bound on the total information available in the auditory nerve. When an efficient estimator is shown to exist, as in the case here, the Cramér-Rao bound quantifies the total information exactly. The Cramér-Rao-bound method by itself does not describe the type of processing required to achieve the performance limit, which is often significantly better than human performance (e.g., see Figure 4). Derived optimal processors (e.g., using a likelihood ratio test) typically include many physiologically unrealistic operations on the observed population of AN discharges, such as the ability to compare all discharge times across the entire duration of the stimulus, as well as the ability to compare the discharges across all AN fibers (Colburn, 1969, 1973; Siebert, 1970). It is often the case that physiologically realistic, but nonoptimal, processors can be described that perform at levels more consistent with human performance (e.g., Colburn, 1969, 1973, 1977a, 1977b; Goldstein & Srulovicz, 1977; see Delgutte, 1996). This approach provides intuition regarding the actual processing performed in the auditory system and is thus a natural second step after the fundamental performance limits have been described based on the total AN information.

There is particular interest in the potential to evaluate quantitatively the effect of noise maskers in experiments that have been used to test rate-place and temporal encoding schemes (e.g., Viemeister, 1983; Moore & Glasberg,

1989). However, the SDT analysis used in this study is limited to single-parameter discrimination tasks with deterministic stimuli (e.g., frequency or level discrimination). Analysis of performance limits in tasks with random stimuli requires several extensions from this study. The influence of stimulus (external) variability on the information in individual AN fibers must be calculated in addition to the influence of the physiological (internal) variability that was evaluated in the study. Random stimuli also result in correlations between activity of AN fibers with overlapping frequency tuning due to the common stimulus drive (Young & Sachs, 1989). Thus, the independence assumption used in this study is invalid for random stimuli, and the correlation between AN fibers must be considered in evaluating performance limits based on the entire AN population. Suppression (i.e., nonlinear interaction between different CFs; Sachs & Kiang, 1968; Delgutte, 1990), which is not included in our AN model, influences AN responses to complex stimuli. Suppression could potentially produce significant interactions between a target and an off-frequency noise masker that are not intuitively clear or consistent with the assumption that all fiber information is eliminated within the noise masker (e.g., Viemeister, 1983; Moore & Glasberg, 1989). Therefore, suppression must be included in AN models to evaluate performance limits for tasks with random-noise stimuli. There are now computational AN models that include the effects of suppression and produce responses to arbitrary stimuli (e.g., Robert & Eriksson, 1999; Zhang et al., 2001).

An initial extension of the SDT analysis to random stimuli is presented in the companion paper in this issue, which evaluates performance limits for psychophysical discrimination tasks in which a single nuisance parameter is randomly varied (e.g., random-level frequency discrimination). A more general extension of the SDT analysis to random stimuli has been developed and applied to the detection of tones in noise maskers (Heinz, 2000). In addition to the effects of suppression and across-fiber correlation, the information contained in the temporal patterns (either fine-time or envelope) of AN fiber responses was shown to be significant in this study.

7 Conclusion

Signal detection theory can be combined with computational models to predict psychophysical performance limits based on an entire neural population. Auditory discrimination performance limits based on the computational AN model matched predictions from previous studies over the parameter range for which analytical AN models were applicable. The parameter range over which psychophysical performance can be predicted was extended using the computational AN model.

The following three conclusions related to the encoding of frequency are not likely to depend on the details of the AN model used in this study:

- Optimal use of rate-place information is consistent with human frequency discrimination performance at low frequencies but inconsis-

tent with the trends in performance versus frequency at high frequencies.

- Frequency discrimination based on optimal use of all-information is roughly two orders of magnitude better than human performance at all frequencies but demonstrates the correct trends versus frequency, based on the frequency dependence of phase locking in the cat.
- There is significant temporal information in the AN for frequency discrimination up to at least 10 kHz, and thus temporal schemes cannot be rejected at high frequencies based on the decrease in phase locking in the AN.

Future studies using the computational approach with more complex AN models are now justified and will be particularly interesting for the task of level discrimination. In addition, this approach is applicable to any neural population in a sensory system for which a model of the discharge pattern statistics is available.

Acknowledgments

We thank Bertrand Delgutte, Christopher Long, Christine Mason, Martin McKinney, Susan Moscynski, Bill Peake, Tom Person, Timothy Wilson, Xuedong Zhang, and Ling Zheng for providing valuable comments on an earlier version of this article. We greatly appreciate the constructive comments and suggestions provided by three anonymous reviewers. We thank Mary Florentine for providing the data from Florentine (1986) shown in Figure 4b and Don Johnson for his synchrony data from cat shown in Figure 1c. This study was part of a graduate dissertation in the Speech and Hearing Sciences Program of the Harvard-MIT Division of Health Sciences and Technology (Heinz, 2000). Portions of this work were presented at the 21st and 22nd Midwinter Meeting of the Association for Research in Otolaryngology. The simulations in this study were performed on computers provided by the Scientific Computing and Visualization group at Boston University. This work was supported in part by the National Institutes of Health, Grants T32DC00038 and R01DC00100, as well as by the Office of Naval Research, Grant Agmt Z883402.

References

Abbott, L. F., & Sejnowski, T. J. (1999). Introduction. In L. F. Abbott & T. J. Sejnowski (Eds.), *Neural codes and distributed representations: Foundations of neural computation*. Cambridge, MA: MIT Press.

Anderson, D. J., Rose, J. E., Hind, J. E., & Brugge, J. F. (1971). Temporal position of discharges in single auditory nerve fibers within the cycle of a sinewave stimulus: Frequency and intensity effects. *J. Acoust. Soc. Am., 49*, 1131–1139.

Braida, L. D., & Durlach, N. I. (1988). Peripheral and central factors in intensity perception. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Auditory function: Neurobiological bases of hearing* (pp. 559–583). New York: Wiley.

Buus, S., & Florentine, M. (1991). Psychometric functions for level discrimination. *J. Acoust. Soc. Am., 90*, 1371–1380.

Buus, S., & Florentine, M. (1992). Possible relation of auditory-nerve adaptation to slow improvement in level discrimination with increasing duration. In Y. Cazals, L. Demany, & K. Horner (Eds.), *Auditory physiology and perception* (pp. 279–288). New York: Pergamon.

Cariani, P. A., & Delgutte, B. (1996a). Neural correlates of the pitch of complex tones: I. Pitch and pitch salience. *J. Neurophysiol., 76*, 1698–1716.

Cariani, P. A., & Delgutte, B. (1996b). Neural correlates of the pitch of complex tones: II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *J. Neurophysiol., 76*, 1717–1734.

Carlyon, R. P., & Moore, B. C. J. (1984). Intensity discrimination: A severe departure from Weber's law. *J. Acoust. Soc. Am., 76*, 1369–1376.

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

Carney, L. H. (1994). Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness. *Hear. Res., 76*, 31–44.

Colburn, H. S. (1969). *Some physiological limitations on binaural performance*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Colburn, H. S. (1973). Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. *J. Acoust. Soc. Am., 54*, 1458–1470.

Colburn, H. S. (1977a). Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. *J. Acoust. Soc. Am., 61*, 525–533.

Colburn, H. S. (1977b). *Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. Supplementary material*. (AIP Document No. PAPS JASMA-61-525-98). Available online at: http://www.aip.org/epaps/howorder.html.

Colburn, H. S. (1981). *Intensity perception: Relation of intensity discrimination to auditory-nerve firing patterns* (Internal Memorandum). Cambridge, MA: Research Laboratory of Electronics, Massachusetts Institute of Technology.

Cramér, H. (1951). *Mathematical methods of statistics*. Princeton: Princeton University Press.

Dau, T., Kollmeier, B., & Kohlrausch, A. (1997a). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *J. Acoust. Soc. Am., 102*, 2892–2905.

Dau, T., Kollmeier, B., & Kohlrausch, A. (1997b). Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *J. Acoust. Soc. Am., 102*, 2906–2919.

Dau, T., Püschel, D., & Kohlrausch, A. (1996a). A quantitative model of the "effective" signal processing in the auditory system I. Model structure. *J. Acoust. Soc. Am., 99*, 3615–3622.

Dau, T., Püschel, D., & Kohlrausch, A. (1996b). A quantitative model of the 'effective' signal processing in the auditory system II. Simulations and measurements. *J. Acoust. Soc. Am., 99*, 3623–3631.

Delgutte, B. (1987). Peripheral auditory processing of speech information: implications from a physiological study of intensity discrimination. In M. E. H. Schouten (Ed.), *The psychophysics of speech perception* (pp. 333–353). Dordrecht: Nijhoff.

Delgutte, B. (1990). Two-tone rate suppression in auditory-nerve fibers: Dependence on suppressor frequency and level. *Hear. Res., 49*, 225–246.

Delgutte, B. (1996). Physiological models for basic auditory percepts. In H. L. Hawkins, T. A. McMullen, A. N. Popper, & R. R. Fay (Eds.), *Auditory computation* (pp. 157–220). New York: Springer-Verlag.

Duifhuis, H. (1973). Consequences of peripheral frequency selectivity for nonsimultaneous masking. *J. Acoust. Soc. Am., 54*, 1471–1488.

Durlach, N. I., & Braida, L. D. (1969). Intensity perception I: Preliminary theory of intensity resolution. *J. Acoust. Soc. Am., 46*, 372–383.

Dye, R. H., & Hafter, E. R. (1980). Just-noticeable differences of frequency for masked tones. *J. Acoust. Soc. Am., 67*, 1746–1753.

Erell, A. (1988). Rate coding model for discrimination of simple tones in the presence of noise. *J. Acoust. Soc. Am., 84*, 204–214.

Evans, E. F., Pratt, S. R., Spenner, H., & Cooper, N. P. (1992). Comparisons of physiological and behavioural properties: Auditory frequency selectivity. In Y. Cazals, L. Demany, & K. Horner (Eds.), *Auditory physiology and perception* (pp. 159–169). New York: Pergamon.

Fitzhugh, R. (1958). A statistical analyser for optic nerve messages. *J. Gen. Physiol., 41*, 675–692.

Florentine, M. (1986). Level discrimination of tones as a function of duration. *J. Acoust. Soc. Am., 79*, 792–798.

Florentine, M., & Buus, S. (1981). An excitation pattern model for intensity discrimination. *J. Acoust. Soc. Am., 70*, 1646–1654.

Florentine, M., Buus, S., & Mason, C. R. (1987). Level discrimination as a function of level for tones from 0.25 to 16 kHz. *J. Acoust. Soc. Am., 81*, 1528–1541.

Freyman, R. L., & Nelson, D. A. (1986). Frequency discrimination as a function of tonal duration and excitation-pattern slopes in normal and hearing-impaired listeners. *J. Acoust. Soc. Am., 79*, 1034–1044.

Geisler, C. D., & Rhode, W. S. (1982). The phases of basilar-membrane vibrations. *J. Acoust. Soc. Am., 71*, 1201–1203.

Giguère, C., & Woodland, P. C. (1994a). A computational model of the auditory periphery for speech and hearing research. I. Ascending path. *J. Acoust. Soc. Am., 95*, 331–342.

Giguère, C., & Woodland, P. C. (1994b). A computational model of the auditory periphery for speech and hearing research. II. Descending paths. *J. Acoust. Soc. Am., 95*, 343–349.

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

Goldstein, J. L., & Srulovicz, P. (1977). Auditory-nerve spike intervals as an adequate basis for aural spectrum analysis. In E. F. Evans & J. P. Wilson (Eds.),

Psychophysics and physiology of hearing (pp. 337–347). New York: Academic Press.

Grantham, D. W., & Yost, W. A. (1982). Measures of intensity discrimination. *J. Acoust. Soc. Am., 72*, 406–410.

Green, D. M. (1988). *Profile analysis: Auditory intensity discrimination*. New York: Oxford University Press.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

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

Gresham, L. C., & Collins, L. M. (1998). Analysis of the performance of a model-based optimal auditory signal processor. *J. Acoust. Soc. Am., 103*, 2520–2529.

Heinz, M. G. (2000). *Quantifying the effects of the cochlear amplifier on temporal and average-rate information in the auditory nerve*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Houtsma, A. J. M. (1995). Pitch perception. In B. C. J. Moore (Ed.), *Hearing*. New York: Academic Press.

Huettel, L. G., & Collins, L. M. (1999). Using computational auditory models to predict simultaneous masking data: Model comparison. *IEEE Trans. Biomed. Eng., 46*, 1432–1440.

Javel, E., & Mott, J. B. (1988). Physiological and psychophysical correlates of temporal processing in hearing. *Hear. Res., 34*, 275–294.

Jesteadt, W., Wier, C. C., & Green, D. M. (1977). Intensity discrimination as a function of frequency and sensation level. *J. Acoust. Soc. Am., 61*, 169–177.

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.

Johnson, D. H., & Kiang, N. Y. S. (1976). Analysis of discharges recorded simultaneously from pairs of auditory-nerve fibers. *Biophys. J., 16*, 719–734.

Joris, P. X., Carney, L. H., Smith, P. H., & 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.

Kaernbach, C., & Demany, L. (1998). Psychophysical evidence against the autocorrelation theory of auditory temporal processing. *J. Acoust. Soc. Am., 104*, 2298–2306.

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

Kiang, N. Y. S. (1990). Curious oddments of auditory-nerve studies. *Hear. Res., 49*, 1–16.

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

Kiang, N. Y. S., Watanabe, T., Thomas, E. C., & Clark, L. F. (1965). *Discharge patterns of single fibers in the cat's auditory nerve*. Cambridge, MA: MIT Press.

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. (1980). Morphological differences among radial afferent fibers in the cat cochlea: An electron microscopic study of serial sections. *Hear. Res., 3*, 45–63.

Liberman, M. C., Dodds, L. W., & 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., & Goldstein, J. L. (1995). Quantifying 2-factor phase relations in nonlinear responses from low characteristic-frequency auditory-nerve fibers. *Hear. Res., 90*, 126–138.

May, B. J., & Sachs, M. B. (1992). Dynamic range of neural rate responses in the ventral cochlear nucleus of awake cats. *J. Neurophysiol., 68*, 1589–1602.

McGill, W. J., & Goldberg, J. P. (1968). A study of the near-miss involving Weber's law and pure-tone intensity discrimination. *Perception Psychophysics, 4*, 105–109.

Miller, G. A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *J. Acoust. Soc. Am., 19*, 606–619.

Miller, M. I., Barta, P. E., & Sachs, M. B. (1987). Strategies for the representation of a tone in background noise in the temporal aspects of the discharge patterns of auditory-nerve fibers. *J. Acoust. Soc. Am., 81*, 665–679.

Miller, R. L., Schilling, J. R., Franck, K. R., & Young, E. D. (1997). Effects of acoustic trauma on the representation of the vowel /ɛ/ in cat auditory nerve fibers. *J. Acoust. Soc. Am., 101*, 3602–3616.

Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *J. Acoust. Soc. Am., 54*, 610–619.

Moore, B. C. J. (1989). *An introduction to the psychology of hearing*. New York: Academic Press.

Moore, B. C. J. (1995). *Perceptual consequences of cochlear damage*. New York: Oxford University Press.

Moore, B. C. J., & Glasberg, B. R. (1989). Mechanisms underlying the frequency discrimination of pulsed tones and the detection of frequency modulation. *J. Acoust. Soc. Am., 86*, 1722–1732.

Moore, B. C. J., & Oxenham, A. J. (1998). Psychoacoustic consequences of compression in the peripheral auditory system. *Psychol. Rev., 105*, 108–124.

Mountain, D. C., & Hubbard, A. E. (1996). Computational analysis of hair cell and auditory nerve processes. In H. L. Hawkins, T. A. McMullen, A. N. Popper, & R. R. Fay (Eds.), *Auditory computation* (pp. 121–156). New York: Springer-Verlag.

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annu. Rev. Neurosci., 21*, 227–277.

Parzen, E. (1962). *Stochastic processes*. San Francisco: Holden-Day.

Patterson, R. D., Allerhand, M. H., & Giguère, C. (1995). Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform. *J. Acoust. Soc. Am., 98*, 1890–1894.

Patterson, R. D., Nimmo-Smith, I., Holdsworth, J., & Rice, P. (1987, Dec. 14–15). *An efficient auditory filterbank based on the gammatone function*. Paper presented at a meeting of the IOC Speech Group on Auditory Modeling at RSRE.

Patuzzi, R., & 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.

Pickles, J. O. (1988). *An introduction to the physiology of hearing*. New York: Academic Press.

Rabinowitz, W. M., Lim, J. S., Braida, L. D., & Durlach, N. I. (1976). Intensity perception VI: Summary of recent data on deviations from Weber's law for 1000-Hz tone pulses. *J. Acoust. Soc. Am., 59*, 1506–1509.

Rasmussen, G. L. (1940). Studies of the VIIIth cranial nerve in man. *Laryngoscope, 50*, 67–83.

Recio, A., Rich, N. C., Narayan, S. S., & Ruggero, M. A. (1998). Basilar-membrane responses to clicks at the base of the chinchilla cochlea. *J. Acoust. Soc. Am., 103*, 1972–1989.

Rhode, W. S. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *J. Acoust. Soc. Am., 49*, 1218–1231.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

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

Ruggero, M. A. (1992). Physiology and coding of sound in the auditory nerve. In A. N. Popper & R. R. Fay (Eds.), *The mammalian auditory pathway: Neurophysiology* (pp. 34–93). New York: Springer-Verlag.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., & Robles, L. (1997). Basilar-membrane responses to tones at the base of the chinchilla cochlea. *J. Acoust. Soc. Am., 101*, 2151–2163.

Ruggero, M. A., Robles, L., & Rich, N. C. (1992). Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression. *J. Neurophysiol., 68*, 1087–1099.

Ryugo, D. K. (1992). The auditory nerve: Peripheral innervation, cell body morphology, and central projections. In D. B. Webster, A. N. Popper, & R. R. Fay (Eds.), *The mammalian auditory pathway: Neuroanatomy* (pp. 23–65). New York: Springer-Verlag.

Sachs, M. B., & Kiang, N. Y. S. (1968). Two-tone inhibition in auditory-nerve fibers. *J. Acoust. Soc. Am., 43*, 1120–1128.

Schouten, J. F. (1940). The residue and the mechanism of hearing. *Proc. Kon. Nedel. Akad. Wetenschap, 43*, 991–999.

Siebert, W. M. (1965). Some implication of the stochastic behavior of primary auditory neurons. *Kybernetik, 2*, 206–215.

Siebert, W. M. (1968). Stimulus transformation in the peripheral auditory system. In P. A. Kolers & M. Eden (Eds.), *Recognizing patterns* (pp. 104–133). Cambridge, MA: MIT Press.

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

Snyder, D. L., & Miller, M. I. (1991). *Random point processes in time and space*. New York: Springer-Verlag.

Spoendlin, H. (1969). Innervation patterns in the organ of Corti of the cat. *Acta. Otolaryngol. (Stockh), 67*, 239–254.

Teich, M. C., & 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.

Teich, M. C., & Lachs, G. (1979). A neural-counting model incorporating refractoriness and spread of excitation. I. Application to intensity discrimination. *J. Acoust. Soc. Am., 66*, 1738–1749.

van Trees, H. L. (1968). *Detection, estimation, and modulation theory: Part I*. New York: Wiley.

Viemeister, N. F. (1974). Intensity discrimination of noise in the presence of band-reject noise. *J. Acoust. Soc. Am., 56*, 1594–1600.

Viemeister, N. F. (1983). Auditory intensity discrimination at high frequencies in the presence of noise. *Science, 221*, 1206–1208.

Viemeister, N. F. (1988a). Psychophysical aspects of auditory intensity coding. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Auditory function: Neurobiological bases of hearing* (pp. 213–241). New York: Wiley.

Viemeister, N. F. (1988b). Intensity coding and the dynamic range problem. *Hear. Res., 34*, 267–274.

Viemeister, N. F., & Bacon, S. P. (1988). Intensity discrimination, increment detection, and magnitude estimation for 1-kHz tones. *J. Acoust. Soc. Am., 84*, 172–178.

von Helmholtz, H. L. F. (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (F. Vieweg und Sohn, Braunschweig, Germany). Translated as: *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, by A. J. Ellis from the 4th German edition, 1877, Leymans, London, 1885 (reprinted by Dover, New York, 1954).

Wakefield, G. H., & Nelson, D. A. (1985). Extension of a temporal model of frequency discrimination: Intensity effects in normal and hearing-impaired listeners. *J. Acoust. Soc. Am., 77*, 613–619.

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

Westerman, L. A., & 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.

Wever, E. G. (1949). *Theory of hearing*. New York: Wiley.

Wier, C. C., Jesteadt, W., & Green, D. M. (1977). Frequency discrimination as a function of frequency and sensation level. *J. Acoust. Soc. Am., 61*, 178–184.

Winslow, R. L., & Sachs, M. B. (1988). Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. *Hear. Res., 35*, 165-190.

Winter, I. M., & Palmer, A. R. (1991). Intensity coding in low-frequency auditory-nerve fibers of the guinea pig. *J. Acoust. Soc. Am., 90*, 1958–1967.

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

Young, E. D., & Sachs, M. B. (1989). Auditory nerve fibers do not discharge independently when responding to broadband noise. *Abstracts of the 12th Midwinter Meeting of the Association for Research in Otolaryngology*, 121.

Zhang, X., Heinz, M. G., Bruce, I. C., & Carney, L. H. (2001). A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J. Acoust. Soc. Am., 109*, 648–670.

Received January 11, 2000; accepted January 3, 2001.