Quantifying the implications of nonlinear cochlear tuning for auditory-filter estimates

Michael G. Heinz[a)]

Speech and Hearing Sciences Program, Division of Health Sciences and Technology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139 and Hearing Research Center, Biomedical Engineering Department, Boston University, 44 Cummington Street, Boston, Massachusetts 02215

H. Steven Colburn and Laurel H. Carney[b)]

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

The relation between auditory filters estimated from psychophysical methods and peripheral tuning was evaluated using a computational auditory-nerve (AN) model that included many of the response properties associated with nonlinear cochlear tuning. The phenomenological AN model included the effects of dynamic level-dependent tuning, compression, and suppression on the responses of high-, medium-, and low-spontaneous-rate AN fibers. Signal detection theory was used to evaluate psychophysical performance limits imposed by the random nature of AN discharges and by random-noise stimuli. The power-spectrum model of masking was used to estimate psychophysical auditory filters from predicted AN-model detection thresholds for a tone signal in fixed-level notched-noise maskers. Results demonstrate that the role of suppression in broadening peripheral tuning in response to the noise masker has implications for the interpretation of psychophysical auditory-filter estimates. Specifically, the estimated psychophysical auditory-filter equivalent-rectangular bandwidths (ERBs) that were derived from the nonlinear AN model with suppression always overestimated the ERBs of the low-level peripheral model filters. Further, this effect was larger for an 8-kHz signal than for a 2-kHz signal, suggesting a potential characteristic-frequency (CF) dependent bias in psychophysical estimates of auditory filters due to the increase in strength of cochlear nonlinearity with increases in CF. © *2002 Acoustical Society of America.*
[DOI: 10.1121/1.1436071]

I. INTRODUCTION

The ability of the auditory system to partially resolve frequency components in a complex stimulus has been widely used as the basis for many fundamental theories of auditory perception (e.g., von Helmholtz, 1863; Fletcher, 1940, 1953). Thus, much effort has gone into developing psychophysical techniques for the measurement of auditory frequency selectivity. Fletcher (1940) observed that the detection threshold for a tone in bandlimited noise increases as the bandwidth of the noise increases up to a *critical bandwidth*, beyond which detection threshold is roughly constant. This observation led to the concept of the psychophysical *auditory filter*, which forms the basis for many modern psychophysical methods for estimating auditory tuning (for a review see Moore, 1995a).

Most psychophysical methods for measuring auditory frequency selectivity are based on the power-spectrum model of masking, which assumes that performance for tone-in-noise detection is constant when the long-term signal-to-noise power ratio is constant at the output of a linear filter centered at (or close to) the frequency of the tone (Moore, 1995a). Fletcher (1940) noted that the auditory-filter bandwidth could be estimated as the ratio of the signal power at masked threshold to the power-spectral density of a broadband noise, now referred to as the critical ratio (Moore, 1995a), if it were assumed that threshold corresponded to a signal-to-noise ratio of 0 dB at the output of the auditory filter. Auditory-filter bandwidths have also been estimated from Fletcher's band-widening experiment as the noise bandwidth (critical band) beyond which detection thresholds are constant, based on the assumption of rectangular filters. Psychophysical tuning curves have been derived by measuring the level of a variable-frequency tone or narrow-band masker that is required to just mask a low-level tone signal at the frequency of interest (e.g., Moore, 1978; Vogten, 1978). Despite the relative simplicity of these methods, the use of notched-noise maskers to estimate auditory-filter shapes has proven to be a much more reliable psychophysical method (Moore, 1995a). In the notched-noise method, detection thresholds are measured for a tone in the presence of two bandlimited noise maskers that are above and below the tone frequency (Patterson, 1976; Glasberg and Moore, 1990, 2000; Rosen *et al.*, 1998). Detection threshold is measured as

---

a function of the separation between the two noise bands (i.e., the notch width), and the power-spectrum model is used to derive an auditory-filter shape that accounts for the rate of threshold decrease as notch width is increased. Asymmetric notches are often used to derive the upper and lower sides of the auditory filter separately, and thus allow estimates of auditory-filter asymmetry to be made. While the notched-noise method provides estimates of auditory filters that are successful in predicting detection performance in a variety of masking conditions (e.g., Derleth and Dau, 2000), a fundamental question is whether the frequency selectivity measured psychophysically is primarily determined by peripheral tuning.

Glasberg and Moore (1990) derived equations that describe the variation in psychophysical estimates of auditory-filter equivalent-rectangular bandwidths (ERBs) with characteristic frequency (CF) and with masker level. The variation in auditory-filter ERB with CF was consistent with the idea that a psychophysical ERB represents a constant distance along the basilar membrane (Greenwood, 1961). The variation in psychophysical estimates of auditory filters with stimulus level was consistent with the well-established result that cochlear tuning broadens with increases in stimulus level (e.g., Patuzzi and Robertson, 1988; Ruggero et al., 1997). In addition, psychophysical estimates of auditory filters were typically broader for hearing-impaired listeners than for normal-hearing listeners (Glasberg and Moore, 1986; Moore, 1995b; Moore et al., 1999b), consistent with the loss of sharp frequency tuning in basilar-membrane and auditory-nerve (AN) responses of damaged cochleae (Patuzzi et al., 1989; Ruggero and Rich, 1991; Liberman and Dodds, 1984). Moore et al. (1999b) have shown that psychophysical estimates of auditory-filter ERB from the notched-noise method are correlated with several other psychophysical measures assumed to be related to cochlear nonlinearity. Thus, it appears that psychophysical estimates of frequency selectivity from the notched-noise method are closely related to peripheral tuning in humans. Consistent with this conclusion, Evans et al. (1992) found a high degree of correspondence between psychophysical and physiological ERBs both measured in guinea pigs.

In spite of this general agreement about the importance of psychophysical auditory filters, there is continued debate about the proper method for psychophysically estimating auditory-filter shapes as a function of level (Lutfi and Patterson, 1984; Moore and Glasberg, 1987; Glasberg and Moore, 1990, 2000; Rosen and Baker, 1994; Moore, 1995a; see Rosen et al., 1998 for a review). This debate has focused on which aspect of the stimulus, signal or masker level, controls the auditory filter shape. Moore (1995a) has proposed masker level per ERB as the controlling variable, rather than overall masker level or spectrum level. Rosen and Baker (1994) and Rosen et al. (1998) argued that deriving auditory filters across a range of fixed noise spectrum levels (e.g., Glasberg and Moore, 1990) is only appropriate if the level dependence of the auditory filter is determined by the noise spectrum level. If auditory filters were to vary with signal level rather than masker level, then the auditory filter would vary across notch widths because the signal level decreases

as notch width is increased, and the psychophysically estimated auditory filter would represent an average auditory filter across a range of signal levels. Rosen et al. (1998) have shown that notched-noise detection data across a wide range of levels were fitted better by the power-spectrum model when the auditory filter was assumed to depend on the signal level rather than on the masker level. Glasberg and Moore (2000), who used uniformly exciting noise that was designed to provide equal excitation for each psychophysical ERB, also concluded that models in which the signal level was assumed to control auditory-filter shape were better able to predict the detection data. Despite the debate over whether the signal or masker level controls auditory-filter shape, methods based on both views result in nonlinear changes in auditory-filter shapes that qualitatively match nonlinear trends in physiologically measured tuning. While the debate has focused on assumptions about which aspects of the stimulus control nonlinear tuning, a direct comparison between psychophysical measures of frequency selectivity and physiological measures of nonlinear peripheral tuning has not been made in the same subject.

The present study uses a computational AN model to relate peripheral tuning to estimates of frequency selectivity from psychophysical methods. This phenomenological AN model provides a useful description of nonlinear peripheral tuning (Heinz et al., 2001c; see also Zhang et al., 2001) and is used to evaluate how different nonlinear AN properties affect psychophysical methods for estimating auditory filters.

Many of the observed nonlinear AN response properties appear to result from a single physiologically vulnerable mechanism that controls peripheral tuning (Sachs and Abbas, 1974; Sewell, 1984; Patuzzi et al., 1989; Ruggero and Rich, 1991; Ruggero et al., 1992; see the review by Ruggero, 1992). This mechanism is widely believed to be related to outer-hair-cell (OHC) electromotility; however, the underlying biophysical basis for the role of the OHCs in this mechanism is still unknown (Allen, 2001). The most prevalent view is that the OHCs provide the high sensitivity and sharp tuning that characterize normal hearing through an active process, often referred to as the *cochlear amplifier*, which enhances the vibration of the basilar membrane in response to low-level sounds (e.g., Yates, 1995; Moore, 1995b). The results from the present study do not depend on the biophysical basis for the underlying mechanism; rather, they depend only on the idea that a single physiologically vulnerable mechanism produces many of the nonlinear peripheral response properties that have been described. Basilar-membrane tuning has been shown to broaden with increases in level and to demonstrate associated compressive magnitude responses and nonlinear phase responses near CF (Ruggero et al., 1997). These three nonlinear properties are related in terms of a filter gain/bandwidth trade-off, i.e., the peripheral filter gain at CF decreases as the bandwidth increases with increases in stimulus level. The nonlinear phase shifts correspond to the peripheral-filter phase-versus-frequency response becoming shallower as level increases, and have also been observed in inner-hair-cell (IHC) responses (Cheatham and Dallos, 1998) and AN responses (Anderson et al., 1971). Recio et al. (1998) have shown that

J. Acoust. Soc. Am., Vol. 111, No. 2, February 2002

Heinz *et al.*: Effects of cochlear nonlinearity on auditory filters   997

the dynamics of cochlear compression are extremely fast, with a time constant on the order of 200 μs.

Two-tone suppression, another response property of the auditory periphery associated with nonlinear cochlear tuning, refers to suppression of basilar-membrane and AN responses to CF tones by an off-CF tone (Sachs and Kiang, 1968; Delgutte, 1990b; Ruggero *et al.*, 1992). Delgutte (1990a) demonstrated that AN suppression plays an important role in the masking of signals by off-frequency stimuli, a finding that has been supported by related psychophysical experiments (Oxenham and Plack, 1998). Suppression has been hypothesized to play a role in psychophysical estimates of auditory tuning based on the common finding that psychophysical estimates of frequency selectivity are sharper when using non-simultaneous masking than when using simultaneous masking (e.g., Houtgast, 1977; Moore, 1978; Moore and Glasberg, 1981, 1982, 1986; Moore and O'Loughlin, 1986; Moore *et al.*, 1987). The role of suppression is often discussed in terms of how the reduction in signal response due to the off-frequency masker may affect psychophysical estimates of auditory tuning. However, Moore and O'Loughlin (1986) argue that this simple-attenuation view is not as well justified as a distributed-attenuation view, which is supported by both psychophysical (e.g., Moore and Glasberg, 1982) and physiological (Pickles, 1984) studies. A direct measure of the effect of off-frequency noise on the underlying peripheral tuning was reported by Kiang and Moxon (1974), who showed that the tuning curves of high-CF AN fibers became broader and CF-tone thresholds were raised when a near-threshold, low-frequency masking noise band was presented simultaneously with the tone. The phenomenological AN model used in the present study (Heinz *et al.*, 2001c; see also Zhang *et al.*, 2001) represents suppression as the ability of off-frequency energy to reduce the excitatory-filter gain at CF and thus to broaden peripheral tuning.

Thus, consideration of the peripheral response properties associated with nonlinear cochlear tuning suggests that psychophysical methods would have difficulty in directly estimating AN-fiber frequency selectivity. This does not necessarily pose a problem for psychophysical studies that aim to characterize behavioral frequency selectivity without regard for the underlying mechanisms. However, this issue is important for interpreting changes in psychophysical estimates of frequency selectivity as a function of level or cochlear state, as well as for modeling studies that use human psychophysical data to specify peripheral tuning.

The present study relates many of the response properties associated with nonlinear cochlear tuning to estimates of psychophysical auditory filters by combining a phenomenological AN model with signal detection theory (SDT). Peripheral tuning (as specified in the AN model) can be directly compared to estimates of psychophysical auditory filters based on SDT analysis of AN-model population discharge patterns. Previous studies have used SDT with analytical AN-population models to quantify psychophysical performance limits based on the stochastic nature of AN discharge patterns (e.g., Siebert, 1968, 1970; Colburn, 1973, 1981; Delgutte, 1987; Heinz *et al.*, 2001b). A benefit of this approach is that the AN forms a bottleneck in the information

pathway from the periphery to the brain, and thus the effects of peripheral response properties can be characterized based on the total information that is available to the central nervous system for the psychophysical task. A second benefit of considering AN information is that the effects of several significant peripheral transformations (between the basilar membrane and AN) are included, e.g., that AN information is encoded by intrinsically random all-or-none action potentials that have a spontaneous rate, and that the driven rate of AN fibers typically saturates about 30–40 dB above threshold. Heinz *et al.* (2001b) used this approach to demonstrate that compressive magnitude responses and nonlinear phase responses associated with nonlinear cochlear tuning are significant for the encoding of sound level based on responses of a narrow range of CFs; however, their analytical AN model was limited to pure-tone stimuli, and several extensions of this approach were necessary to evaluate detection in random notched-noise maskers in the present study. Heinz *et al.* (2001a) demonstrated the use of computational AN models with this SDT approach, and a theoretical extension of the SDT analysis was developed to quantify the influence of random stimulus variation (in addition to AN variability) on psychophysical performance (Heinz, 2000).

In the present study, psychophysical auditory filters were estimated from the AN model using the notched-noise method by predicting psychophysical detection thresholds based on the information available in the population discharge patterns of AN-model fibers. Two specific questions were addressed: (1) How closely are estimates of psychophysical auditory filters related to peripheral tuning, given that psychophysical detection is based on a population of 30 000 AN fibers that have a wide range of CFs, different spontaneous rates (SRs), and different thresholds (Liberman, 1978)? (2) How do AN response properties associated with nonlinear cochlear tuning influence estimates of psychophysical auditory filters? Notched-noise maskers with a fixed spectrum level (as in Glasberg and Moore, 1990) were used in order to present the most straightforward demonstration of the expected effects of nonlinear AN properties on estimates of psychophysical auditory filters. Four versions of the computational AN model were used to demonstrate the separate contributions of compression and suppression to the estimates of psychophysical auditory filters.

II. METHODS

A. Computational auditory-nerve model

Zhang *et al.* (2001) developed a phenomenological model for cat AN responses. This model has a single nonlinear mechanism that accounts for many nonlinear AN response properties, including the on- and off-frequency control of peripheral tuning, i.e., compression and suppression. Heinz *et al.* (2001c) modified the original model in several ways to make it more appropriate for evaluating human psychophysical performance. These modifications included a description of human tuning, implementation of three SR groups (Liberman, 1978), and several implementations of peripheral-tuning control to allow the separate effects of compression and suppression to be evaluated. The AN model

Heinz *et al.*: Effects of cochlear nonlinearity on auditory filters

used in the present study was identical (including all parameter values) to the Heinz *et al.* (2001c) model. The model properties that are relevant to the current study are described below, while the details of the model implementation are described elsewhere (Heinz *et al.*, 2001c; Zhang *et al.*, 2001).

The AN model has a signal path that consists of a nonlinear, third-order narrow-band filter followed by a broader linear, first-order filter and an IHC/synapse module that produces the time-varying discharge rate $r(t)$ for a given CF. The bandwidth and gain of the nonlinear signal-path filter are varied according to the output of a control path, which consists of a nonlinear wideband filter followed by an OHC module. As the level of the stimulus energy that passes through the control-path filter increases, the bandwidth of the nonlinear signal-path filter is increased and the filter gain at CF is reduced, representing the reduction in cochlear gain at CF as stimulus level is increased above roughly 20 dB SPL (Ruggero *et al.* 1997). This single control mechanism in the AN model produces nonlinear tuning with both compression and suppression. Compression represents the reduction in cochlear gain at CF as the level of the stimulus near CF is increased (Ruggero *et al.*, 1997). Suppression represents the ability of off-CF energy to reduce the cochlear gain at CF (Ruggero, 1992). Stimulus energy that passes through the wideband control-path filter but not the narrow-band excitatory filter causes suppression (e.g., Geisler and Sinex, 1980). The wideband control-path filter was centered slightly higher in frequency than the signal-path filter (a basal shift of 1.2 mm along the basilar membrane; Heinz *et al.*, 2001c; see Zhang *et al.*, 2001 for details). The size of the basal shift and the bandwidth of the control-path filter were chosen to match AN two-tone suppression tuning curves from the literature (e.g., Sachs and Kiang, 1968; Delgutte, 1990b). The dynamics of the AN-model control path are extremely fast, with the ability of the control signal to follow cycle-by-cycle fluctuations in the stimulus determined by an 800-Hz low-pass filter in the OHC module. The fast dynamics in the model control path are consistent with the 200-μs basilar-membrane compression time constant reported by Recio *et al.* (1998) and the nearly instantaneous effects of two-tone suppression (e.g., Arthur *et al.*, 1971).

The low-level tuning of the nonlinear AN model was set based on low-level psychophysical estimates of human auditory tuning (Glasberg and Moore, 1990). The ERBs of the nonlinear signal-path filters at low levels were set to be 1.2 times smaller than the standard midlevel psychophysical estimates of auditory-filter ERBs described by Glasberg and Moore (1990) as a function of CF, a factor that is roughly consistent with their more complicated level-dependent equation for auditory-filter ERB (see Heinz *et al.* 2001c for details). High-level tuning in the AN model is broadened because the cochlear gain at CF is reduced as stimulus level increases, i.e., due to the filter gain/bandwidth trade-off. The maximal reduction in cochlear gain at CF across level (often referred to as the cochlear-amplifier gain), ranges from 20 dB at and below 500 Hz to roughly 55 dB at and above 8 kHz, which is consistent with basilar-membrane data from the chinchilla basal and apical turns (Ruggero *et al.*, 1997; Coo-
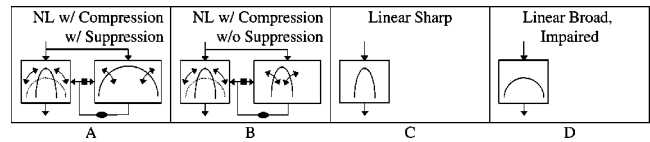

FIG. 1. Four versions of the phenomenological auditory-nerve model used to isolate the effects of compression and suppression on tuning. Only the section of the model that includes the peripheral tuning is illustrated; it includes the signal path (in A, B, C, D; only the initial third-order signal-path filter is shown; see the text) and the control path (in A, B; filled squares and ovals represent control-path modules that are the same in both nonlinear model versions). The curved arrows in A and B indicate that the filters fluctuate with level. A: The *nonlinear with compression and suppression* model version has a broad control-path filter. B: The *nonlinear with compression and without suppression* model has a narrow control-path filter. C: The *linear-sharp* model has sharp tuning and high gain at all stimulus levels, consistent with models A and B at low levels. D: The *linear-broad, impaired* model has broad tuning and low gain, consistent with the average tuning in models A and B at high stimulus levels.

per and Rhode, 1997) and with human psychophysical data (Hicks and Bacon, 1999; Glasberg and Moore, 2000). The bandwidth of the linear, first-order filter in the signal path equaled the time-averaged bandwidth of the nonlinear signal-path filter at high levels, i.e., with full reduction of the cochlear gain (Heinz *et al.*, 2001c; Zhang *et al.*, 2001). The effective ERB of the total signal-path filter at low levels was calculated computationally to be about 20% larger than the low-level ERB of the nonlinear signal-path filter for a 2-kHz CF, and about 25% larger for an 8-kHz CF.

Predictions from four versions of the AN model (Fig. 1) are compared to evaluate the role of several response properties associated with nonlinear cochlear tuning, including compression and suppression (Heinz *et al.*, 2001c). (1) The *nonlinear with compression and suppression* AN model is the standard version of the model with a level-dependent broad control-path filter [Fig. 1(A)]. (2) The *nonlinear with compression and without suppression* AN model controls the level-dependent gain and bandwidth of the signal-path filter by the stimulus energy within a narrow-band control-path filter that has the same level-dependent bandwidth and center frequency as the nonlinear signal-path filter [Fig. 1(B)]. The nonlinear models (#1 and 2) differ from one another in the spectral content of the stimulus that controls the level-dependent tuning. (3) The *linear-sharp* version of the AN model has low thresholds and linear signal-path filters with bandwidths set to the low-level tuning in the nonlinear AN model [Fig. 1(C)]. (4) The *linear-broad, impaired* version of the AN model has high thresholds and linear filters with broad bandwidths that match the time-averaged high-level tuning in the nonlinear AN model [Fig. 1(D)]; this model represents an impaired cochlea in which the cochlear-amplifier mechanism is absent (e.g., total OHC loss). Loss of the cochlear-amplifier gain, which increases as a function of CF (as discussed above), produces a sloping high-frequency hearing loss for this version of the AN model (Heinz *et al.*, 2001c).

Each version of the AN model was used to generate a population response for each of the three SR groups described by Liberman (1978). All AN fibers within an SR group had the same SR, threshold, and dynamic range for a CF-tone response (Heinz *et al.*, 2001c). Spontaneous rates of

60, 5, and 1 spikes/s were used for the high-, medium-, and low-SR populations, respectively. The physiological thresholds for high-SR fibers were chosen for simplicity to be near 0 dB SPL at all CFs, i.e., no external- or middle-ear filtering was included in this AN model. The lowest physiological thresholds observed in cat AN fibers are roughly 0 dB SPL and are found for high-SR fibers (Liberman, 1978; Miller *et al.*, 1997). Auditory-nerve thresholds were inversely related to SR in the model (Liberman, 1978). The elevated thresholds for low-SR fibers (~ 20 dB above the high-SR class) resulted in wide dynamic ranges for the low-SR fibers due to the compression in the signal-path filter response (Sachs and Abbas, 1974). The increased amount of compression at high CFs resulted in "straight" rate-level curves for high-CF low-SR fibers and "sloping saturation" rate-level curves at low CFs, consistent with the lack of straight rate-level curves observed for CFs below 1.5 kHz in guinea pigs (Winter and Palmer, 1991).

Sixty model CFs ranging from 200 Hz to 20 kHz were used to simulate the AN population response and were uniformly spaced in location according to a human cochlear map (Greenwood, 1990; as in Heinz *et al.*, 2001a).[1] This spacing of fiber CFs corresponded to roughly a 0.5-mm separation on the basilar membrane and is estimated to be about one-half of a human psychophysical ERB (Glasberg and Moore, 1990). The tone frequency was always chosen to be equal to one of the model CFs (e.g., the closest model CFs to the 2- and 8-kHz conditions used in the present study were 1927 and 8079 Hz, respectively, based on the roughly log-spaced cochlear map). Populations of high-, medium-, and low-SR fibers were simulated by assuming that 200 high-SR, 75 medium-SR, and 50 low-SR fibers were represented by each model CF. This represents all AN fibers within the frequency range from 200 Hz to 20 kHz [based on a total population of 30 000 AN fibers from 20 Hz to 20 kHz in human (Rasmussen, 1940; Heinz *et al.*, 2001a)], and is consistent with the 61%, 23%, and 16% distributions for high-SR, medium-SR, and low-SR fibers reported by Liberman (1978). All AN fibers were assumed to have independent Poisson discharge-generating mechanisms. Thus, the model AN-fiber responses were treated as conditionally independent stochastic point processes given the stimulus (Heinz *et al.*, 2001a); however, there were correlations across AN-fiber responses due to random stimulus fluctuations associated with the noise maskers (see below).

B. Signal detection theory

Heinz *et al.* (2001a) used SDT analysis with computational AN models to predict psychophysical performance limits based on the intrinsic variability in AN-discharge responses; however, this analysis was limited to deterministic stimuli. For masking studies with random-noise maskers, the effects of random stimulus fluctuations on the variability of AN responses must also be taken into account. The present study used the SDT analysis developed by Heinz (2000) to quantify detection performance limits due to random variations in the notched-noise stimuli as well as in the AN responses. Predicted performance was calculated for a processor based on *rate-place* information (as shown in Fig. 2 of

Heinz *et al.*, 2001a). In the rate-place analysis, the observations on which the detection decision is made are the population of spike counts $\{K_i\}_{i=1,\dots,M}$, where M is the total number of AN fibers. The counts are produced by M homogeneous Poisson processes (conditionally independent, given the stimulus) with rates equal to the average rates r_i produced by the AN model. This section provides an overview of the detection-in-noise analysis for rate-place information that was used in the current study, while the details of this approach for both rate-place information and *all information* (both temporal and average-rate information) are described by Heinz (2000).

A likelihood-ratio test (LRT) can be used to derive the form of an optimal processor based on a set of random observations (van Trees, 1968). For the case considered here in which the stimulus is random in addition to the AN discharges, the form of the optimal population processor can only be specified analytically if the processor is assumed to have knowledge of the average AN discharge rates $r_i(n)$ for each noise waveform n from the random-noise ensemble (Heinz, 2000); however, this is an unrealistic assumption for human listeners in a random-noise masking task. In order to remove this assumption, the same processor form as the optimal processor was used with the assumption that the processor only has knowledge of the *average-noise-response* properties (i.e., $x_i = E_n[r_i(n)]$, where E_n represents the expected value across the random-noise ensemble), and not the individual-noise responses $r_i(n)$. Thus, the predictions in the present study represent a *suboptimal* processor that is assumed to have knowledge of the average (across noise waveforms) AN discharge properties in response to the masked tone $[x_i(SN)]$ and in response to the notched-noise masker alone $[x_i(N)]$. The processor uses this knowledge of average noise responses in the same way that the optimal processor uses the knowledge of individual-noise responses.

The form of this general processor, which evaluates the number of observed discharges from the ith AN fiber based on the assumed *a priori* knowledge, $x_i(SN)$ and $x_i(N)$, is given by

$$Y_i(K_i) \triangleq \left[\ln \frac{x_i(SN)}{x_i(N)} \right] K_i + T \left[x_i(N) - x_i(SN) \right], \quad (1)$$

where T is the duration of the stimulus. The decision variable $Y_i(K_i)$ is a function of the random AN discharge count K_i, and thus is a random variable itself that depends on both the stimulus and AN variability. The reliability of this decision variable for indicating the presence of the tone depends on the difference in the mean value of $Y_i(K_i)$ between the signal-plus-masker (SN) and masker-alone (N) observation intervals, and on the variance of $Y_i(K_i)$. A useful metric for quantifying the sensitivity of a decision variable Y is

$$Q = \frac{(E[Y|SN] - E[Y|N])^2}{\text{Var}[Y|N]}, \quad (2)$$

where detection threshold is defined as the signal level for which the sensitivity index $Q = 1$. This sensitivity metric Q is similar to $(d')^2$, and represents a complete characterization of performance when the decision variable Y is Gaussian and has equal variance in both observation intervals

(Green and Swets, 1966; van Trees, 1968). These two assumptions are reasonably accurate for characterizing just-detectable differences in decision variables that are based on the total population of all AN discharges (Siebert, 1968, 1970; Colburn, 1969, 1973, 1981; Heinz *et al.*, 2001a). Actual deviations from these assumptions do not significantly affect the characterization of performance based on the sensitivity metric Q (Colburn, 1981).

Performance based on M AN fibers was calculated for a population decision variable

$$Y = \sum_{i=1}^{M} Y_i(K_i). \tag{3}$$

The equal-weighting combination in Eq. (3) behaves in an intuitive way based on Eq. (1). Fibers for which the signal does not (on average) change the discharge rate in response to the masker [i.e., $x_i(SN) = x_i(N)$] do not contribute to the population decision variable Y because $Y_i(K_i) = 0$ for any observed discharge count K_i. However, as described above the equal-weighting combination is not necessarily optimal, because across-fiber correlations are not accounted for in the way information is combined across fibers. The population sensitivity index Q based on this suboptimal processor [Eqs. (1) and (3)] is given by

$$Q = \left\{ \sum_{i=1}^{M} \left[\ln \frac{x_i(SN)}{x_i(N)} \right] [x_i(SN) - x_i(N)] \right\}^2 \Bigg/$$
$$\left(\frac{1}{T} \sum_{i=1}^{M} \left[\ln \frac{x_i(SN)}{x_i(N)} \right]^2 x_i(N) + \mathrm{Var}_n \left\{ \sum_{i=1}^{M} \left[\ln \frac{x_i(SN)}{x_i(N)} \right] r_i(n|N) \right\} \right), \tag{4}$$

where Var_n represents the variance across the random-noise ensemble (Heinz, 2000). Note that the variance of the decision variable Y [the denominator of Eq. (4)] is separated into two terms, where the first term represents the contribution of *AN variability* and the second term represents the contribution of *stimulus variability*.

Detection performance was predicted based on the individual high-, medium-, and low-SR populations of AN fibers (as described above), as well as on the total AN-fiber population. For predictions based on a population of AN fibers, the effect on performance of any potential correlation between AN fibers of different SR or CF due to a common random-stimulus drive was accounted for by including all AN fibers in the three summations in Eq. (4). The summation across AN fibers in the numerator and in both denominator terms differs from summing individual values of Q_i. The sensitivity indices Q_i for individual fibers cannot be simply added to obtain the total sensitivity because the fibers' responses are potentially correlated when stimulated with random-noise stimuli (see Heinz, 2000). In general, if the stimulus-induced contribution to the variance [second denominator term in Eq. (4)] of the decision variable Y dominates the intrinsic-AN contribution [first denominator term in Eq. (4)], then AN fibers with similar CFs will be correlated.

Psychophysical detection thresholds based on the AN model were simulated using Eq. (4). For a fixed signal level and masker notch width, the sensitivity index Q was esti-mated based on AN-population responses to the signal-plus-masker and masker-alone conditions for ten individual noise waveforms n from the random-noise ensemble. The adequacy of using only ten noise waveforms to estimate E_n and Var_n was evaluated by verifying in several conditions that the results did not change significantly when more noise waveforms $(20–40)$ were used. The sensitivity index Q was evaluated as a function of level for each notch width. The signal level at which $Q = 1$ was defined as representing the psychophysical detection threshold and was determined by interpolation.

C. Notched-noise method for estimating auditory filters

Most psychophysical methods for estimating auditory-filter shapes use the power-spectrum model of masking to derive the best auditory filter to explain a set of detection data (Moore, 1995a). The power-spectrum model assumes that the psychophysical detection threshold corresponds to a fixed long-term signal-to-noise ratio (SNR) at the output of the auditory filter. To estimate the psychophysical auditory filter, signal power at detection threshold, P_S, is measured for various masker spectra $N(f)$, and the best auditory-filter shape $W(f)$ to explain the set of detection data is derived based on the constant-SNR assumption. To simplify the fitting of the detection data, a class of auditory-filter shapes $W(f)$, which can be specified by a few parameters, is often assumed.

In the present study, psychophysical detection thresholds were predicted from the AN model (as described above) for a tone in the presence of two noise bands that were above and below the tone frequency f, had bandwidths equal to $0.4f$, and had a *fixed* noise spectrum level ($N_0 = 20$ dB SPL). The tone was always centered arithmetically between the two noise bands, i.e., only symmetric notches were used in the present study. The notch width Δf was defined as the frequency separation between the tone and the edge of each noise band that was closest to the tone frequency, and is referred to in terms of the normalized frequency separation $\Delta f/f$. Psychophysical detection thresholds were predicted for normalized frequency separations of 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6.

The predicted AN-model detection thresholds were then used to derive an estimate of the psychophysical auditory filter $W(f)$ from a common class of auditory filters using the power-spectrum model in the usual manner (Glasberg and Moore, 1990). The class of rounded-exponential filters, roex(p,r), is specified by a parameter p that describes the slope of the filter and a parameter r that controls the filter dynamic range (Patterson *et al.*, 1982; also see Moore, 1995a).[2] The estimated psychophysical auditory filter represents the roex(p,r) filter that best describes the AN-model detection thresholds predicted from the SDT analysis. The goodness of fit is quantified in terms of the root-mean-squared deviation in dB between the predicted thresholds from the power-spectrum model and the AN-model de-

J. Acoust. Soc. Am., Vol. 111, No. 2, February 2002

Heinz *et al.*: Effects of cochlear nonlinearity on auditory filters   1001

tection thresholds. The estimated psychophysical auditory filters from the fitting procedure are described in terms of the roex-filter parameters p and r, and the ERB (which equals $4f/p$). The power-spectrum model also includes the variable K_{psm}, which represents the constant signal-to-noise ratio at the output of the auditory filter at detection threshold. The corrections described by Glasberg and Moore (1990) for the transmission characteristics of the middle and external ear and for the earphone frequency response were not needed in the fitting procedure because the AN model had constant thresholds as a function of frequency (Heinz *et al.*, 2001c).

III. RESULTS

A. The relation between model peripheral filters and estimates of psychophysical auditory filters

First, estimates of psychophysical auditory filters were derived from predicted detection performance based on the total AN population using a version of the AN model with linear tuning. The ERBs of these estimated psychophysical auditory filters can be compared directly to the ERBs of the linear peripheral filters of the AN model at the signal frequency.

Predicted performance for the *linear-sharp* version of the AN model is shown in Fig. 2 for the three SR populations (triangles) as well as for the total AN population (circles). Detection thresholds for the rate-place model are plotted as a function of notch width for fixed-spectrum-level ($N_0 = 20$ dB SPL) notched-noise maskers. Note that the predicted detection thresholds for the total-AN population did not correspond to one SR group for all notch widths. Rather, detection performance was dependent on the low-SR population for small notch widths and on the high-SR population for large notch widths due to the differences in threshold and dynamic range across SR groups. The on-frequency masker energy in the zero-notch-width condition drove the high- and medium-SR fibers closer to saturation than the low-SR fibers, which have a higher threshold and wider dynamic range. The dependence of performance on low-SR fibers was stronger at higher levels (not shown), consistent with AN-fiber responses to tones in noise (Young and Barta, 1986). Thus, estimates of psychophysical auditory filters are derived from detection performance that is based on different AN fibers as a function of notch width. It is unlikely that estimated auditory filters based on individual AN fibers (or on single SR classes, as discussed below) would consistently correspond with those estimated psychophysically. However, this does not necessarily imply that psychophysical estimates of auditory filters do not represent peripheral tuning, because same-CF AN fibers with different SRs have similar tuning as a function of CF (Liberman, 1978).

The power-spectrum model was used to derive estimates of psychophysical auditory filters from the AN model predictions based on the total population (combined across all three SR groups). The resulting fit is shown in Fig. 2 (solid curve), and a summary of the roex-filter parameters used by the power-spectrum model to fit these data (circles) is given in Table I. The fit of the power-spectrum model to the predicted AN-model detection thresholds is quite good (1.4 dB rms
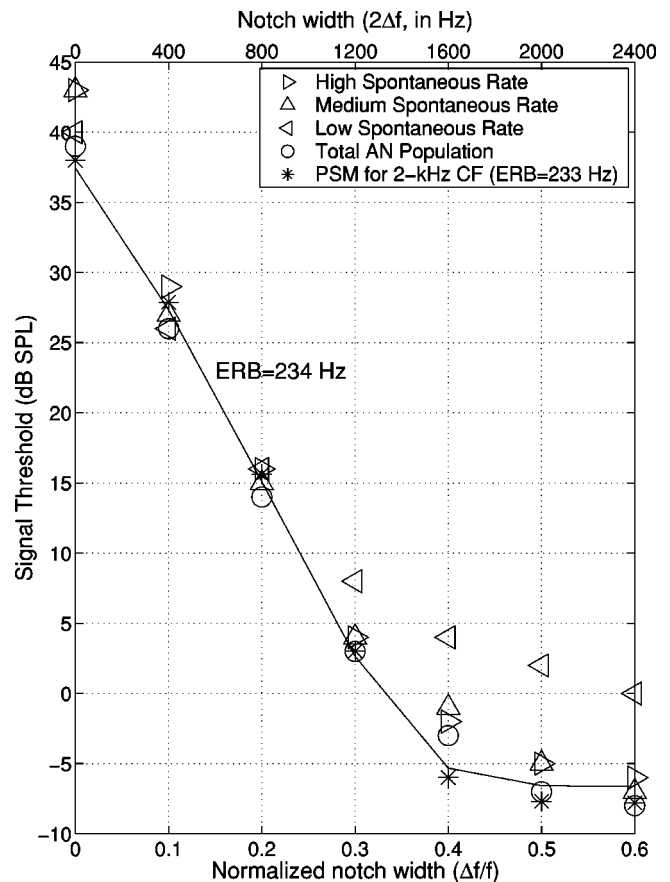

FIG. 2. Predicted detection thresholds for a tone in a notched-noise masker as a function of notch width based on the *linear-sharp* AN model. Model predictions are for a 2-kHz, 100-ms (20-ms rise/fall) tone presented simultaneously with the masker, which had a spectrum level of $N_0 = 20$ dB SPL. Predicted thresholds for the AN model are shown for each of the three SR populations (triangles), as well as for the combination of all three SR groups (circles). The solid curve represents the best fit from the power-spectrum model to the total-AN-population detection thresholds (circles) that were used to derive the estimated psychophysical auditory filter [ERB given in the figure; all auditory-filter parameters given in Table I]. The detection thresholds and estimated psychophysical-auditory-filter ERB derived from the AN model match very closely to the predicted thresholds (stars) from the power-spectrum model (PSM) based on the effective ERB of the 2-kHz peripheral model filter (see the text), which was 233 Hz.

deviation). Shailer *et al.* (1990) reported rms deviations between power-spectrum model fits and human psychophysical detection thresholds that were typically between 0.9 and 2.2 dB and were always less than or equal to 2.7 dB. The ERB of

TABLE I. Psychophysical-auditory-filter parameters derived from the power-spectrum model for Fig. 2. The derived estimate of the psychophysical auditory filter represents the roex(p,r) filter that best describes the rate-place detection thresholds predicted from the AN model. The roex-filter parameters used to predict thresholds from the power-spectrum model based on the effective 2-kHz peripheral model filter bandwidth (stars, Fig. 2) are also listed. ERB: equivalent-rectangular bandwidth; rms: root-mean-squared error; p: filter slope; r: filter dynamic range; K_{psm}: signal-to-noise at the auditory-filter output required for detection threshold.

	ERB (Hz)	rms (dB)	p	r (dB)	K_{psm} (dB)
Rate-place model	234	1.4	33.0	-52.3	-6.2
2-kHz peripheral model filter	233	$\cdots$	33.1	-54.0	-5.7

Heinz *et al.*: Effects of cochlear nonlinearity on auditory filters

the estimated psychophysical auditory filter derived from the total-population AN-model thresholds was 234 Hz. This estimate of the psychophysical-auditory-filter ERB matches very closely to the effective ERB (233 Hz; see Sec. II) of the 2-kHz-CF peripheral filter in the linear-sharp AN model. As discussed above, detection performance predicted from the total AN population was based on different SR classes at different notch widths. Thus, as expected, none of the estimated psychophysical-auditory-filter ERBs derived from individual SR classes [e.g., ERB(HSR)=214 Hz, ERB(MSR) =216 Hz, or ERB(LSR)=245 Hz, not shown] corresponded to the 234-Hz ERB of the estimated psychophysical auditory filter based on the total AN population.

For reference, detection thresholds predicted from the power-spectrum model based on the effective 2-kHz peripheral model filter with ERB=233 Hz are shown in Fig. 2 (stars). The roex-filter parameters used to predict the detection thresholds are given in Table I and were chosen to be consistent with critical ratios and absolute thresholds. Detection threshold for the zero-notch-width condition was set to 38 dB SPL, which is equal to the masker spectrum level (20 dB SPL) plus the human critical ratio at 2 kHz (18 dB, Fletcher, 1940), by adjusting the variable K_{psm}. Asymptotic detection thresholds for large notch widths, controlled by the parameter r, were set to be consistent with the AN-model rate-place absolute threshold of roughly -7 dB SPL at 2 kHz. The rate-place detection threshold is slightly less than the lowest AN-fiber physiological threshold because the threshold definitions differ and because information is combined across many AN fibers in the rate-place model.

Overall, the detection thresholds based on the total AN population were very close to the detection thresholds predicted from the power-spectrum model based on the 2-kHz peripheral model filter. Thus, estimates of psychophysical auditory filters based on rate-place detection thresholds from the linear-sharp AN model appear to be closely related to peripheral tuning at the signal frequency.[3]

B. The influence of compression and suppression on estimates of psychophysical auditory filters

If estimates of psychophysical auditory filters are directly related to peripheral tuning, as shown above for the linear-sharp AN model, then it is expected that changes in peripheral tuning associated with compression and suppression would affect psychophysical methods for estimating auditory filters. A demonstration of how AN response properties associated with nonlinear cochlear tuning influence estimates of psychophysical auditory filters is provided by comparing auditory-filter estimates based on detection thresholds from the four versions of the AN model.

Predicted detection thresholds for the four versions of the AN model are compared in Fig. 3 for a 2-kHz signal and a fixed masker spectrum level of $N_0=20$ dB SPL. Performance was predicted based on the combination of the three AN-model SR groups and plotted as a function of notch width. The power-spectrum model was used to derive estimates of psychophysical auditory filters from the AN-model detection thresholds, and the resulting fits (curves) and ERBs are shown in Fig. 3. A summary of the roex-filter parameters
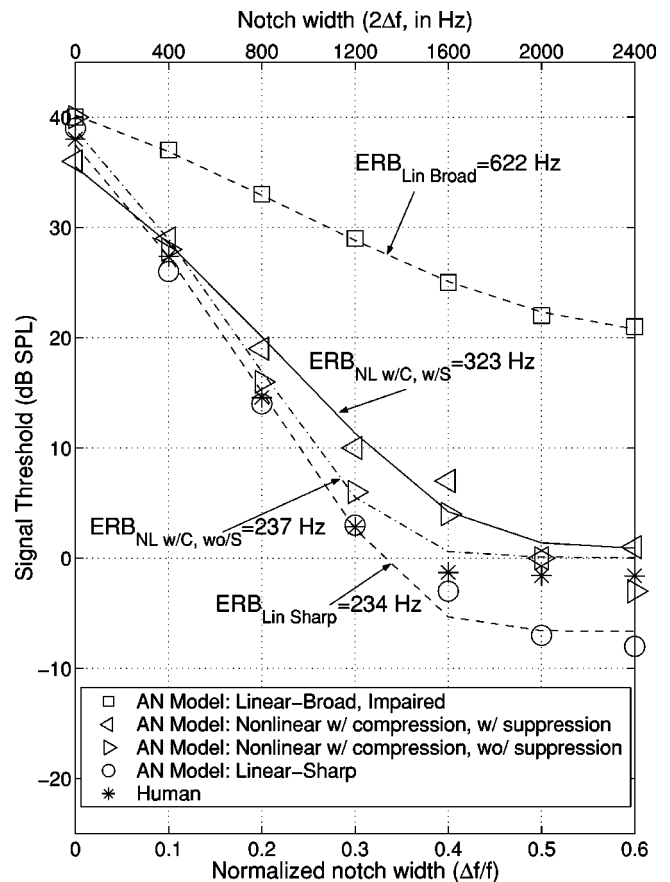

FIG. 3. Predicted detection thresholds for a 2-kHz tone in a notched-noise masker as a function of notch width for the four versions of the AN model. Same conditions as in Fig. 2. Predicted thresholds for the AN model are based on the combination of discharge counts across the three SR populations. Curves represent the best fits from the power-spectrum model [auditory-filter parameters given in Table II]. Also shown are typical human detection thresholds (stars) for $N_0=20$ dB SPL predicted from the power-spectrum model (see the text; Glasberg and Moore, 1990). The 2-kHz peripheral filter in the linear-sharp AN model had an effective ERB of 233 Hz and corresponds to the low-level tuning in the nonlinear AN models. The ERB of the peripheral filter in the linear-broad, impaired AN model was 613 Hz, corresponding to the ERB of the filter that represents a full reduction in the cochlear gain. The ERB of the estimated psychophysical auditory filter derived for the nonlinear AN model with compression and suppression was 323 Hz, corresponding to an overestimation of the ERB of the peripheral model filter describing low-level tuning to the signal.

used by the power-spectrum model to fit the AN-model detection thresholds is given in Table II. Also shown in Fig. 3 are typical human psychophysical detection thresholds (stars) for a masker spectrum level of $N_0=20$ dB SPL predicted from the power-spectrum model based on a psychophysical-auditory-filter ERB of 223 Hz (Glasberg and Moore, 1990). The roex-filter parameters used to predict typical human detection thresholds are given in Table II and were chosen to be consistent with human critical ratios (as in Fig. 2) and with an absolute threshold of 0 dB SPL at 2 kHz (Robinson and Dadson, 1956).

The lowest predicted detection threshold for each notch width (except zero) was for the linear-sharp AN-model version, while the highest threshold was always for the linear-broad, impaired version (Fig. 3). Detection thresholds from the two nonlinear versions of the AN model (with and without suppression) were in between those for the two linear

TABLE II. Psychophysical-auditory-filter parameters derived from the power-spectrum model for the 2-kHz (Fig. 3) and the 8-kHz (Fig. 4) signals. The roex-filter parameters used to predict typical human detection thresholds (stars) from the power-spectrum model for $N_0 = 20$ dB SPL are also listed (Glasberg and Moore, 1990).

	ERB (Hz)	rms (dB)	p	r (dB)	K_{psm} (dB)	Peripheral model filter ERB (Hz)
2 kHz						
Linear-sharp	234	1.4	33.0	-52.3	-6.2	233
Nonlinear, w/o suppression	237	1.9	32.6	-46.9	-4.9	233–613
Nonlinear, w/suppression	323	1.4	23.9	-41.4	-9.6	233–613
Linear-broad, impaired	622	0.2	12.6	-24.5	-7.6	613
Human	223	⋯	34.6	-48.0	-5.5	⋯
8 kHz						
Linear-sharp	932	2.1	34.7	-53.5	-6.4	934
Nonlinear, w/o suppression	931	1.7	34.7	-57.1	-2.0	934–6168
Nonlinear, w/suppression	1486	0.4	21.8	-48.7	-8.3	934–6168
Human	889	⋯	36.4	-57.0	-2.5	⋯

model versions and were within 5 dB of one another for all notch widths. Furthermore, for all notch widths except zero, the detection thresholds for the nonlinear AN model with suppression were higher than those for the nonlinear model without suppression, which in turn were higher than those for the linear-sharp model. Thus, for all off-frequency masking conditions, both compression and suppression in the AN model led to higher predicted psychophysical thresholds for detecting a tone in a notched-noise masker.

The fits from the power-spectrum model to the predicted AN-model detection thresholds in Fig. 3 were good for all versions of the AN model (Table II). For both linear versions of the AN model, the ERBs of the estimated psychophysical auditory filters were close to the ERBs of the peripheral model filters for the 2-kHz CF. The peripheral filters in the nonlinear versions of the AN model vary with level and time. At low stimulus levels, the tuning in the nonlinear AN model is linear and corresponds to the tuning in the linear-sharp AN model version. As stimulus level increases, tuning in the nonlinear AN model tends to broaden as the signal-path-filter gain is reduced; however, the fast dynamics of the control path result in a peripheral filter in the nonlinear AN model that fluctuates as a function of time. Thus, the instantaneous ERB of the 2-kHz nonlinear signal-path filter in the AN model can vary between 233 Hz, corresponding to the effective ERB of the filter that describes low-level tuning, and 613 Hz, corresponding to the ERB of the filter that represents a full reduction of the signal-path-filter gain (i.e., the filter in the linear-broad, impaired AN-model version). The time-average value of the nonlinear signal-path-filter ERB increases as the stimulus level through the control-path filter increases. Thus, the peripheral tuning in the nonlinear AN model with compression and suppression can be expected to be broader when the masker is present than when the signal is alone.

Despite higher predicted detection thresholds at all notch widths, the ERB of the estimated psychophysical auditory filter for the nonlinear AN model without suppression (237 Hz) was essentially the same as the ERB of the estimated psychophysical auditory filter derived for the linear-sharp

AN model (234 Hz). In contrast, the ERB of the estimated psychophysical auditory filter derived for the nonlinear AN model with both compression and suppression (323 Hz) was a factor of 1.38 larger than the ERB of the estimated psychophysical auditory filter for the linear-sharp AN model. This result suggests that the presence of suppression results in an estimated psychophysical-auditory-filter ERB that overestimates the ERB of the peripheral tuning to the signal alone at low levels. In response to the tonal signal alone at levels below roughly 20 dB SPL, the tuning in the nonlinear AN model equals the tuning in the linear-sharp AN model because the compression threshold for CF tones is 20 dB SPL (Heinz *et al.*, 2001c; see also Zhang *et al.*, 2001). Although the signal level at detection threshold for the nonlinear AN models was below 20 dB SPL for notch widths of 0.2 and greater (Fig. 3), the notched-noise masker contains energy that passes through the wideband control path and thus acts to broaden the peripheral tuning in the AN model.

While there was a clear effect of the AN-model nonlinearity on the detection thresholds and estimated psychophysical-auditory-filter ERBs for the 2-kHz CF (Fig. 3), a larger effect could be expected for higher-frequency CFs due to the increased strength of cochlear nonlinearity at higher frequencies. The maximal reduction in cochlear gain at CF across level was 30 dB at 2 kHz in the AN model and 55 dB at 8 kHz (Heinz *et al.*, 2001c). Figure 4 shows predicted detection thresholds for an 8-kHz signal as a function of notch width for the four versions of the AN model. Detection thresholds were predicted for a fixed-level masker ($N_0 = 20$ dB SPL) based on the combination of all three SR groups.

The relative roles of the SR groups across different notch widths (not shown) were the same as for the 2-kHz signal, i.e., low-SR fibers accounted for detection performance at small notch widths, while high-SR fibers accounted for detection at large notch widths. Detection thresholds for the linear-broad, impaired AN model were essentially invariant with notch width and represent the absolute rate-place threshold for an 8-kHz tone in the impaired model (Heinz *et al.*, 2001c). Full reduction of the cochlear gain by 55 dB at

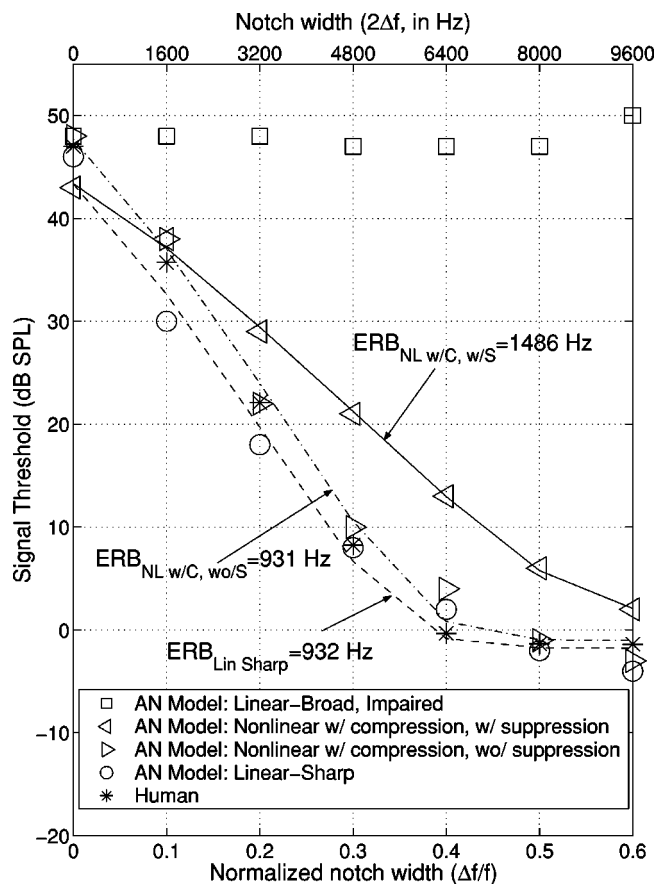Heinz *et al.*: Effects of cochlear nonlinearity on auditory filters

FIG. 4. Predicted detection thresholds for an 8-kHz tone in a notched-noise masker as a function of notch width for the four versions of the AN model (as in Fig. 3). Curves represent the best fits from the power-spectrum model [auditory-filter parameters given in Table II]. Also shown are typical human detection thresholds (stars) for $N_0 = 20$ dB SPL predicted from the power-spectrum model (see the text; Glasberg and Moore, 1990). The 8-kHz peripheral filter in the linear-sharp AN model had an effective ERB of 934 Hz, while the ERB of the peripheral filter in the linear-broad, impaired AN model was 6168 Hz. The factor by which the ERB of the estimated psychophysical auditory filter derived for the nonlinear AN model with suppression overestimated the ERB of the peripheral model filter describing low-level tuning to the signal was greater for the 8-kHz CF than for the 2-kHz CF (Fig. 3).

the 8-kHz CF raised the absolute rate-place detection threshold from roughly -7 dB SPL in the nonlinear AN model to 48 dB SPL in the impaired model. Typical human detection thresholds (stars in Fig. 4) for a masker spectrum level of $N_0 = 20$ dB SPL were predicted from the power-spectrum model based on a psychophysical-auditory-filter ERB of 889 Hz (Glasberg and Moore, 1990). The roex-filter parameters used to predict typical human psychophysical detection thresholds were chosen to be consistent with the 8-kHz human critical ratio (27 dB, Fletcher, 1940) and an absolute threshold of roughly 0 dB SPL. The use of an absolute threshold of 0 dB SPL in the present study is a simplification to match the absolute threshold of the AN model. Human absolute threshold at 8 kHz is about 18 dB SPL (Robinson and Dadson, 1956), primarily due to middle-ear filtering, which is not considered in the present study. The current predictions thus isolate the frequency-dependent effects of cochlear nonlinearity from those of middle-ear filtering; however, the potential effects of frequency-dependent audi-

bility could be evaluated in future studies by including a middle-ear filter in the AN model. Table II summarizes the power-spectrum-model fits (curves) to the predicted AN-model detection thresholds shown in Fig. 4 and the roex-filter parameters used to predict typical human detection thresholds. The power-spectrum model was not fit to the predicted detection thresholds for the linear-broad, impaired AN model because the thresholds were approximately invariant with notch width due to the high absolute tone threshold for this model version.

The same relative effects between predicted psychophysical detection thresholds for the different versions of the AN model were demonstrated for the 8-kHz signal (Fig. 4) as were observed for the 2-kHz signal (Fig. 3). The lowest detection threshold at each notch width except zero was for the linear-sharp AN model. Predicted thresholds for the nonlinear AN model without suppression were only slightly higher than for the linear-sharp model, but the detection thresholds for the nonlinear AN model with suppression were substantially higher for the middle notch widths. The large difference between the nonlinear AN models with and without suppression for notch widths larger than 0.1 was due to off-frequency masking energy passing through the wide-band control-path filter and acting to reduce the signal-path-filter gain, and thus to broaden the peripheral tuning in the model.

The ERB of the estimated psychophysical auditory filter for the linear-sharp AN model (932 Hz) was very close to the effective ERB of the 8-kHz peripheral model filter (934 Hz), similar to the result for the 2-kHz CF (Table II). Again, none of the estimated psychophysical auditory filters derived from the individual SR classes [e.g., ERB(HSR) = 850 Hz, ERB(MSR) = 864 Hz, or ERB(LSR) = 1022 Hz, not shown] provided as accurate an estimate of the underlying peripheral tuning as the estimated psychophysical auditory filter based on the total AN population.

The instantaneous ERB of the 8-kHz peripheral model filter in the nonlinear AN models can vary between 934 Hz at low levels and 6168 Hz at high levels based on the output of the control-path filter. Similar to the 2-kHz signal condition, the ERB of the estimated psychophysical auditory filter for the nonlinear AN model without suppression (931 Hz) was very close to the ERB of the estimated psychophysical auditory filter derived for the linear-sharp AN model (932 Hz). In contrast, when suppression was included in the nonlinear AN model, the ERB of the estimated psychophysical auditory filter (1486 Hz) was a factor of 1.59 larger than the ERB of the estimated psychophysical auditory filter derived for the linear-sharp AN model. This factor was larger than the corresponding factor of 1.38 for the 2-kHz CF; thus, as expected, the degree by which the estimated psychophysical-auditory-filter ERB overestimates the low-level peripheral tuning bandwidth was larger for the 8-kHz CF than for the 2-kHz CF. The larger difference at higher frequencies was due to the increased strength of cochlear nonlinearity at high frequencies in the AN model.

IV. DISCUSSION

For the AN model with linear tuning, the estimated psychophysical auditory filters were shown to be closely related to the peripheral filters in the AN model. Thus, detection predictions from the rate-place model were used in the present study to evaluate the effects of compression and suppression on estimates of psychophysical auditory filters.

A. The influence of nonlinear cochlear tuning on estimates of psychophysical auditory filters

The primary benefit of the modeling approach taken in the present study was to provide a quantitative framework in which to explore the effect of nonlinear AN response properties on estimates of psychophysical-auditory-filter bandwidths. It is not surprising that the inclusion of nonlinear AN response properties associated with broadened peripheral tuning resulted in broader estimates of psychophysical auditory filters. However, the specific properties related to the nonlinear control of peripheral tuning have potentially important implications for the interpretation of psychophysical-auditory-filter estimates. While several of the fundamental implications were demonstrated directly in the present study, many others are clearly suggested from this initial demonstration and it will be important to investigate these quantitatively in future studies.

1. Nonlinear control of peripheral tuning

Many nonlinear AN response properties, often described separately, are consistent with a single underlying mechanism that controls peripheral tuning (e.g. Sachs and Abbas, 1974; Sewell, 1984; Patuzzi et al., 1989; Ruggero and Rich, 1991; Ruggero et al., 1992; see the review by Ruggero, 1992). The phenomenological AN model used in the present study included both compression and suppression properties by accounting for the general property that both on- and off-frequency stimulus energy can act to control peripheral tuning. This property suggests that both the signal and the masker can influence peripheral tuning, and the fast dynamics of nonlinear cochlear tuning (Arthur et al., 1971; Recio et al., 1998) implies that both the temporal and spectral relation of the signal and masker must be considered when interpreting the results of psychophysical masking studies. Thus, it is improper to discuss the tuning at a given CF without specifying the stimulus configuration. For example, Kiang and Moxon (1974) showed that AN tuning curves in cat were broader and CF thresholds were higher in the presence of a bandpass noise centered well below CF than they were in the presence of the signal alone. This effect is consistent with the difference in human psychophysical frequency selectivity estimated using simultaneous masking (for which both the signal and the masker influence the peripheral tuning) and nonsimultaneous masking (for which the signal and the masker influence peripheral tuning independently) (e.g., Houtgast, 1977; Moore, 1978; Moore and Glasberg, 1981, 1982, 1986; Moore et al., 1987). Thus, the nonlinear control of cochlear tuning poses a basic problem for the interpretation of psychophysical estimates of frequency selectivity because the noise maskers affect the tuning of the peripheral filter that is being measured.

2. Implications of the nonlinear control of peripheral tuning for the interpretation of psychophysical-auditory-filter estimates

The results from the present study show that suppression produces an estimated psychophysical auditory filter with an ERB that overestimates the ERB of peripheral tuning to the signal alone at low levels. The low masker spectrum level used in the present study ($N_0 = 20$ dB SPL) was chosen to evaluate whether compression and suppression influenced estimates of psychophysical auditory filters near the lowest spectrum levels that have been used in psychophysical studies. The phenomenological AN model suggests that any condition for which enough stimulus energy (signal and masker) passes through the wideband suppression filter to reduce the cochlear gain would result in estimates of psychophysical auditory filters that were broader than low-level peripheral tuning. The compression threshold for CF tones is about 20 dB SPL in the model, consistent with physiological measures of basilar-membrane compression (Ruggero et al., 1997). Thus, any noise masker that produces roughly 20 dB SPL or more of overall level through the suppression filter would be expected to reduce the cochlear gain and thus to broaden peripheral tuning. Psychophysical studies that estimate auditory filters rarely use noise spectrum levels less than 20 dB SPL, and AN suppression bandwidths are much larger than 100 Hz for both low and high CFs (Delgutte, 1990b). Thus, it is likely that most psychophysical estimates of auditory filters include the effects of suppression, and therefore overestimate the bandwidth of low-level peripheral tuning to the signal alone. Consistent with this prediction, Rosen et al. (1998) found that psychophysical estimates of auditory filters were nonlinear down to the lowest stimulus levels that they could measure. A consequence of this limitation is that the difference between the frequency selectivity of normal and impaired listeners may be underestimated using current psychophysical methods, because the largest difference would be expected to occur for stimulus levels just above absolute threshold, where suppression was not invoked.

The debate over how to estimate psychophysical auditory filters as a function of level has focused on which aspect of the stimulus controls the level dependence of the psychophysical auditory filter, e.g., signal or masker level, overall level or level per ERB (Moore, 1995a). In addition, the debate has included whether the dependence is on the stimulus level prior to filtering (e.g., Glasberg and Moore, 1990) or on the level of the filtered stimulus (e.g., Rosen and Baker, 1994; Rosen et al., 1998). Glasberg and Moore (2000) have recently suggested that the gain of their tip (sharp) filter might depend on the output of their tail (broad) filter. The current results suggest that none of these views is entirely correct because they fail to account for the effect of suppression on peripheral tuning. Contrary to the current debate, peripheral tuning depends on both the signal and the masker, and the level dependence is based on the stimulus energy that passes through the suppression filter rather than the excitatory filter. This property suggests that the underlying peripheral tuning is likely to vary across the different notch widths used to estimate psychophysical auditory filters (also see Sinex and Havey, 1986; Delgutte, 1990a). Thus, as suggested

Heinz et al.: Effects of cochlear nonlinearity on auditory filters

by Rosen *et al.* (1998) for fixed-level masking conditions, psychophysical estimates of auditory filters are likely to represent some sort of average auditory filter over the different notch widths. Future modeling studies that evaluate the effect of stimulus level on estimates of psychophysical auditory filters are needed to investigate this issue further.

The present study predicts that the increase in strength of cochlear nonlinearity with CF could introduce a CF-dependent bias in psychophysical estimates of auditory filters such that the bandwidth of high-CF peripheral tuning is overestimated to a greater extent than the bandwidth of low-CF tuning. This predicted CF-bias is consistent with recent noninvasive measures of human cochlear tuning based on the phase response of stimulus-frequency otoacoustic emissions. Shera and Guinan, (2000) and Shera *et al.* (2002) described the frequency dependence of an emission-based measure of the sharpness of tuning derived from the theory of coherent reflection filtering (Shera and Zweig, 1993; Zweig and Shera, 1995) for humans, cats, and guinea pigs. The emission-based measure of sharpness of tuning increased with frequency for all three species, consistent with neural measures of sharpness ($Q = CF/ERB$) for cats and guinea pigs (e.g., Liberman, 1978; Miller *et al.*, 1997). In contrast, the increase in the human emission-based measure of sharpness with frequency was inconsistent with the essentially constant sharpness of tuning above 1 kHz described by the frequency dependence of psychophysical estimates of auditory-filter ERBs (Glasberg and Moore, 1990).

The role of suppression in broadening peripheral tuning suggests an important implication for the use of asymmetric notched-noise maskers to estimate psychophysical-auditory-filter asymmetry. Studies that have used asymmetric notch widths have consistently shown that psychophysical estimates of auditory filters derived from the power-spectrum model become more asymmetric as stimulus level increases, with the low-frequency side of the derived auditory filters becoming shallower as level increases (Moore and Glasberg, 1987; Glasberg and Moore, 1990, 2000; Rosen and Baker, 1994; Moore, 1995a; Rosen *et al.*, 1998). However, the results from the present study suggest that the relative influence of suppression above and below the signal frequency must be considered in an interpretation of psychophysical estimates of auditory filters derived using asymmetric notch widths. Delgutte (1990b) systematically measured the growth of two-tone suppression in AN fibers as a function of suppressor level and found that suppression growth was asymmetric above and below CF. The growth rate of suppression was much higher for suppressor frequencies below CF (ranging from 1–3 dB/dB) than for suppressor frequencies above CF (ranging from 0.15–0.7 dB/dB). The much faster growth rate of suppression below CF than above CF is consistent with the much greater effect of level on broadening the low-frequency side of psychophysical estimates of auditory filters than the high-frequency side. In addition, the asymmetry in psychophysical estimates of auditory filters for 2- and 4-kHz signals was reported to be greater than that for lower-frequency signals (Glasberg and Moore, 2000), which is consistent with the greater asymmetry in AN suppression growth rates for CFs above 2 kHz (Delgutte, 1990b). The combination of faster growth rates and higher thresholds for suppression below CF than above CF produces suppression that is stronger above CF at low levels and stronger below CF at high levels (Delgutte, 1990b). This result is consistent with psychophysical estimates of auditory filters having shallower high-frequency slopes at low levels, shallower low-frequency slopes at high levels, and symmetric slopes at medium levels (Moore, 1995a). Thus, the present study suggests that the reported changes in the asymmetry of psychophysical-auditory-filter estimates as a function of level may be due to (or at least influenced by) the asymmetry in the growth rate of suppression above and below CF, and not solely to actual changes in excitatory-filter asymmetry. Future studies are needed to explore this issue quantitatively.

The nonlinear control of peripheral tuning may have implications for comparing estimates of frequency selectivity derived from different psychophysical methods. Techniques such as Fletcher's (1940) band-widening (critical-band) method, the critical ratio, psychophysical tuning curves, and the notched-noise method have all been criticized based on methodological issues, such as the role of off-frequency listening, the varying influence of random-noise fluctuations as a function of masker bandwidth, and the assumption of rectangular filters (Moore, 1995a). However, a more fundamental issue is whether the underlying peripheral tuning at the frequency of interest is the same in each experiment, given that these methods use different stimulus configurations that may affect the control of peripheral tuning in different ways. The critical-band and critical-ratio methods use noise maskers of different bandwidths centered at the tone frequency, while the notched-noise and psychophysical-tuning-curve methods primarily use off-frequency maskers. Each of these methods involves the physiological mechanisms of compression and suppression in vastly different ways, and thus it is likely that the bandwidths of the underlying peripheral filters are different for these common psychophysical methods for estimating auditory frequency selectivity. For example, different bandwidths of the underlying peripheral filters are likely to be the cause for the common finding that frequency selectivity measured psychophysically using non-simultaneous masking is typically sharper than when measured using simultaneous masking (e.g., Houtgast, 1977; Moore, 1978; Moore and Glasberg, 1981, 1982, 1986; Moore and O'Loughlin, 1986; Moore *et al.*, 1987; see Moore, 1995a for a review). Thus, differences in psychophysical estimates of frequency selectivity across methods must be evaluated both in terms of methodological issues and in terms of the underlying peripheral tuning (e.g., Lentz *et al.*, 1999).

B. What do current psychophysical estimates of auditory filters represent?

Despite the complications that cochlear nonlinearity places on the interpretation of psychophysical estimates of auditory filters, these estimates have proven useful in predicting psychophysical data from a variety of masking conditions (Moore, 1995a). Derleth and Dau (2000) have suggested that their linear filter bank was successful in predicting masking patterns at midlevels because the effects

of cochlear nonlinearities were likely to be included in the psychophysical-auditory-filter bandwidths estimated using simultaneous notched-noise maskers, and that these filter bandwidths could be considered to represent "effective" auditory filters. The present study suggests that psychophysical auditory filters estimated using simultaneous masking may be appropriate for representing the average peripheral tuning in the presence of *both* the signal and the masker (i.e., the "effective" tuning), which is consistent with their success in accounting for similar masking conditions. Alternatively, nonsimultaneous masking may be more appropriate for estimating peripheral tuning in response to the *signal alone* by allowing the masking noise to be used primarily as an excitatory masker due to the fast dynamics of nonlinear cochlear tuning. However, the signal and the masker may pass through peripheral filters with different bandwidths in nonsimultaneous masking experiments, which would have important implications for comparing the amount of excitation produced by the signal and by the masker. A similar idea was used to model the additivity of nonsimultaneous masking based on the independent effects of compression on the signal and on the masker (Oxenham and Moore, 1994). Thus, even though psychophysical estimates of auditory filters based on nonsimultaneous masking may provide a better estimate of peripheral tuning in response to the signal alone, these estimates are likely to depend on peripheral tuning both in response to the signal alone and to the masker alone. Potential differences between signal-alone and masker-alone peripheral tuning can be minimized by measuring masked detection at stimulus levels as close to absolute threshold as possible.

The predicted detection thresholds in the present study were based on peripheral AN-model filters that were specified according to low-level psychophysical estimates of auditory filters (Glasberg and Moore, 1990); however, the present results suggest that these psychophysical estimates of auditory filters are likely to be broader than human peripheral tuning at low levels. A strength of the present study comes from demonstrating the issues that need to be considered to interpret the effects of cochlear nonlinearity on psychophysical estimates of auditory filters, and this demonstration does not depend significantly on the accuracy of the peripheral-filter ERBs in the AN model. In order to determine the appropriate ERBs to use for low-level human peripheral tuning, the variation in the strength of cochlear nonlinearity as a function of frequency and the bandwidth of the suppression filter must be determined for human listeners. Determination of whether the current model parameters are appropriate for human listeners will require future studies that can evaluate the strength of human cochlear nonlinearity by quantitatively relating physiological response properties to human psychophysical performance. The related physiological and psychophysical masking studies by Delgutte (1990a) and Oxenham and Plack (1998) provide a good basis for relating the strength of compression and suppression in cats and humans. Physiological and psychophysical measures of basilar-membrane compression (e.g., Ruggero *et al.*, 1997; Cooper and Rhode, 1997; Oxenham and Plack, 1997; Plack and Oxenham, 1998) provide useful data for relating

the strength of compression in chinchilla, guinea pig, and humans.

C. Implications of the cochlear-amplifier mechanism for comparing normal and impaired hearing

The present study demonstrated that the nonlinear response properties associated with the physiologically vulnerable cochlear-amplifier mechanism influence estimates of psychophysical auditory filters. This result suggests that a better quantitative understanding of these influences may allow for a direct psychophysical characterization of cochlear status in human listeners. Loss of the cochlear-amplifier mechanism in hearing-impaired listeners is likely to be caused by the loss of OHCs, and it represents a form of sensorineural hearing loss that is likely to be common in humans (Pickles, 1988; Patuzzi *et al.*, 1989; Ruggero and Rich, 1991; van Tasell, 1993; Moore, 1995b). The results from the current study are consistent with the view that listeners with OHC loss have broader psychophysical auditory filters than normal-hearing listeners, but that reduced frequency selectivity is not the only difference between normal and impaired listeners (Moore, 1995b).

The phenomenological AN model used in the present study highlights the idea that the difference between a normal-hearing listener with a cochlear-amplifier mechanism and a hearing-impaired listener without a cochlear-amplifier mechanism is not simply a difference between one system in two states, i.e., with low or high thresholds, narrow or broad tuning, compressive or linear magnitude responses. Rather, the difference is between two different systems: an impaired system that is static, insensitive, and has broad tuning, and a normal system with fast, dynamic tuning that is continuously changing in response to the stimulus. This view is different than the view that has arisen from the power-spectrum model of masking, i.e., that the primary differences in the impaired system are higher thresholds and broader tuning.

The difference between these two views is likely to be most significant for rapidly changing complex stimuli and may be less significant for steady-state stimuli. Peters *et al.* (1998) and Moore *et al.* (1999a) have reported that the difference between the ability of normal-hearing and hearing-impaired listeners to understand speech is much larger in the presence of temporally and spectrally varying backgrounds than for stationary broadband backgrounds, even when advanced amplification algorithms were provided to the hearing-impaired listeners. This result has been interpreted as representing a deficit in the ability of hearing-impaired listeners to "listen in the dips" (e.g., Peters *et al.*, 1998; Moore *et al.*, 1999a). This deficit has been suggested to result from degraded frequency selectivity and reduced temporal resolution, which have been accounted for in terms of broadened tuning and loss of compression, respectively (see Moore, 1995b, and Moore and Oxenham, 1998, for reviews). Thus, the absence of the dynamic peripheral tuning provided by the cochlear-amplifier mechanism is likely to be significant for stimulus conditions in which current amplification algorithms have the most limited benefit. The present modeling approach provides a quantitative method that could be used

in future studies to investigate the significance of the loss or degradation of dynamic cochlear tuning in hearing-impaired listeners.

V. CONCLUSIONS

(1) The ERB of the estimated psychophysical auditory filter that was derived based on the total AN-model population was closely related to the ERB of the peripheral AN-model filter at the signal frequency for model versions with linear peripheral tuning.

(2) The predicted detection performance that was used to derive estimates of psychophysical auditory filters was based on different AN-fiber SR classes at different notch widths due to the variation of threshold and dynamic range across SR. None of the estimated psychophysical auditory filters derived from individual SR classes provided as accurate an estimate of the underlying peripheral tuning as the estimated psychophysical auditory filter based on the total AN population.

(3) The ERBs of the estimated psychophysical auditory filters derived from the nonlinear AN model with suppression were always larger than the ERBs of the estimated psychophysical auditory filters from the linear-sharp AN model. This result suggests that psychophysically estimated auditory-filter ERBs represent an overestimate of the bandwidth of low-level peripheral tuning in response to the signal alone.

(4) The factor by which the ERB of the estimated psychophysical auditory filter derived from the nonlinear AN model overestimated the ERB of the peripheral AN-model filter was larger at 8 kHz (factor of 1.59) than at 2 kHz (factor of 1.38). This results suggests a potential CF-dependent bias in psychophysical estimates of auditory filters due to the increased strength of cochlear non-linearity at high CFs.

(5) The role of suppression in broadening peripheral tuning in response to the noise masker has important implications for the interpretation of psychophysical-auditory-filter estimates. The present study suggests that psychophysical estimates of auditory filters depend on the level of the signal and masker energy at the output of the wideband suppression filter. Thus, psychophysical auditory filters estimated using simultaneous masking are likely to represent some sort of average tuning in the presence of both the signal and the masker.

[1]The human cochlear map described by Greenwood (1990) was used in the present study, and is given by $f(x) = 165.4(10^{0.06x} - 0.88)$, where x is the distance (in mm) along the basilar membrane from the apex, and $f(x)$ is the frequency (in Hz) corresponding to the position x.

[2]The equation for the roex(p,r) filter shape is given by $W(g) = (1-r)(1+pg)\exp(-pg)+r$, where $g = |f-f_c|/f$ represents the normalized frequency relative to the filter center frequency f_c (Moore, 1995a). The parameter p determines the filter slope and bandwidth, while the parameter r controls the filter dynamic range.

[3]Estimates of psychophysical auditory filters derived from the all-information model (based on discharge times and counts) did not represent a good measure of peripheral tuning (see Heinz, 2000), and thus are not presented in this report. The all-information estimates of psychophysical auditory filters were typically narrower than the AN-model peripheral filters due to the differential effects of the random-noise masker on the temporal signal information for on- versus off-frequency masking conditions. In the zero-notch-width condition, the on-frequency random-masker energy degraded the temporal signal information, resulting in an all-information detection threshold that was only about 10 dB below the rate-place threshold. For the off-frequency conditions (0.1–0.6 notch widths), the temporal information about the signal was much more reliable, producing all-information thresholds that were roughly 25–30 dB below the rate-place thresholds.

Allen, J.B. (**2001**). "Nonlinear cochlear signal processing," in *Physiology of the Ear*, edited by A.F. Jahn and J. Santos-Sacchi (Singular Thomson Learning, San Diego), pp. 393–442.

Anderson, D.J., Rose, J.E., Hind, J.E., and 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.

Arthur, R.M., Pfeiffer, R.R., and Suga, N. (**1971**). "Properties of 'two-tone inhibition' in primary auditory neurons," J. Physiol. (London) **212**, 593–609.

Cheatham, M.A., and Dallos, P. (**1998**). "The level dependence of response phase: Observations from cochlear hair cells," J. Acoust. Soc. Am. **104**, 356–369.

Colburn, H.S. (**1969**). "Some physiological limitations on binaural performance," Ph.D. 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. (**1981**). "Intensity perception: Relation of intensity discrimination to auditory-nerve firing patterns," Internal Memorandum, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA.

Cooper, N.P., and Rhode, W.S. (**1997**). "Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea," J. Neurophysiol. **78**, 261–270.

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

Delgutte, B. (**1990a**). "Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers," J. Acoust. Soc. Am. **87**, 791–809.

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

Derleth, R.P., and Dau, T. (**2000**). "On the role of envelope fluctuation processing in spectral masking," J. Acoust. Soc. Am. **108**, 285–296.

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

Fletcher, H. (**1940**). "Auditory patterns," Rev. Mod. Phys. **12**, 47–65.

Fletcher, H. (**1953**). *Speech and Hearing in Communication* (Van Nostrand, New York).

Geisler, C.D., and Sinex, D.G. (**1980**). "Responses of primary auditory fi-

bers to combined noise and tonal stimuli,'' Hear. Res. **3**, 317–334.

Glasberg, B.R., and Moore, B.C.J. (**1986**). ''Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments,'' J. Acoust. Soc. Am. **79**, 1020–1033.

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

Glasberg, B.R., and Moore, B.C.J. (**2000**). ''Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise,'' J. Acoust. Soc. Am. **108**, 2318–2328.

Green, D.M., and Swets, J.A. (**1966**). *Signal Detection Theory and Psychophysics* (Wiley, New York; reprinted 1988 by Peninsula, Los Altos, CA).

Greenwood, D.D. (**1961**). ''Critical bandwidth and the frequency coordinates of the basilar membrane,'' J. Acoust. Soc. Am. **33**, 1344–1356.

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

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

Heinz, M.G., Colburn, H.S., and Carney, L.H. (**2001a**). ''Evaluating auditory performance limits. I. One-parameter discrimination using a computational model for the auditory nerve,'' Neural Comput. **13**, 2273–2316.

Heinz, M.G., Colburn, H.S., and Carney, L.H. (**2001b**). ''Rate and timing cues associated with the cochlear amplifier: Level discrimination based on monaural cross-frequency coincidence detection,'' J. Acoust. Soc. Am. **110**, 2065–2084.

Heinz, M.G., Zhang, X., Bruce, I.C., and Carney, L.H. (**2001c**). ''Auditory-nerve model for predicting performance limits of normal and impaired listeners,'' ARLO **2**, 91–96.

Hicks, M.L., and Bacon, S.P. (**1999**). ''Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing,'' J. Acoust. Soc. Am. **105**, 326–338.

Houtgast, T. (**1977**). ''Auditory-filter characteristics derived from direct-masking data and pulsation-threshold data with a rippled-noise masker,'' J. Acoust. Soc. Am. **62**, 409–415.

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

Lentz, J.J., Richards, V.M., and Matiasek, M.R. (**1999**). ''Different auditory filter bandwidth estimates based on profile analysis, notched noise, and hybrid tasks,'' J. Acoust. Soc. Am. **106**, 2779–2792.

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., and Dodds, L.W. (**1984**). ''Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves,'' Hear. Res. **16**, 55–74.

Lutfi, R.A., and Patterson, R.D. (**1984**). ''On the growth of masking asymmetry with stimulus intensity,'' J. Acoust. Soc. Am. **76**, 739–745.

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

Moore, B.C.J. (**1978**). ''Psychophysical tuning curves measured in simultaneous and forward masking,'' J. Acoust. Soc. Am. **63**, 524–532.

Moore, B.C.J. (**1995a**). ''Frequency analysis and masking,'' in *Hearing*, edited by B.C.J. Moore (Academic, New York), Chap. 5.

Moore, B.C.J. (**1995b**). *Perceptual Consequences of Cochlear Damage* (Oxford University Press, New York).

Moore, B.C.J., and Glasberg, B.R. (**1981**). ''Auditory filter shapes derived in simultaneous and forward masking,'' J. Acoust. Soc. Am. **70**, 1003–1014.

Moore, B.C.J., and Glasberg, B.R. (**1982**). ''Interpreting the role of suppression in psychophysical tuning curves,'' J. Acoust. Soc. Am. **72**, 1374–1379.

Moore, B.C.J., and Glasberg, B.R. (**1986**). ''Comparisons of frequency selectivity in simultaneous and forward masking for subjects with unilateral cochlear impairments,'' J. Acoust. Soc. Am. **80**, 93–107.

Moore, B.C.J., and O'Loughlin, B.J. (**1986**). ''The use of nonsimultaneous masking to measure frequency selectivity and suppression,'' in *Frequency Selectivity in Hearing*, edited by B.C.J. Moore (Academic, London), pp. 179–250.

Moore, B.C.J., and Glasberg, B.R. (**1987**). ''Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns,'' Hear. Res. **28**, 209–225.

Moore, B.C.J., Poon, P.W.F., Bacon, S.P., and Glasberg, B.R. (**1987**). ''The temporal course of masking and the auditory filter shape,'' J. Acoust. Soc. Am. **81**, 1873–1880.

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

Moore, B.C.J., Peters, R.W., and Stone, M.A. (**1999a**). ''Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips,'' J. Acoust. Soc. Am. **105**, 400–411.

Moore, B.C.J., Vickers, D.A., Plack, C.J., and Oxenham, A.J. (**1999b**). ''Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism,'' J. Acoust. Soc. Am. **106**, 2761–2778.

Oxenham, A.J., and Moore, B.C.J. (**1994**). ''Modeling the additivity of nonsimultaneous masking,'' Hear. Res. **80**, 105–118.

Oxenham, A.J., and Plack, C.J. (**1997**). ''A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing,'' J. Acoust. Soc. Am. **101**, 3666–3675.

Oxenham, A.J., and Plack, C.J. (**1998**). ''Suppression and the upward spread of masking,'' J. Acoust. Soc. Am. **104**, 3500–3510.

Patterson, R.D. (**1976**). ''Auditory filter shapes derived with noise stimuli,'' J. Acoust. Soc. Am. **59**, 640–654.

Patterson, R.D., Nimmo-Smith, I., Weber, D.L., and Milroy, R. (**1982**). ''The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold,'' J. Acoust. Soc. Am. **72**, 1788–1803.

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

Patuzzi, R.B., Yates, G.K., and Johnstone, B.M. (**1989**). ''Outer hair receptor currents and sensorineural hearing loss,'' Hear. Res. **42**, 47–72.

Peters, R.W., Moore, B.C.J., and Baer, T. (**1998**). ''Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people,'' J. Acoust. Soc. Am. **103**, 577–587.

Pickles, J.O. (**1984**). ''Frequency threshold curves and simultaneous masking functions in single fibres of the guinea pig auditory nerve,'' Hear. Res. **14**, 245–256.

Pickles, J.O. (**1988**). *An Introduction to the Physiology of Hearing* (Academic, New York).

Plack, C.J., and Oxenham, A.J. (**1998**). ''Basilar-membrane nonlinearity and the growth of forward masking,'' J. Acoust. Soc. Am. **103**, 1598–1608.

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

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

Robinson, D.W., and Dadson, R.S. (**1956**). ''A redetermination of the equal-loudness relations for pure tones,'' Br. J. Appl. Phys. **7**, 166–181.

Rosen, S., and Baker, R.J. (**1994**). ''Characterising auditory filter nonlinearity,'' Hear. Res. **73**, 231–243.

Rosen, S., Baker, R.J., and Darling, A. (**1998**). ''Auditory filter nonlinearity at 2 kHz in normal hearing listeners,'' J. Acoust. Soc. Am. **103**, 2539–2550.

Ruggero, M.A. (**1992**). ''Physiology and coding of sound in the auditory nerve,'' in *The Mammalian Auditory Pathway: Neurophysiology*, edited by A.N. Popper and R.R. Fay (Springer, New York), pp. 34–93.

Ruggero, M.A., and Rich, N.C. (**1991**). ''Furosemide alters organ of Corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane,'' J. Neurosci. **11**, 1057–1067.

Ruggero, M.A., Robles, L., and 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.

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

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

Sachs, M.B., and Abbas, P.J. (**1974**). ''Rate versus level functions for auditory nerve fibers in cats: Tone burst stimuli,'' J. Acoust. Soc. Am. **56**, 1835–1847.

Sewell, W.F. (**1984**). ''The effects of furosemide on the endocochlear potential and auditory-nerve fiber tuning curves in cats,'' Hear. Res. **14**, 305–314.

Shailer, M.J., Moore, B.C.J., Glasberg, B.R., Watson, N., and Harris, S. (**1990**). ''Auditory filter shapes at 8 and 10 kHz,'' J. Acoust. Soc. Am. **88**, 141–148.

Shera, C.A., and Zweig, G. (**1993**). ''Order from chaos: Resolving the para-

dox of periodicity in evoked otoacoustic emission," in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 54–63.

Shera, C.A., and Guinan, J.J., Jr. (**2000**). "Frequency dependence of stimulus-frequency-emission phase: Implications for cochlear mechanics," in *Recent Developments in Auditory Mechanics*, edited by H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (World Scientific, Singapore), pp. 381–387.

Shera, C.A., Guinan, Jr., J.J., and Oxenham, A.J. (**2002**). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proc. Natl. Acad. Sci. USA, in press.

Siebert, W.M. (**1968**). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns*, edited by P.A. Kolers and M. Eden (MIT Press, Cambridge, MA), pp. 104–133.

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

Sinex, D.G., and Havey, D.C. (**1986**). "Neural mechanisms of tone-on-tone masking: Patterns of discharge rate and discharge synchrony related to rates of spontaneous discharge in the chinchilla auditory nerve," J. Neurophysiol. **56**, 1763–1780.

van Tasell, D.J. (**1993**). "Hearing loss, speech, and hearing aids," J. Speech Hear. Res. **36**, 228–244.

van Trees, H.L. (**1968**). *Detection, Estimation, and Modulation Theory: Part I* (Wiley, New York), Chap. 2.

Vogten, L.L.M. (**1978**). "Low-level pure-tone masking: A comparison of 'tuning curves' obtained with simultaneous and forward masking," J. Acoust. Soc. Am. **63**, 1520–1527.

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).

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

Yates, G.K. (**1995**). "Cochlear structure and function," in *Hearing*, edited by B.C.J. Moore (Academic, New York), pp. 41–74.

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

Zhang, X., Heinz, M.G., Bruce, I.C., and 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.

Zweig, G., and Shera, C.A. (**1995**). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," J. Acoust. Soc. Am. **98**, 2018–2047.