ACTA ACUSTICA UNITED WITH ACUSTICA
Vol. 91 (2005) 980–991

Quadratic-Compression Model of Auditory Discrimination and Detection

Stephen T. Neely, Walt Jesteadt
Boys Town National Research Hospital, Omaha, Nebraska 68131, USA

Summary
Our model of detection and discrimination is based on the common assumption that, for psychoacoustic tasks dominated by a single auditory comparison, d-prime is equal to a perceived-intensity difference divided by the standard deviation of this perceived intensity. In general, the transformation from the external, physical intensity of an acoustic signal to an internal, perceived intensity is nonlinear, compressive, and similar to the transformation from physical intensity to loudness. We represent the compression associated with this intensity transformation by a quadratic function. This leads to explicit mathematical representation of the external-to-internal intensity transformation and its inverse. For psychoacoustic tasks such as intensity discrimination, increment detection, and forward masking, we can write explicit expressions for d-prime and signal threshold. Model predictions based on these expressions are in good agreement with experimental data from the literature, as well as observations from our laboratory. Some of these comparisons between model and data lead to predictions regarding the dependence of internal noise (or variance) on stimulus intensity for intensity discrimination and increment detection. Other comparisons suggest that the influence of neural adaptation on forward masking is greater than the influence of masker persistence. The quadratic-compression model may provide a useful framework for understanding a variety of simple psychoacoustic tasks.

PACS no. 43.66.Ba, 43.66.Cb, 43.66.Dc, 43.66.Fe

1. Introduction

This paper describes a modeling framework that may provide a guide to understanding simple psychoacoustic tasks. Our modeling approach is similar in many respects to the one suggested by Durlach and Braida [1] in that we derive an expression for the decision variable d' in terms of stimulus parameters and assume that performance is limited by variability in perceived sound intensity. Our modeling objectives parallel those of several recent models of detection and discrimination that specify multiple stages of signal processing followed by a decision process (e.g., [2, 3, 4, 5, 6]).

To arrive at explicit formulas that can be used in place of simulations, we have adopted a simplified model of the decision process. We assume that performance on psychoacoustic tasks can be described in a detection-theory framework in terms of d' [7] and that performance is limited by variability in perceived intensity, which may come from either external or internal sources. We are concerned with estimating the magnitude of perceived- intensity variance in all experimental conditions, but we make no assumptions about the decision process beyond the standard assumption that d' is determined by the ratio of the mean value of a random variable to its standard deviation.

Received 1 March 2005, Revised 7 October 2005,
accepted 10 October 2005.

We are particularly concerned with understanding the consequences of compression in the peripheral auditory system, the focus of modeling efforts described by Yates *et al.* [8], Moore and Oxenham [4], Plack and Oxenham [6], and others. Our approach to modeling compression differs from other recent models in that we have selected a mathematically-tractable functional form to represent the nonlinear transformation from physical intensity to perceived intensity. Likewise, we have chosen to represent temporal integration in a simplified form that can be expressed as an element in an equation. We show how this approach can be used to derive explicit formulas for signal threshold in intensity discrimination, increment detection and forward-masking tasks. By *increment detection* we mean the detection of a change in level of an ongoing stimulus, either a continuous tone or one that is turned on before the increment occurs, while *intensity discrimination* refers to the detection of a change in level where the listener is asked to compare two stimuli of equal frequency and duration that differ only in level. Although our model could be used to generate predictions for other stimulus configurations and for a wide range of stimulus parameters, we have focused initially on data obtained recently in our laboratory and closely related data reported in the literature.

One drawback of our approach that should be mentioned is the need to tune various parameters for different tasks or specific conditions. The requirement to know

© S. Hirzel Verlag · EAA

stimulus parameters inevitably limits the use of any model to predict performance for new conditions. We cannot offer general rules to describe how every signal characteristic influences an auditory decision, because the underlying physics and physiology are much more complex than our simple model allows. At this point, our results should be viewed as preliminary and our model regarded as a set of working hypotheses that we hope will offer new insights or lead to further progress in understanding auditory detection and discrimination.

2. Model derivation

In order to derive a model of detection and discrimination, we assume that the auditory system transforms an external sound with physical intensity I_s into a perceived intensity. The perceived intensity is a random variable that has a mean value N_s and a standard deviation σ_s. Although N_s is essentially the same as *loudness*, we call it *perceived intensity* to avoid assuming an exact equivalence between these two concepts. The sensory transformation $\mathcal{I}$ determines the mean perceived intensity N_s for any physical intensity I_s.

$$N_s = \mathcal{I}(I_s) \tag{1}$$

We assume that the sensory transformation can be deduced, approximately, by experimental measurements. In general, the sensory transformation is not invariant across auditory tasks and may depend on specific characteristics of the sound, such as its center frequency or bandwidth. Although it is of interest to know how this transformation varies across normal-hearing subjects and how it is affected by hearing loss, these issues are not the focus of this paper. Here we restrict our consideration to tasks involving tones at one frequency and do not consider intersubject variability. By doing this, we are able to keep the sensory transformation constant through most of cases that we consider.

We define the decision variable d' as the ratio of perceived-intensity difference ΔN_s to standard deviation σ_s.

$$d' = \Delta N_s / \sigma_s. \tag{2}$$

The perceived-intensity difference ΔN_s in the numerator of equation (2) may represent the difference between *two* perceived intensities or may represent a *single* perceived intensity, depending on the specific task. We will focus attention on simple auditory tasks in which a decision could be based on information that is available within a single auditory "channel". The total variance of the perceived intensity σ_s^2 is the sum of contributions from two sources: (1) variance due to internal processing variability and (2) variance due to the presence of external noise. In our comparisons of model results with psychophysical measurements, we select cases that allow σ_s^2 to be simplified as much as possible, in order to allow explicit expressions to be derived for the perceived intensity at threshold. The validity of our expressions for d' will be tested by comparing our model results with representative data.

If we take the external sound level to be the physical sound intensity expressed in decibels and the internal sound level to be the perceived intensity similarly expressed in decibels, then, for most steady-state sounds, the sensory transformation has a compressive nature, such that the internal sound level has nearly a logarithmic dependence on the external sound level [9]. For convenience, we will arbitrarily define both the external and internal sound level to be zero at perceptual threshold in quiet, when $d' = 1$, which is sometimes called *absolute threshold*. We use the same log transformation $\mathcal{L}$ to represent the transformation from *intensity to level* for both external and internal versions of these quantities.

$$L_s = \mathcal{L}[I_s] = 10 \log_{10} \frac{I_s}{I_{st}}, \tag{3}$$

$$E_s = \mathcal{L}[N_s] = 10 \log_{10} \frac{N_s}{N_{st}}. \tag{4}$$

The $\mathcal{L}$ transformation includes an implicit threshold parameter (such as I_{st} or N_{st} in the equations above) to make its value zero at threshold.

In order to construct our sensory transformation $\mathcal{I}$, it is useful to define the intermediate transformation from external level L_s to internal level (or excitation) E_s.

$$E_s = \mathcal{E}[L_s]. \tag{5}$$

Note that $\mathcal{E}$ in equation (5) denotes a transformation between two levels, whereas $\mathcal{I}$ in equation (1) denotes a transformation between two intensities. It is also important to note that the present model incorporates any influence of spread of excitation into the excitation transformation $\mathcal{E}$. We define *compression* $\alpha(L_s$ as the reciprocal of the slope of $\mathcal{E}$ with respect to the external sound level.

$$\alpha(L_s) = \left(\frac{\mathrm{d}E_s}{\mathrm{d}L_s} \right)^{-1}. \tag{6}$$

Compression can strongly influence performance on psychoacoustic tasks and it plays an important role in any model of auditory perception.

The excitation transformation $\mathcal{E}$ describes auditory response growth. It is shaped to a large extent by peripheral auditory processing. Two examples of auditory response growth are shown in the upper panel of Figure 1. These estimates are based on the loudness growth measurements of Fletcher and Munson [10, dashed line] and Neely *et al.* [11, squares]. The solid lines in Figure 1 will be discussed below. The dashed line in the lower panel is the reciprocal of the numerical slope of the Fletcher and Munson loudness data (shown in the upper panel) and provides an empirical estimate of compression.

Neely *et al.* [11] used measurements of the number of dB required to quadruple loudness to derive estimates of loudness growth, which are represented by the squares in the upper panel of Figure 1. In this experiment, subjects listened to a single tone in one interval and the sum of

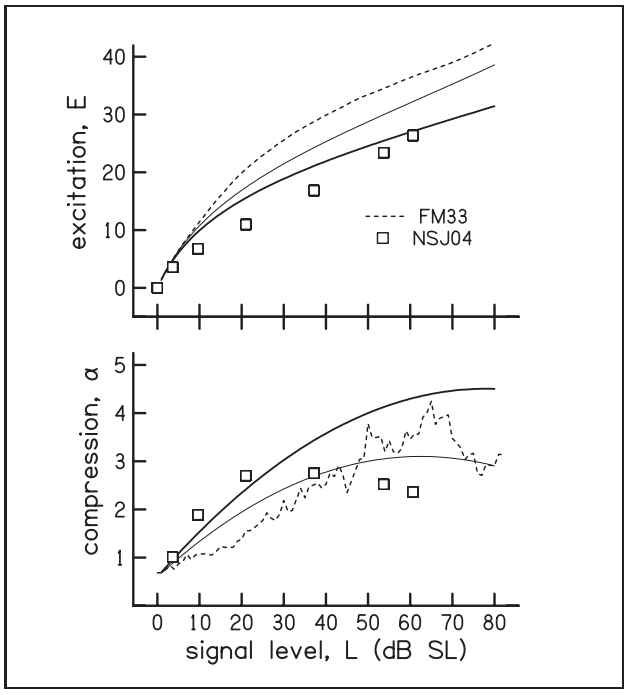

Figure 1. Excitation and compression. Response growth (or excitation, upper panel) and compression (lower panel) based on psychophysical data are compared with the quadratic compression used in the present model. Response growth (excitation) curves shown in the upper panel correspond to each of the compression curves in the lower panel. The dashed line represents $10\log_{10}$ of the loudness values of Fletcher and Munson [10]. The squares represent loudness growth data of Neely *et al.* [11]. The normal quadratic-compression curve (thick solid line, lower panel) was computed using equation (7) with $a = 0.6$, $b = 0.1$, and $c = -0.00064$. An alternate quadratic-compression curve (thin solid line, lower panel) was computed with $b = 0.075$ to represent compression that is less than normal. The corresponding model response growth functions (thin and thick solid lines, upper panel) were computed using equation (5) with the same model parameters.

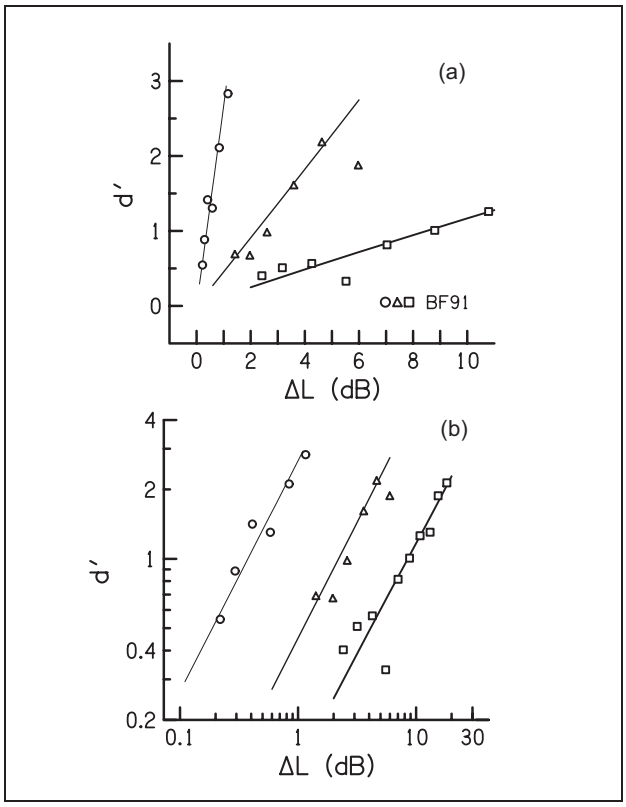

Figure 2. Psychometric functions for intensity discrimination. The left panel uses linear axes and the right panel uses logarithmic axes. Equation (12) was used to obtain d' values with the same compression parameters as in Figure 1, $a = 0.6$, $b = 0.1$, and $c = -0.00064$. The thin lines are for $L_1 = 85$ dB SL and $\delta_s = 1$. The medium-weight lines are for $L_1 = 85$ dB SL and $\delta_s = 0.17$. The thick lines are for $L_1 = 25$ dB SL and $\delta_s = 0.17$. In each case, the perceived-intensity variance coefficient was $n_2 = 0.82$. For comparison, the symbols represent data obtained by Buus and Florentine [14, Figure 1] from one subject (RR) for comparable stimulus conditions, assuming a threshold of 5 dB SPL. The circles are for 500-ms tones at 90 dB SPL, the triangles are for 10-ms tones at 90 dB SPL, and the squares are for 10-ms tones at 30 dB SPL.

four, equally-loud tones, at four different frequencies (1, 2, 4, and 8 kHz), in the other interval. They were asked to judge which of the two stimuli was louder. The *number of dB required to quadruple loudness* was derived from these relative-loudness judgments. These data were averaged across five subjects and then averaged across four single-tone frequencies (1, 2, 4, and 8 kHz). Compression was estimated as *the number of dB required to quadruple loudness* divided by 6. The division by 6 is appropriate because an ideal energy detector would see 6 dB as quadruple the energy. The squares in the lower panel of Figure 1 represent the average of the compression estimates shown in Figure 2 (lower panel) of Neely *et al.* [11]. The squares in the upper panel of Figure 1 represent a numerical integration of these compression values (The integrated values (upper panel) are approximately equal to $10\log_{10}$ of the loudness values in Figure 3 (lower panel) of Neely *et al.* [11], with an adjustment to the x-axis for the average threshold level (7 dB SPL) and an adjustment to the y-axis to make excitation zero at threshold).

The solid lines in the lower panel of Figure 1 are quadratic functions that approximate the empirical compression estimates.

$$\alpha(L_s) = a + bL_s + cL_s^2. \tag{7}$$

Here L is the sound level in dB relative to threshold. Parameter values for this quadratic function were determined primarily by fitting intensity discrimination data that will be described later. Consequently, the quadratic compression functions shown in Figure 1 do not provide the best possible fit to any of the empirical compression estimates; however, they do have some similar characteristics. Because it incorporates three parameters, quadratic compression provides a reasonable compromise between simplicity and flexibility when applied to modeling psychophysical performance.

From our quadratic representation of compression we can derive a corresponding transformation from external

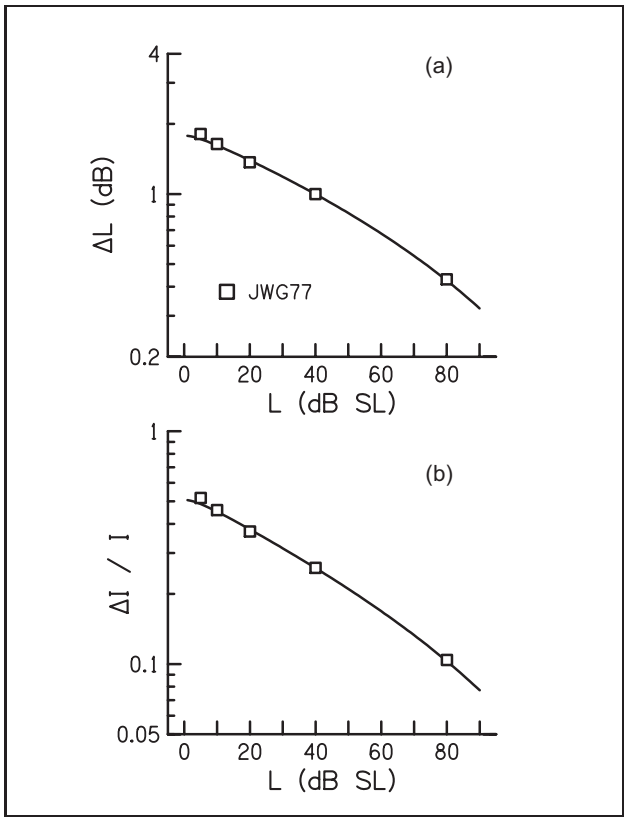

Figure 3. Intensity discrimination JND and Weber fraction. The solid lines were computed from equation (14) and equation (15) with $\delta_s = 1$. All other model parameters were the same as in Figure 2. The symbols represent intensity discrimination data of Jesteadt *et al.* [15] at 1 kHz.

level to internal level, by integrating the reciprocal of compression with respect to the external level. The result is an inverse hyperbolic tangent function of sound level.

$$E = \mathcal{E}(L_s) = -\frac{2}{q}\tanh^{-1}\left(\frac{2cL_s + b}{q}\right) + C. \quad (8)$$

Here $q = \sqrt{b^2 + 4ac}$ and C is an integration constant that we will arbitrarily select to make $\mathcal{E}(0) = 0$. This function is plotted in the upper panel of Figure 1 for two different values of the compression parameter b. Equation (8) provides a mathematically tractable representation of the peripheral nonlinearity that agrees in general form with our loudness growth data and other representations in the literature. A useful feature of this expression is that its inverse can also be simply expressed in terms of a hyperbolic tangent. This feature facilitates conversion between internal (perceptual) and external (physical) domains. The fact that our nonlinearity represents the total output of the cochlea, including spread of excitation, distinguishes it from measures of peripheral nonlinearity at a single point, such as observed in direct basilar membrane measurements. However, these two nonlinearities have many similar features because the integrated output of the cochlea is dominated by response at a single point whenever the stimulus is a low-level tone.

We assume cochlear filtering precedes cochlear nonlinearity in order to avoid the difficulty of modeling nonlinear filters. When referring to external sounds, we sometimes use the word signals instead of stimuli or sounds, to distinguish between signal and masker components. The signal is the part of the sound that carries the information required for detection or discrimination. In a given frequency region, we assume that the whole sound undergoes the same compression as it is processed by the cochlea.

3. Model results

The modeling principles described above may be applied to several, simple psychoacoustic tasks. The modeling approach that we adopt here is to use basic assumptions to derive an equation for d' that describes task performance, then solve for the perceived intensity at threshold by setting $d' = 1$. Finally, we assess the validity of our d' and threshold equations by comparing model results with empirical data.

Before we consider specific tasks, we will describe a general expression for d' that may be applied to all of the tasks considered in this paper. The perceived-intensity difference ΔN_s in the numerator of equation (2) is, in general, a linear combination of two perceived intensities that may be present in the stimulus.

$$\Delta N_s = \nu_2 \mathcal{I}(I_2) - \nu_1 \mathcal{I}(I_1). \quad (9)$$

In most cases, ΔN_s is simply the difference between two perceived tone intensities and we let $\nu_2 = \nu_1 = 1$. The standard deviation σ_s in the denominator of equation (2) is the square-root of the variance, which is the sum of several possible contributions.

$$\sigma_s^2 = \eta_0^2 + \eta_1^2 \mathcal{I}(I_1) + \eta_2^2 \mathcal{I}(I_2) + \eta_{en}^2 \mathcal{I}(I_{en}). \quad (10)$$

The first term in equation (10) is independent of any stimulus component and can be attributed to internal noise added by the central detector. The second and third terms are proportional to the perceived intensities of the stimulus tones. These intensity-dependent contributions to the variance might be due to variability in the neural-spike generation process. Modeling the intensity-dependent contribution to variance as being proportional to the perceived intensity was first suggested by McGill and Goldberg [12] and Siebert [13]. This appropriateness of this choice will be seen when we compare results of our d' expression to representative data.

The fourth term in equation (10) is proportional to the square of the perceived intensity of external noise. Modeling the noise contribution to the total variance as being proportional to the square of the perceived noise intensity may be consistent with it adding to the variability of total perceived intensity at the central detector, in contrast to being associated with neural-spike generation. However, this formulation was an empirical choice, based on comparing model behavior with experimental data. The fourth term is zero when external noise is not present.

Our general expression for d' is

$$d' = \frac{\nu_2 \mathcal{I}(I_2) - \nu_1 \mathcal{I}(I_1)}{\sqrt{\eta_0^2 + \eta_1^2 \mathcal{I}(I_1) + \eta_2^2 \mathcal{I}(I_2) + \eta_{en}^2 \mathcal{I}(I_{en})}}. \quad (11)$$

This expression will be simplified for each specific task considered below.

3.1. Intensity discrimination

Consider a two-interval, two-alternative forced-choice (2AFC) task in which the listener is asked which of two tones is louder and the tones differ only in intensity. We assume that psychoacoustic performance for this task is governed by the difference between the two perceived intensities. The two tones contribute equally to this task, so we let $\nu_2 = \nu_1 = 1$. We assume that variance is dominated by the perceived intensity of the more intense of the two tones I_2, so we set $\eta_2 = n_2/\delta_s$. This variance coefficient includes a duration dependence δ_s that will be assigned a value between 0 and 1 to increase the variance when tones are short. We will not consider the presence of external noise in this task, so $\eta_0 = \eta_1 = \eta_{en} = 0$.

With these assumptions, performance on the *intensity discrimination* task is described by the following equation.

$$d' = \frac{\mathcal{I}(I_2) - \mathcal{I}(I_1)}{(n_2/\delta_s)\sqrt{\mathcal{I}(I_2)}}. \quad (12)$$

We have determined a value for n_2 by fitting experimental intensity discrimination data from the literature. In Figure 2, we plot d' as a function of $\Delta L = 10\log(I_2/I_1)$. When I_1 is fixed and I_2 is varied, we see (in the lower panel of Figure 2) that the log of d' is approximately a linear function of the log of ΔL, with a slope of about 1. Changes in the value of n_2 shift the curves (in the right panel) vertically. Specifically, an increase in n_2 shifts the curves downward. Changes in the value of I_1 shift the curves (in the right panel) horizontally. The thick and medium-weight lines in Figure 2 represent tone levels of L_1 = 25 and 85 dB SL, respectively, for the same tone duration ($\delta_s = 0.17$). It is interesting to note that these d' curves maintain their linear shape regardless of whether they are plotted on linear axes (left panel) or logarithmic axes (right panel). The model results provide a good fit to intensity discrimination data of Buus and Florentine [14].

Our next step is to derive an equation for I_2 as a function of I_1, in order to describe threshold intensity for this task. We begin by redefining σ_s in terms of the perceived intensity of the first (lower level) stimulus instead of the second (higher level) stimulus by letting $\eta_1 = n_2/\delta_s$ and $\eta_2 = 0$ to obtain $\sigma_s = (n_2/\delta_s)\sqrt{\mathcal{I}(I_2)}$. The reason for this redefinition is to simplify our expression for I_2. When I_1 and I_2 are similar intensities, as they generally are when $d' = 1$, this redefinition has little effect on the value of d'. However, when I_2 is much larger than I_1, as it is for some conditions shown in Figure 2, it is better to use our initial definition of σ_s.

We set $d' = 1$ in equation (12) and solve for I_2. The following expression describes I_2 as a function of I_1 when $d' = 1$.

$$I_2 = \mathcal{I}^{-1}\left[\mathcal{I}(I_1) + \frac{n_2}{\delta_s}\sqrt{\mathcal{I}(I_1)}\right]. \quad (13)$$

Here $\mathcal{I}^{-1}$ is the inverse of the intensity transformation $\mathcal{I}$. Equation (13) describes the value of I_2 that is *just noticeably different* from the less intense stimulus I_1.

We can use equation (13) to derive an expression for the just-noticeable-difference (JND) in terms of stimulus level ΔL.

$$\begin{aligned}\Delta L &= \mathcal{L}\left(\frac{I_2}{I_1}\right) \\ &= \mathcal{L}\left[\mathcal{I}^{-1}\left(\mathcal{I}(I_1) + \frac{n_2}{\delta_s}\sqrt{\mathcal{I}(I_1)}\right)\right].\end{aligned} \quad (14)$$

Recall that $\mathcal{L}$ in this equation represents a log transformation from intensity to level. Results for the intensity JND from equation (14) are shown in the upper panel of Figure 3. Model results are in good agreement with the intensity discrimination data of Jesteadt *et al.* [15]. This agreement supports our choice to represent the perceived-intensity variance as being proportional to the perceived intensity. In the process of fitting the intensity discrimination data we determined not only the variance parameter n_2, but also the three compression parameters a, b, and c.

The Weber fraction is an alternative metric for describing intensity discrimination performance.

$$\begin{aligned}\frac{\Delta I}{I} &= \frac{I_2 - I_1}{I_1} \\ &= \mathcal{I}^{-1}\left[\mathcal{I}(I_1) + \frac{n_2}{\delta_s}\sqrt{\mathcal{I}(I_1)}\right] I_1^{-1} - 1.\end{aligned} \quad (15)$$

The lower panel of Figure 3 compares the same model results with intensity discrimination data in terms of the Weber fraction. When plotted on log axes, the Weber fraction looks remarkably similar to ΔL.

3.2. Single-tone detection

Consider the task in which a listener is asked detect the presence of a short tone of duration τ_s and intensity I_s. Our focus here will be to explore the influence of duration on the threshold of tone detection. We have no second tone or external noise in this task, so we set $\nu_2 = \eta_2 = \eta_{en} = 0$. To represent the single tone appropriately, we set $\nu_1 = -1$ and $\eta_1 = n_2/\delta_s$. To describe the duration dependence in more detail, we let $\delta_s^2 = 1 - a_i e^{-\tau_s/\tau_i}$, where τ_s is the signal duration and τ_i is a temporal integration time constant. The functional form of δ_s represents a possible integration of the perceived intensity that effectively decreases the perceived-intensity variance as the signal duration increases [16]; however, this functional form was selected for use in the current model primarily because it produces model results that fit the data.

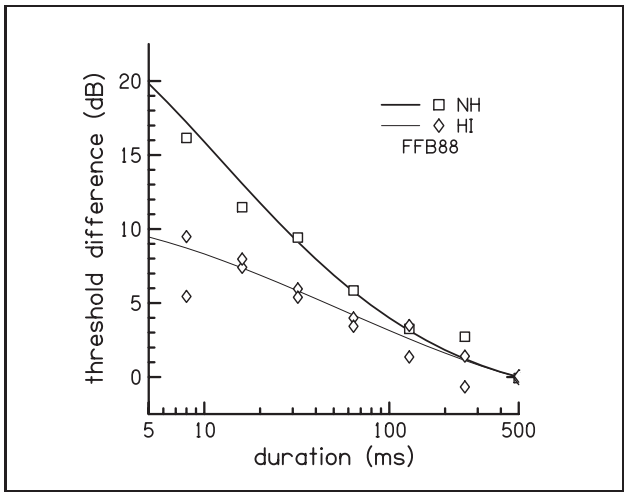

Figure 4. Temporal integration demonstrated by threshold difference (dB re 500 ms) as a function of single-tone duration. The thick line represents normal-hearing (NH) performance and was computed from equation (17) with temporal integration parameters $a_i = 0.99$, $\tau_i = 450$ ms. The compression and variance parameters were the same as in Figure 2. Duration on the abscissa is equivalent to τ_s in the model. The thin line represents hearing-impaired (HI) performance and were computed using the same parameters except $b = 0.01$. This parameter change reduces the amount of compression at moderate and high levels. The symbols show data from [17, Figure 3]. The diamonds represent two individual hearing-impaired listeners. The squares represent the average of five normal-hearing listeners.

With these assumptions, performance on the *single-tone detection* task is described by the following equation.

$$d' = \frac{\mathcal{I}(I-1)}{n_2\sqrt{\mathcal{I}(I_1)/(1-a_i\mathrm{e}^{-\tau_s/\tau_i})}}. \tag{16}$$

Setting $d' = 1$ in equation (16) and solving for I_1 yields an equation that describes the dependence of threshold on the duration of the tone.

$$I_1 = \mathcal{I}^{-1}\left[\frac{n_2^2}{1-a_i\mathrm{e}^{-\tau_s/\tau_i}}\right]. \tag{17}$$

Results of equation (17) are shown in Figure 4 for normal-hearing (NH) and hearing-impaired (HI) listeners. In this example, hearing impairment is simulated by changing one of the compression parameters to reduce the amount of compression. The trends in the model results are similar to the temporal-integration data of Florentine *et al.* [17]. Note that no change in the temporal integration time constant is required to account for the differences in the data between the two categories of listeners. The present model is similar in this respect to the model proposed by Oxenham *et al.* [18]; however, the two models differ in the slope of the temporal integration curve. In the present model (Figure 4), for signal durations of 10 ms and less, the slope is much larger for the NH listeners than for HI listeners. In the model described by Oxenham *et al.* (their Figure 4), slopes are more similar between these two groups.

3.3. Increment detection

The decision process for increment detection is generally thought to be based on detection of a change in level of the stimulus within a single interval rather than a comparison of the overall levels of the stimuli presented in two separate intervals [19]. For the present purposes, we will ignore details of the decision process and simply assume that internal noise differs between increment detection and intensity discrimination.

To arrive at a model of increment detection, let I_1 be the pedestal intensity and I_2 be the incremented intensity of duration τ_s. This means that the intensity increment is $I_2 - I_1$. The two perceived-intensity means contribute equally to this task, so $\nu_2 = \nu_1 = 1$. We include the influence of detector noise and the variance contribution due to the pedestal, so $\eta_0 = n_0/\delta_s$, $\eta_1 = n_1/\delta_1$, and $\eta_2 = 0$. The influence of external noise is also included by letting $\eta_{en} = n_3/\delta_s$. We include the effect of varying increment duration, so $\delta_s^2 = 1 - a_i\mathrm{e}^{-\tau_s/\tau_i}$.

With these assumptions performance on the *increment detection* task is governed by the following equation.

$$d' = \frac{\mathcal{I}(I_2) - \mathcal{I}(I_1)}{\dfrac{\sqrt{n_0^2 + n_1^2\mathcal{I}(I_1) + n_3^2\mathcal{I}^2(I_{en})}}{\left(1-a_i\mathrm{e}^{-\tau_s/\tau_i}\right)}}. \tag{18}$$

Equation (18) is very similar in form to equation (12), which describes performance for the intensity discrimination task. The only feature in our model that distinguishes *increment detection* from *intensity discrimination* is a reduced influence from the intensity-dependent contribution to the variance. Specifically, parameter n_1 for increment detection in the denominator of equation (18) is less than one-third the value of parameter n_2 for intensity discrimination in the denominator of equation (12). This reduction agrees with our expectation that the cross-interval comparison required for intensity discrimination introduces more variability than the within-interval comparison required for increment detection.

We derive the signal threshold as a function of duration and external noise level by setting $d' = 1$ in equation (18) and solving for I_2.

$$I_2 = \mathcal{I}^{-1}\left(\mathcal{I}(I_1) + \sqrt{\frac{n_0^2 + n_1^2\mathcal{I}(I_1) + n_3^2\mathcal{I}^2(I_{en})}{1-a_i\mathrm{e}^{-\tau_s/\tau_i}}}\right). \tag{19}$$

In order to compare model results with data, the signal level at threshold can be expressed in terms of the stimulus-level difference ΔL.

$$\Delta L = \mathcal{L}\left(\frac{I_2}{I_1}\right) = \mathcal{L}\left[\mathcal{I}^{-1}\left(\mathcal{I}(I_1) + \sqrt{\frac{n_0^2 + n_1^2\mathcal{I}(I_1) + n_3^2\mathcal{I}^2(I_{en})}{1-a_i\mathrm{e}^{-\tau_s/\tau_i}}}\right)\right] - \mathcal{L}(I_1). \tag{20}$$

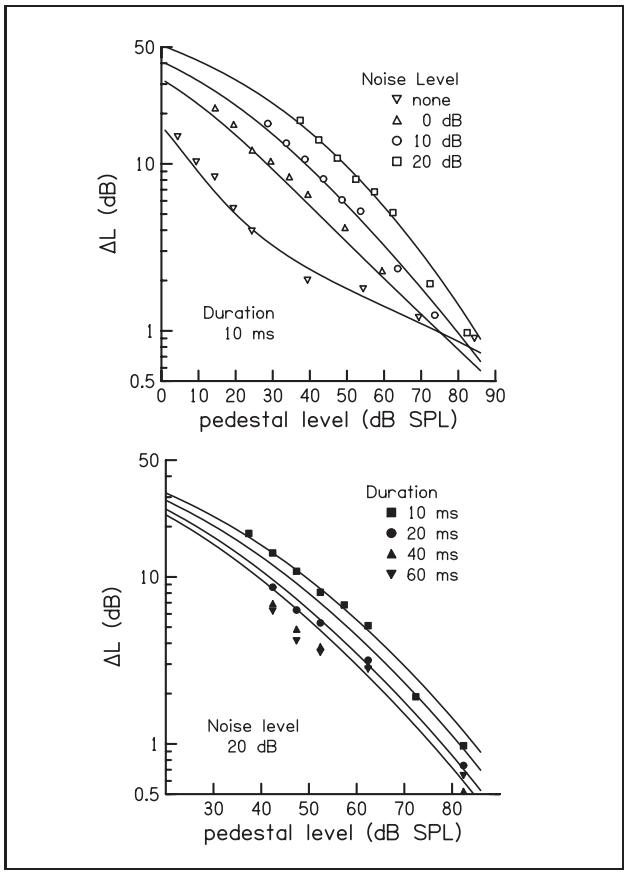

Figure 5. Increment detection in noise. Model results with no noise and with two levels of external noise (solid lines) are compared with data from our laboratory at 4 kHz [38]. The temporal-integration parameters were the same as in Figure 4. For the case without noise, the compression parameters were the same as in Figure 2, with $b = 0.1$. Compression was reduced for the three external noise conditions by setting b = 0.09, 0.085, and 0.08 for L_n = 0, 20, and 40 dB, respectively. Model thresholds for the external noise and pedestal levels were −42 dB and 6 dB SPL, respectively. Other model parameters were $n_0 = 2$, $n_1 = 0.25$, and $n_3 = 0.05$.

Results of equation (20) are shown in the upper panel of Figure 5 as functions of pedestal level for increment detection with three levels of external noise and for increment detection with no external noise. The right panel shows the effect on ΔL of varying the increment duration τ_s. We see good agreement between the model results and the experimental data in the left panel with model results reproducing most of the trends observed in the experimental data. Agreement is not as good in the right panel, because the model results are not influenced as much by duration as the experimental data.

In Figure 6, results of equation (20) are compared with increment detection data of Glasberg *et al.* [3] as a function of increment duration. The total decrease in ΔL over this range of τ_s is less for the data than for the model. This is, perhaps, an indication that the variance due to external noise should not be subjected to the same temporal integration as the variance due to internal noise. Instead of further complicating the model to explore better fits to these data, we leave this discrepancy as an apparent weakness of the current model implementation.

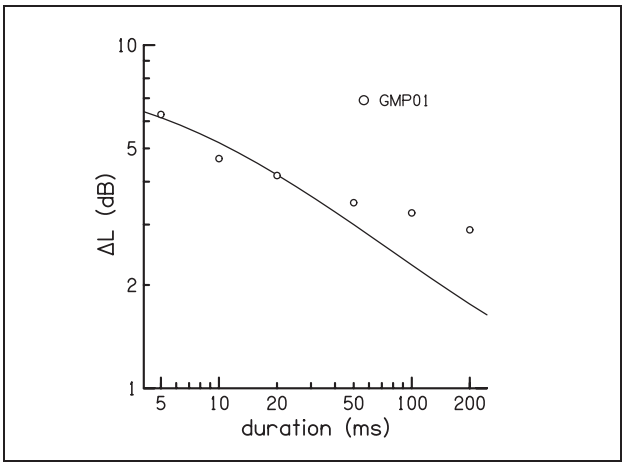

Figure 6. Increment-detection JND as a function of increment duration. The model results (solid line) were computed from equation (20) with $L_p = 66$ and $L_n = 73$ dB re threshold. All other model parameters were the same as in Figure 5, with $b = 0.08$. The circles represent increment detection data from [3] for comparable stimulus conditions at 1 kHz.

3.4. Forward masking

We use a similar approach to arrive at a model of forward masking, which can be viewed as a special case of increment detection in which the increment follows the pedestal. The pedestal is now called the masker. Let I_1 be the intensity of the masker and I_2 be the intensity of the signal, which follows the masker with a delay of τ_d. Signal delay is defined as the time between the trailing edge of the masker offset ramp and the leading edge of the signal onset ramp. Unlike increment detection, I_2 in forward masking is usually less than I_1. In preliminary versions of the forward-masking model, we included the influence of masker persistence by letting $\nu_1 = a_p \, \mathrm{e}^{-\tau_d/\tau_p}$; however, in the process of optimizing model parameters, it was discovered that the best results were obtained when the influence of persistence was negligible. In the interest of simplicity, we have dropped masker persistence in the present version of the model by setting $\nu_1 = 0$. Masker influence is presently modeled only as neural adaptation by specifying $\nu_2 = 1/(1 + \mathcal{I}(I_1) a_a \, \mathrm{e}^{-\tau_d/\tau_a}$, where τ_a is a time constant associated with this adaptation. This specification decreases the perceived intensity of the signal as signal delay decreases. For simplicity, we will not consider external noise and will not include any intensity- dependent contributions to the variance; however, we retain the same dependence on signal duration. In this case, variance will simply be represented by a constant $\eta_0 = n_{fm}/\delta_s$, with $\eta_1 = \eta_2 = \eta_{en} = 0$.

With these assumptions performance on the *forward-masked detection* task is governed by the following equation.

$$d' = \frac{\mathcal{I}(I_2)/\left(1 + \mathcal{I}(I_1) a_a \, \mathrm{e}^{-\tau_d/\tau_a}\right)}{n_{fm}/\delta_s}. \qquad (21)$$

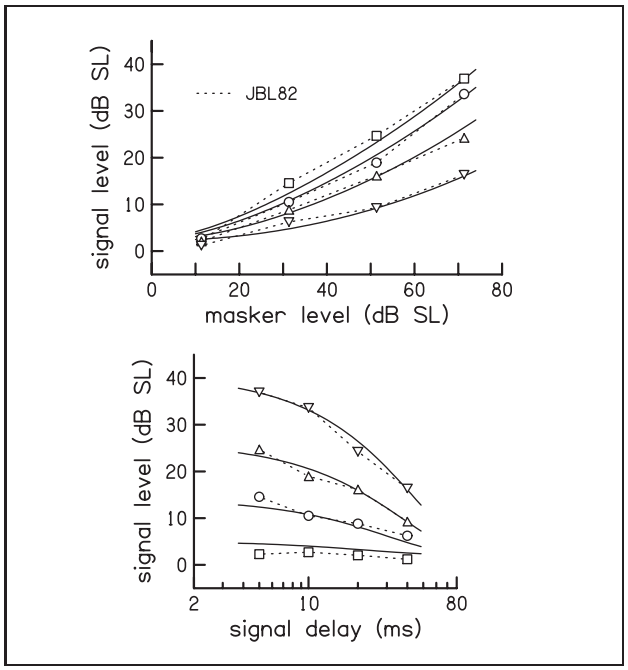

Figure 7. Signal thresholds for forward masking as a function of masker level (top) and signal delay (bottom). The solid lines were computed from equation (22) with $n_{fm} = 0.4$, $a_a = 0.09$, $\tau_a = 19$ ms, and $\delta_s = 0.2$. The compression parameters were the same as in Figure 2. In the left panel, signal delay was varied by setting τ_s = 5, 10, 20 and 40 ms. In the right panel, masker level was varied by setting L_m = 20, 40, 60 and 80 dB re threshold. The dashed lines are linear fits to forward-masking data of [20] at 1 kHz for equivalent stimulus parameters.

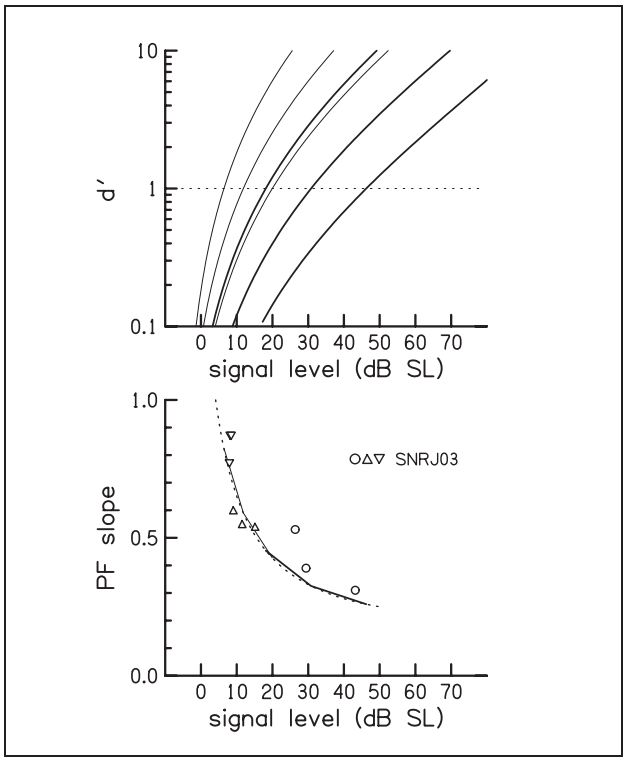

Figure 8. Psychometric functions (PF) for forward masking as a function of signal level. The value of d' (top) was computed from equation (21) using the same model parameters as in Figure 2. The thick and thin solid lines represent signal delays τ_s = 2 and 40 ms, respectively. At each signal delay, three different values of masker level are shown, L_m = 45, 65, and 85 dB SL. PF slope (bottom) is defined for the model as the derivative of $10\log_{10}(d')$ with respect to signal level, evaluated at $d' = 1$. The symbols (bottom) represent comparable forward-masking, PF-slope data from [21]. The dashed line is the reciprocal of the quadratic compression function.

The use of identical transformations for the signal and masker assumes that they are both at the same frequency and, therefore, undergo the same compression. To model off-frequency masking would require an additional set of compression parameters, because response-growth rate may differ at other frequencies.

We see how the signal level at threshold varies with masker level and signal delay by setting $d' = 1$ in equation (21) and solving for I_2.

$$I_2 = \mathcal{I}^{-1}\left[\frac{n_{fm}}{\delta_s}\left(1 + \mathcal{I}(I_1)a_a\,\mathrm{e}^{-\tau_d/\tau_a}\right)\right], \tag{22}$$

$$L_2 = \mathcal{L}\left[\mathcal{I}^{-1}\left(\frac{n_{fm}}{\delta_s}\left(1 + \mathcal{I}(I_1)a_a\,\mathrm{e}^{-\tau_d/\tau_a}\right)\right)\right]. \tag{23}$$

Signal thresholds are shown in Figure 7 as a function of masker level for selected values of signal delay. Parameter values were chosen to yield results similar to the forward masking measurements of Jesteadt *et al.* [20], which are represented by symbols in this figure. Good agreement was achieved, despite using the same compression parameters here that were used for all of the previous tasks.

Psychometric functions (PF) for the forward-masking task, defined as d' versus signal level, are shown in the left panel of Figure 8. These model results used the same parameter values as in Figure 7. Note that PF slope decreases with increasing signal level. The model PF slope is shown in the right panel and is in good agreement with the comparable PF slopes for forward masking reported by Schairer *et al.* [21].

4. Discussion

4.1. Model derivation

The present model is similar in approach and formulation to the work of Durlach and Braida [1, 22]. Their general formula for d', which they call sensitivity, has a difference between transformed intensities in the numerator and a standard deviation in the denominator, similar to our equation (11). Their intensity transformation differed in being simply a log function, in contrast to the concatenation of log, inverse hyperbolic tangent, and anti-log transformation that we use. Their assumed contributions to the standard deviation differed from ours both in dependence on stimulus intensity and rationale used to justify these contributions, but were similar in assuming multiple sources of internal noise.

The intensity transformation in our model builds on previous work relating intensity discrimination performance

to the rate of growth of loudness (e.g., [23, 19]). Our representation of compression by a quadratic function was motivated by the loudness growth data of Neely *et al.* [11]. Loudness data have often been used to estimate compression [24, 25]. Other methods to estimate compression include measurements of forward masking [26, 27, 6], otoacoustic emissions [28], and basilar membrane displacement [8, 29]. Other suggested compression functions are generally similar in form to the quadratic function used here, but tend to be less tractable, and, therefore, less conducive to the derivation of explicit equations for d' and signal threshold.

Note that when the compression parameter c is zero, the excitation transformation $\mathcal{E}$ becomes logarithmic. In this case, the anti-log transformation from perceptual level E_s to perceived intensity N_s nearly straightens out the nonlinearity in the transformation from L_s to E_s. When this happens, N_s is approximately proportional to L_s. In other words, the transformation from physical intensity to perceived intensity is nearly a log transformation, so perceived *intensity* can be nearly proportional to external *level*. This observation may explain the apparent correspondence that is sometimes observed between additive internal noise, which presumably adds to perceived *intensity*, and decibel variations in the external sound *level*.

4.2. Intensity discrimination

The observation that the Weber fraction decreases with level for intensity discrimination of tones is called the "near miss" to Webers law [12]. It is generally acknowledged that the near miss is due to spread of excitation within the cochlea as the tone level increases. The present model is consistent with this interpretation; however, the influence of spread of excitation is implicitly incorporated in the excitation transformation $\mathcal{E}$.

It is frequently assumed that compression in the cochlea starts at about 40 dB SL. Compression in our model differs from this assumption by starting at threshold and increasing gradually. A possible explanation for this apparent discrepancy is that compression at low levels in our model may reflect spread of excitation, which is incorporated differently in other models of auditory function.

The representation of perceived-intensity variance as being proportional to perceived-intensity mean for the intensity-discrimination task is consistent with the source of internal variability being due to generation of neural spikes, which is thought to be a Poisson process (e.g., [12, 30, 13]). However, the main reason for selecting this functional form for the perceived-intensity variance is that it produces model results that are in better agreement with the Jesteadt *et al.* [15] intensity discrimination data than could be obtained by selecting either (1) variance that is independent of perceived intensity or (2) variance that is proportional to the square of perceived intensity.

Other alternatives were also considered. Better agreement with the data at low levels might have been achieved by adding a constant term to the equation for perceived-intensity variance and better agreement at high levels might be achieved by adding a term that is proportional to the square of the perceived intensity. However, comparison with the Jesteadt *et al.* [15] data suggests that linear dependence of perceived-intensity variance on perceived intensity dominates over almost the entire range of signal levels.

4.3. Temporal integration

One interpretation of temporal integration is that the central estimate of the mean perceived intensity N_s is obtained by integrating a randomly fluctuating version of this intensity, so that the variance of this estimate decreases as the duration of the sound increases. The functional form that we used to represent temporal integration in our model is similar to the one suggested by Plomp and Bouman [16]. Although we apply the integration here to the variance instead of to the signal, the effect on the model is similar.

Viemeister and Wakefield [31], in order to explain experimental results that appeared to be inconsistent with temporal integration, proposed a "multiple looks" model that is consistent with the Plomp and Bowman formulation for signals that have a contiguous duration. As a hypothetical implementation of their "multiple looks" model, Viemeister and Wakefield suggested that short-term memory could allow the central detector to combine d' estimates across (possibly discontiguous) portions of a signal. The important role of perceived-intensity variance in the present model suggests that an alternative model for "multiple looks" might be to estimate intensity mean and variance separately across discontiguous portions of a signal and compare them only as a final step in the decision process.

In Figure 4, the influence of hearing impairment on temporal integration was simulated by changing the value of one of the compression parameters. This is consistent with the current view that loss of compression is the primary factor underlying many differences between listeners with normal hearing and those with cochlear hearing loss, including loss of frequency selectivity (e.g., [4, 32]). The physiological evidence supporting loss of compression associated with hearing loss is clear; whereas broadening of tuning curves is not always observed (e.g., [33]). Launer *et al.* [39] (1997) have argued that loss of compression is an adequate model for the influence of hearing loss on loudness summation across frequency. Broadened tuning has no meaning in our model because it has no filters. In any case, reduction in compression appears to be an appropriate representation of hearing impairment.

Note in Figure 4 that the slope of the threshold difference curve is less negative in the hearing impaired (HI) condition, suggesting an increase in the temporal integration time constant. This is an illusion created by expressing threshold in external (physical) units (e.g., [4]), because the time constant associated with temporal integration in the model was not changed.

It might seem natural to explore the effect of external noise on the model for single-tone detection and compare model results with data (e.g., [17]); however, the introduction of external noise into equation (16) makes the corre-

sponding version of equation (17) considerably more complicated, perhaps requiring iterative methods to find a solution. For this reason, we decided to limit our consideration of external noise to the case of increment detection.

4.4. Forward masking

As noted above, the influence of masker persistence on threshold signal level for the forward-masking task was explored in other versions of the model. Masker persistence (MP) was represented by a term that was proportional to the perceived masker intensity $\mathcal{I}(I_1)$ and was subtracted from the perceived intensity $\mathcal{I}(I_2)$ in the numerator of the equation for d'. This representation of masker persistence can, by itself, produce model results for signal threshold that are identical to those obtained by using neural adaptation (NA) alone, such as shown in Figure 7. However, the corresponding model results for d', such as shown in the left panel of Figure 8, are much too steep when MP is the only masking mechanism used in the model. Furthermore, when MP was combined with NA, the best fit to the data was achieved by giving MP only one-tenth of the influence of NA. Based on these results, it was decided, for simplicity, to exclude MP from the present model because NA provided a much better fit to the forward masking PF data. NA was suggested as the mechanism of forward masking by Duifhuis [34].

Some of the model curves (solid lines) in the left panel of Figure 7 appear to have too much curvature compared to the Jesteadt *et al.* data (dashed lines). Such upward curvature is also produced by the model of forward masking proposed by Plack and Oxenham [6] and can be seen in other sets of forward masking data (e.g., [35, 21]). Further exploration involving comparisons with other forward masking data could determine whether the degree of curvature observed in the present model results is appropriate.

Note the overlap between the two solid lines in the right panel of Figure 8. The thick line represents model results with signal delay of 40 ms and the thin line with 2 ms. Apparently, PF slope (at $d' = 1$) is independent of signal delay (or masker level) for fixed signal level. This provides another prediction of our forward-masking model and supports an assumption made by Schairer *et al.* [21].

Because it only depends on signal level, the shape of the PF slope curve should be closely related to the amount of compression. This prediction is confirmed by comparison of the solid lines in Figure 8 with the dashed line, which represents the reciprocal of the quadratic compression function that was used to model the forward-masking task. In other words, when $d' = 1$, PF slope provides a very good approximation to the reciprocal of compression. This relationship should allow compression to be estimated empirically as the reciprocal of the forward-masking PF slope.

Although forward masking and increment detection are typically implemented experimentally as two-interval tasks, it is interesting to note that recent evidence [19] indicates these tasks do not require cross-interval comparisons. Unlike intensity discrimination, forward masking and increment detection can be performed without comparing overall level across intervals. This difference may help to explain differences in our model across these tasks in the amount of intensity-dependent variance.

4.5. Time-domain model

The model presented here operates on parametric representations of acoustic stimuli. It should be possible to develop an equivalent model that would operate on stimulus waveforms and produce similar results. As part of a multistage, time-domain auditory model, our new representation of the peripheral nonlinearity could offer advantages over other representations in the literature because it introduces an appropriate amount of compression over a wide range of signal levels. We feel that basing the compressive nonlinearity on psychophysical data, as we have done in this paper, is preferable to basing it on basilar membrane measurements. Implementation of this compression in the time-domain should be straightforward because the reciprocal of the quadratic representation of compression can be used as an exponent in a "power-law" compression stage of the model. Another attractive feature of this model for time- domain implementation is that it provides explicit equations for the mean and variance of the perceived intensity. This feature makes it possible to avoid Monte-Carlo simulations and still obtain distributions of random variables. Decision variables in a time-domain model would be time-varying versions of the equations for d' that are provided in this paper.

It may be difficult to extend our modeling approach to the sounds that stimulate many auditory filters. One approach would be to implement a filter bank with a separate compressive nonlinearity at the output of each filter; however, it seems unlikely that this approach could ever represent interactions among frequency components in a completely realistic manner. It may be preferable to use physically-base models of cochlear mechanics (e.g., [36, 37]) to gain a deeper understanding of how complex stimuli are processed by the auditory system.

5. Conclusions

The quadratic-compression modeling framework appears to offer a useful tool for understanding auditory detection and discrimination. An advantageous feature of our model is that it provides a formulation in which internal and external noises are combined appropriately. A possible disadvantage of our model is that we are unable to provide rules for applying it to new stimulus conditions. Such rules may become apparent as we gain more experience with fitting the model to a wider range of experimental data. However, the complexity of the underlying physics and physiology may make it difficult to generalize our model.

Model results can lead to new insights regarding auditory perception. For example, the dependence of perceived-intensity variance on perceived intensity appears to be a distinguishing feature between the decision-making

processes involved in intensity discrimination and increment detection tasks. Also, the influence of masker persistence on forward masking appears to be insignificant compared to neural adaptation.

The mathematical framework provided here can be applied to new sets of data and modified as necessary. The assumptions underlying the model are explicit and more readily testable than the assumptions underlying other recent models that require computer code for their specification. More work needs to be done to explore the limitations of this modeling approach and discover possible extensions that might enable it to handle a wider range of psychoacoustic tasks.

Acknowledgement

Partially supported by grant R01-006648 from the National Institutes of Health, NIDCD. We thank Hongyang Tan for her help with figures and Kim Schairer for sharing her data. We thank Armin Kohlrausch for his many helpful suggestions. Portions of this study were presented at the Spring 2004 meeting of the Acoustical Society, in New York and at Summerschool 2004 sponsored by Oldenburg University in Bad Zwischenahn, Germany.

References

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

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

[3] B. R. Glasberg, B. C. J. Moore, R. W. Peters: The influence of external and internal noise on the detection of increments and decrements in the level of sinusoids. Hear. Res. **155** (2001) 41–53.

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

[5] B. C. Moore, R. W. Peters, B. R. Glasberg: Effects of frequency and duration on psychometric functions for detection of increments and decrements in sinusoids in noise. J. Acoust. Soc. Am. **106** (1999) 3539–3552.

[6] C. J. Plack, A. J. Oxenham: Basilar-membrane nonlinearity and the growth of forward masking. J. Acoust. Soc. Am. **103** (1998) 1598–608.

[7] D. M. Green, J. A. Swets: Signal detection theory and psychophysics. Wiley, New York, 1966.

[8] G. K. Yates, I. M. Winter, D. Robertson: Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range. Hear. Res. **45** (1990) 203–219.

[9] F. G. Zeng, R. V. Shannon: Loudness balance between electric and acoustic stimulation. Hear. Res. **60** (1992) 231–235.

[10] H. Fletcher, W. Munson: Loudness, its definition, measurement, and calculation. J. Acoust. Soc. Am. **5** (1933) 82–108.

[11] S. T. Neely, K. S. Schairer, W. Jesteadt: Cochlear compression estimates from loudness growth data. – In: Auditory Signal Processing: Physiology, Psychophysics, and Models. D. Pressnitzer, A. de Cheveigne, S. McAdams, L. Collet (eds.). Springer-Verlag, New York, 2005, 42–56.

[12] W. J. McGill, J. P. Goldberg: Pure-tone intensity discrimination and energy detection. J. Acoust. Soc. Am. **44** (1968) 576–581.

[13] W. M. Siebert: Stimulus transformations in the peripheral auditory system. – In: Recognizing Patterns. P. Kolers, M. Eden (eds.). MIT Press, Cambridge, Mass., 1968.

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

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

[16] R. Plomp, M. A. Bouman: Relation between hearing threshold and duration for tone pulses. J. Acoust. Soc. Am. **31** (1959) 749–758.

[17] M. Florentine, H. Fastl, S. Buus: Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking. J. Acoust. Soc. Am. **84** (1988) 195–203.

[18] A. J. Oxenham, B. C. Moore, D. A. Vickers: Short-term temporal integration: evidence for the influence of peripheral compression. J. Acoust. Soc. Am. **101** (1997) 3676–3687.

[19] W. Jesteadt, K. S. Schairer, D. L. Neff: Effect of variability in level on forward masking and on increment detection. J. Acoust. Soc. Am. **118** (2005) 325–337.

[20] W. Jesteadt, S. P. Bacon, J. R. Lehman: Forward masking as a function of frequency, masker level, and signal delay. J. Acoust. Soc. Am. **71** (1982) 950–962.

[21] K. S. Schairer, L. Nizami, J. F. Reimer, W. Jesteadt: Effects of peripheral nonlinearity on psychometric functions for forward-masked tones. J. Acoust. Soc. Am. **113** (2003) 1560–1573.

[22] L. D. Braida, N. I. Durlach: Intensity perception. II. Resolution in one-interval paradigms. J. Acoust. Soc. Am. **51** (1972) 483–502.

[23] J. B. Allen, S. T. Neely: The relation between the intensity JND and loudness for pure tones and wide-band noise. J. Acoust. Soc. Am. **102** (1997) 3628–3646.

[24] B. C. Moore, B. R. Glasberg: A revised model of loudness perception applied to cochlear hearing loss. Hear. Res. **188** (2004) 70–88.

[25] R. S. Schlauch, J. J. DiGiovanni, D. T. Ries: Basilar membrane nonlinearity and loudness. J. Acoust. Soc. Am. **103** (1998) 2010–2020.

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

[27] A. J. Oxenham, C. J. Plack: Effects of masker frequency and duration in forward masking: further evidence for the influence of peripheral nonlinearity. Hear. Res. **150** (2000) 258–266.

[28] S. T. Neely, M. P. Gorga, P. A. Dorn: Cochlear compression estimates from measurements of distortion-product otoacoustic emissions. J. Acoust. Soc. Am. **114** (2003) 1499–1507.

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

[30] R. D. Luce, D. M. Green: A neural timing theory for response times and the psychophysics of intensity. Psychol. Rev. **79** (1972) 14–57.

[31] N. F. Viemeister, G. H. Wakefield: Temporal integration and multiple looks. J. Acoust. Soc. Am. **90** (1991) 858–865.

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

[33] M. P. Gorga, S. T. Neely, D. M. Dierking, P. A. Dorn, B. M. Hoover, D. F. Fitzpatrick: Distortion product otoacoustic emission suppression tuning curves in normal-hearing and hearing-impaired human ears. J. Acoust. Soc. Am. **114** (2003) 263–278.

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

[35] B. C. Moore, B. R. Glasberg: Growth of forward masking for sinusoidal and noise maskers as a function of signal delay; implications for suppression in noise. J. Acoust. Soc. Am. **73** (1983) 1249–1259.

[36] S. T. Neely, D. O. Kim: An active cochlear model shows sharp tuning and high sensitivity. Hear. Res. **9** (1983) 123–130.

[37] S. T. Neely, M. P. Gorga, P. A. Dorn: Distortion product and loudness growth in an active, nonlinear model of cochlear mechanics. – In: Proceedings of the International Symposium on Recent Developments in Auditory Mechanics. H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, T. Koike (eds.). World-Scientific, Singapore, 2000, 237–243.

[38] W. Jesteadt, L. Nizami, S. T. Neely, K. S. Schairer: Effect of noise on the detection of intensity increments. J. Acoust. Soc. Am. **115** (2004) 2600 (abstract).

[39] S. Launer, V. Hohmann, B. Kollmeier: Modeling loudness growth and loudness summation in hearing impaired listeners. – In: Modeling Sensorineural Hearing Loss. W. Jesteadt (ed.). Laurence Earlbaum, Mahwah, New Jersey, 1997, 175-185.