

# Extension of a temporal model of frequency discrimination: Intensity effects in normal and hearing-impaired listeners

Gregory H. Wakefield

*Department of Psychology and Department of Electrical Engineering, University of Minnesota, Minneapolis, Minnesota 55455*

David A. Nelson

*Hearing Research Laboratory, Department of Otolaryngology and Department of Communication Disorders, University of Minnesota, Minneapolis, Minnesota 55455*

(Received 6 April 1984; accepted for publication 4 September 1984)

The effects of intensity on the difference limen for frequency (DLF) in normal-hearing and in hearing-impaired listeners are incorporated into the temporal model of frequency discrimination proposed by Goldstein and Sruлович [ *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, New York, 1977)]. A simple extension of the temporal model, which includes the dependence of phase locking on intensity, is sufficient to predict the effects of intensity on the DLF in normal-hearing listeners. To account for elevated DLFs in hearing-impaired listeners the impairment is modeled as a reduction in the synchrony of the discharge from VIIIth-nerve fibers that innervate the region of hearing loss. Constraints on the optimal processor and the validity of the temporal model at high frequencies are discussed.

PACS numbers: 43.66.Ba, 43.66.Fe, 43.66.Sr

## INTRODUCTION

One recent temporal model of frequency discrimination is that proposed by Goldstein and Sruлович in 1977. In that model, frequency is encoded in the distribution of interspike intervals of VIIIth-nerve discharges. Given a quantitative description of such a probability distribution, Goldstein and Sruлович have shown that human performance is consistent with that of an optimal processor that bases its estimate on an observed set of interspike intervals. Such consistency does not, in itself, imply that the human auditory system functions as an optimal estimator, but does demonstrate that the timing information present in interspike intervals is sufficient to encode frequency with the same precision as the human auditory system.

In their 1977 paper, Goldstein and Sruлович were concerned with the effects of frequency and duration at supra-threshold intensity levels on the difference limen for frequency (DLF). More recently, Sruлович and Goldstein (1983) have extended their model to include both temporal cues and place cues in order to explain a broader range of psychoacoustic phenomena. Within this extended model, the effects of intensity on the DLF have also been assessed and have been shown to be consistent with human data.

The present paper focuses on the effects of hearing impairment on the DLF. Our modeling efforts have been applied to the original temporal model of Goldstein and Sruлович rather than to their more recent temporal-place model. As the authors argue in their later paper, such tasks as frequency discrimination involve primarily the temporal rather than the spatial coding properties of their more comprehensive model. Therefore, we expect that a similar analysis using the latter model would yield conclusions that are comparable to ours.

In the following, we review the original model of Goldstein and Sruлович, extend the model to include intensity effects, and show in what way the model must be modified to

predict the effects of hearing impairment on the DLF. We will consider the sensitivity of the model to perturbations of the underlying representation of VIIIth-nerve discharge, since the dependence of this representation on intensity is not completely known. We will also consider, in greater detail, the additional constraints on the optimal processor that are necessary to yield performance consistent with human data.

## I. REVIEW OF THE MODEL

Siebert was the first to analyze human auditory discrimination as a fundamental limit imposed on the system by the stochastic nature of the VIIIth-nerve response (Siebert, 1968, 1970). In his work, the response of VIIIth-nerve units to pure-tone stimulation is represented as a nonhomogeneous Poisson process, which has since been referred to as the exponential model of VIIIth-nerve response. Bounds on the performance of an optimal processor that is designed to estimate parameters of the physical stimulus from the VIIIth-nerve response are derived and compared with human performance in various discrimination tasks. He concluded that the DLF measured in humans was consistent with the sensitivity of an optimal processor that ignored the timing information contained in the phase-locked activity of neural units and based its estimate solely on the distribution of average discharge rate across a population of units. Although this place model predicts accurately the effects of frequency on the DLF, it does not predict the effects of duration on the DLF (Goldstein and Sruлович, 1977).

Goldstein and Sruлович (1977) revived the possibility of a temporal model of frequency discrimination by demonstrating that the performance of an optimal processor that observed only the second-order statistics of the discharge record, e.g., the interspike intervals, was consistent with the performance of humans when frequency and duration was varied. Like Siebert, they modeled the discharge activity of a

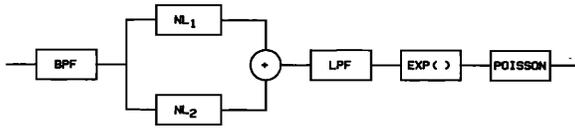


FIG. 1. The exponential model of the VIIIth-nerve response includes two nonlinearities that determine the growth in average discharge rate and in synchrony, respectively, followed by a low-pass filter that limits maximum synchrony. An exponentially rectified version of the output of the low-pass filter drives a Poisson generator.

stimulate neuron as a nonhomogeneous Poisson process. This representation is shown in Fig. 1 and is similar to that found in Johnson (1974).

### A. Exponential model of VIIIth nerve response

The pure-tone stimulus is initially bandpass filtered by a filter with an amplitude characteristic similar to a neural tuning curve. The output of this bandpass filter is then passed through two parallel nonlinearities. The output of the first of these nonlinearities is a dc term, which determines the growth of average discharge rate with intensity. The output of the second nonlinearity is an ac term, which depends on both stimulus amplitude and frequency, and determines the rate of growth of discharge synchrony with stimulus intensity.<sup>1</sup> (Throughout the rest of this paper, these two components will be termed the dc and the ac nonlinearities, respectively.) These two outputs are added and passed through a low-pass filter that limits the degree to which discharges are phase locked to the input. Thus, at high frequencies, where the ac term is attenuated, the output of the low-pass filter will be dominated by the dc term, or average discharge rate, while at lower frequencies, where the ac term is passed with little attenuation, the output will be sinusoidally modulated about the dc term. Finally, another nonlinearity exponentially rectifies the signal (from which the exponential model receives its name) before driving a Poisson generator.

### B. Optimal processor

We assume that the optimal processor is restricted in its use of information provided by the VIIIth nerve. Specifically, the optimal processor cannot base its estimate of frequency on the spread of activity across the population of fibers, but must monitor, instead, the interspike intervals from those fibers that are tuned to the stimulus frequency. Therefore, without loss of generality, we can ignore the initial bandpass filter characteristic and focus entirely on the nature of the remaining components of the model. Johnson (1974) fit neural data to a more general exponential model and developed empirical formulas for the ac and dc nonlinearities and the low-pass filter. Given these expressions, the driving function of the Poisson process can be written as

$$\lambda(t) = \frac{R(x)}{I_0[Z(f,x)]} \exp\{Z(f,x)\cos[2\pi ft + \theta(f)]\}, \quad (1)$$

where

$x$  = stimulus amplitude;

$f$  = stimulus frequency;

$I_0(\ )$  = modified Bessel function of order zero;

$R(x) = R_{sp} + R_d[x/(2+x)]^2$  is the average discharge rate with  $R_{sp}$  = spontaneous average discharge rate (25 sp/s), and  $R_d$  = driven discharge rate (125 sp/s);

$Z(f,x) = A(x)H_L(f)$  is the amplitude of the ac output

with  $A(x) = x/(1+x)$ ,

and  $H_L(f) = K(f)/\{[1 + (f/630)^2][1 + (f/3000)^2]\}^{1/2}$  is the amplitude characteristic of one low-pass filter;

$K(f) = 6.5$ , and  $\theta(f)$  = the phase characteristic of the system (Johnson, 1974; p. 174).<sup>2</sup>

Thus, the driving function is periodic, with a period equal to that of the pure-tone stimulus.

From the driving function, we can write an expression for the probability distribution of interspike intervals during an observation interval:

$$p(\tau) = \left\{ \frac{R(x)\exp[-R(x)\tau]I_0[G(f,x,\tau)]}{I_0[Z(f,x)]^2} \right\} \times \left\{ \frac{RT(1-\tau/T)}{RT-1+\exp(-RT)} \right\}, \quad (2)$$

where

$$G(f,x,\tau) = 2Z(f,x)\cos\pi f\tau.$$

and the second term in braces represents a correction for relatively short duration signals (Goldstein and Sruлович, 1977).

The Crámer-Rao bound on the variance of the optimal estimate can be written as

$$\sigma_{CR}^2 = \left[ NRT \int_0^T \left( \frac{\partial}{\partial f} \ln p(\tau) \right)^2 p(\tau) d\tau \right]^{-1}. \quad (3)$$

Substituting the expression for  $p(\tau)$  into Eq. (3), the Crámer-Rao bound becomes

$$\sigma_{CR}^2 = \left[ NRT \int_0^T \left( \frac{\partial G}{\partial f} \frac{I_1(G)}{I_0(G)} - 2 \frac{\partial Z}{\partial f} \frac{I_1(Z)}{I_0(Z)} \right)^2 \times \left( \frac{RT(1-\tau/T)}{RT-1+\exp(-RT)} \right) p(\tau) d\tau \right]^{-1}, \quad (4)$$

where  $\partial G/\partial f$  = the partial derivative of  $G(f,x,\tau)$  with respect to frequency  $f$ , and  $I_1(\ )$  = the modified Bessel function of the first order. In the above, we have assumed that the estimate is based on the output of  $N = 9$  independent fibers tuned to the standard frequency (Goldstein and Sruлович, 1977) and that each fiber contributes, on the average,  $RT$  independent spikes over the stimulus duration  $T$ . The DLF is determined by numerical evaluation of Eq. (4) using a composite Simpson's rule (Johnson and Riess, 1977) and setting the DLF equal to the square root of the Crámer-Rao bound.

## II. EXPERIMENT AND RESULTS

Frequency difference limens were measured at six frequencies: 300, 600, 1200, 2000, 4000, and 8000 Hz, using a four-interval four-alternative forced-choice (4AFC) adaptive procedure with feedback (for more details, see Nelson *et al.*, 1983). Stimuli were generated by a programmable oscil-

lator (Krohn-Hite 4141R) and gated by an electronic switch with 10-ms rise and decay times before monaural presentation over TDH-49 earphones. Stimulus duration was 300 ms. Programmable attenuators were used to control the intensity of the signals.

Three normal-hearing and two hearing-impaired subjects participated in the experiment. The data for the three normal listeners were originally presented in Nelson *et al.* (1983). Prior to the frequency discrimination task, absolute thresholds at each of the six frequencies were measured on each subject. These measures were used to determine the sound pressure level at each frequency for sensation levels of 10 to, at most, 80 dB.

## A. Effects of intensity on the DLF of normal listeners

### 1. Performance of the model at suprathreshold levels

Results averaged over the three normal listeners at the highest sensation level are shown by the symbols in Fig. 2 with frequency as the parameter. Also shown are the DLFs predicted by the temporal model as a function of stimulus duration. In general, the model *overestimates* performance at 300 ms by at least a factor of 3 for frequencies at or below 2 kHz, and by a factor of 2 at 4 kHz. At 8 kHz, the model *underestimates* performance by about a factor of 2.

Goldstein and Srulovicz (1977) have noted similar discrepancies in the performance of the model for frequencies at or below 4 kHz. They suggest that these discrepancies reflect absolute limitations in the central processing time of interspike intervals. By way of adjusting the model to fit the data, they include an additional component which limits the maximum duration of an interspike interval that can be used by

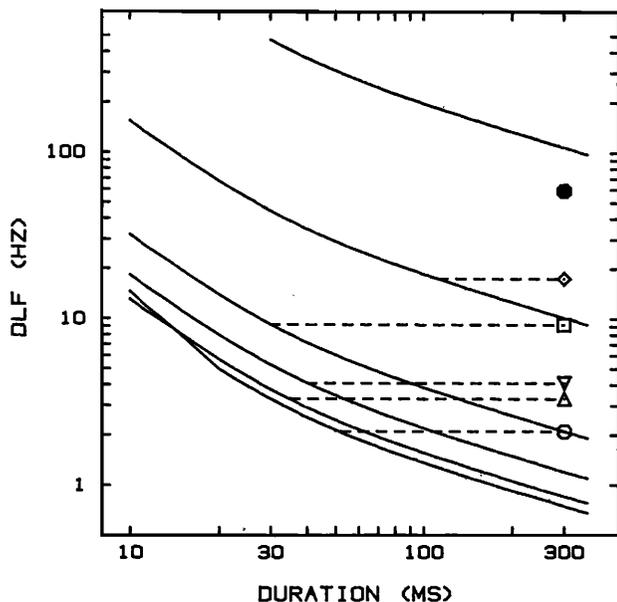


FIG. 2. DLFs averaged over three subjects with normal hearing are shown by the symbols:  $\circ$  (300 Hz),  $\triangle$  (600 Hz),  $\nabla$  (1200 Hz),  $\square$  (2000 Hz),  $\diamond$  (4000 Hz), and  $\bullet$  (8000 Hz). Stimulus duration was 300 ms. Solid curves show the dependence on duration of the DLFs that are predicted by the temporal model. The dashed horizontal lines indicate the estimates of the maximum interspike interval used in the model to predict DLFs that are comparable to those observed for 300 ms.

an optimal estimator. This has the effect of truncating the duration curves at a maximum duration so that for durations greater than the maximum, performance is constant. We can estimate the maximum duration for each of the frequencies by determining the duration along each curve which yields the same value as observed in the average data. These estimates are shown by the dashed lines in Fig. 2. The maximum durations that result are distributed between 30 and 52 ms for frequencies at or below 2 kHz. The maximum duration is 112 ms at 4 kHz.

Such variability in the maximum duration parameter was attributed by Goldstein and Srulovicz to possible differences in sensation level or in experimental procedure. Our results support neither of these explanations. It is apparent that, within experimental error, a single maximum duration may be chosen for frequencies up to 2 kHz, but above this frequency, the maximum duration must be increased substantially in order to agree with the data. However, throughout the rest of this section, we will use the maximum durations estimated for each frequency in fitting the model to the data.

The failure of the model to predict the data at 8 kHz raises a fundamental question concerning phase locking at high frequencies in the VIIIth nerve. According to the exponential model of VIIIth-nerve discharge, phase locking is determined by the low-pass filter that follows the ac and dc nonlinearities. Johnson's empirical formula for such a filter is valid for frequencies up to 5 kHz (Johnson, 1974). Beyond this frequency, in agreement with other physiological data, Johnson observed no phase-locked activity and, based on his analysis of measurement noise in his system, concluded that phase-locked activity does not exist above 5 kHz. Thus, to be consistent with the neural data, the temporal model should be used to predict performance only for those frequencies at which phase locking is observed in the VIIIth nerve. It is interesting that extrapolation of the low-pass filter characteristic beyond 5 kHz yields DLF's for the temporal model that are consistent with one set of psychophysical data (see Fig. 1, Goldstein and Srulovicz, 1977). Nevertheless, such extrapolation is not valid in the context of the physiologically based exponential model.

### 2. Performance of the model as a function of sensation level

The effects of sensation level on the DLF are shown in Fig. 3 along with the predictions of the model. Average results for the three normal-hearing subjects are shown with the DLF plotted in logarithmic units. Frequency is the parameter. These results are similar to those of other investigators (e.g., Harris, 1952; Wier *et al.*, 1977): the DLF is a negatively decelerating function of sensation level which tends to asymptote at levels above 30 dB SL.

The performance bounds of Eq. (4) are expressed in arbitrary units of stimulus amplitude. Therefore, fitting the data requires a conversion from stimulus amplitude to SL at each frequency. This adjustment is equivalent to shifting the entire DLF intensity curve along the intensity axis. In fitting the data, DLF intensity curves were obtained at each of the frequencies tested and then shifted to minimize the mean-square error. The results are shown in Fig. 3.

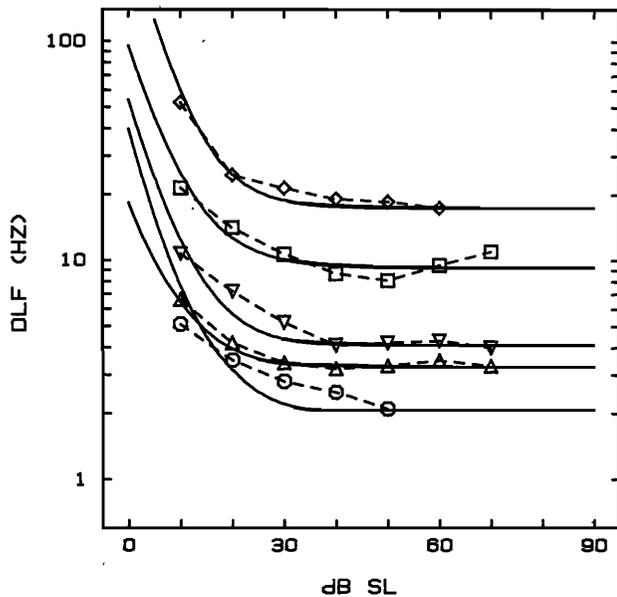


FIG. 3. The effect of intensity on the DLF for five frequencies are shown by the symbols:  $\circ$  (300 Hz),  $\triangle$  (600 Hz),  $\nabla$  (1200 Hz),  $\square$  (2000 Hz), and  $\diamond$  (4000 Hz). Data are averaged over three subjects with normal hearing. The solid curves are the predicted DLFs when the expression for stimulus amplitude in the model is converted to units of SL by minimizing the error.

In general, the model does a satisfactory job of predicting the data. At 0 dB SL, the amplitude of the signal present at the input to the two nonlinearities varied as a function of frequency over a range of 12.5 dB. More specifically, the relative amplitudes expressed in decibels were:  $-20.0$  (300 Hz),  $-12.5$  (1200 and 4000 Hz),  $-10.0$  (600 Hz), and  $-7.5$  (2000 Hz).

Srulovicz and Goldstein (1983) also observed a frequency-dependent offset in absolute threshold for their temporal-place model, but offered no explanation for such dependence. One way to determine whether these offsets are truly dependent on frequency is to consider the value of certain statistics of the discharge pattern at absolute threshold available to the optimal estimator and observe whether these values change, given the offsets determined above. From Eq. (1), we can compute the average discharge rate and the value of synchronization at absolute threshold for the frequencies tested. These calculations show that both average discharge rate  $[R(x)]$  and the amplitude  $[Z(f,x)]$  of the sinusoidal component in the rate function change with increasing frequency over a range of 31 to 51 sp/s and 1.11 to 0.13, respectively. However, if a synchronization index  $[S(f,x)]$  is computed,<sup>3</sup>  $S(f,x)$  at threshold is fairly constant for frequencies at or below 2 kHz (0.31) and drops to 0.06 at 4 kHz. Thus, as in the case of the maximum duration, the additional parameter appears relatively fixed for frequencies up to 2 kHz and must be adjusted substantially at 4 kHz to fit the data.

### B. Effects of hearing impairment on the DLF

The results and the model predictions for two hearing-impaired listeners are shown in Figs. 4 and 5. Both subjects showed moderate to severe sloping losses across the frequencies tested. In general, the DLF intensity curves are shifted toward larger DLFs by at least a factor of 3 but do not neces-

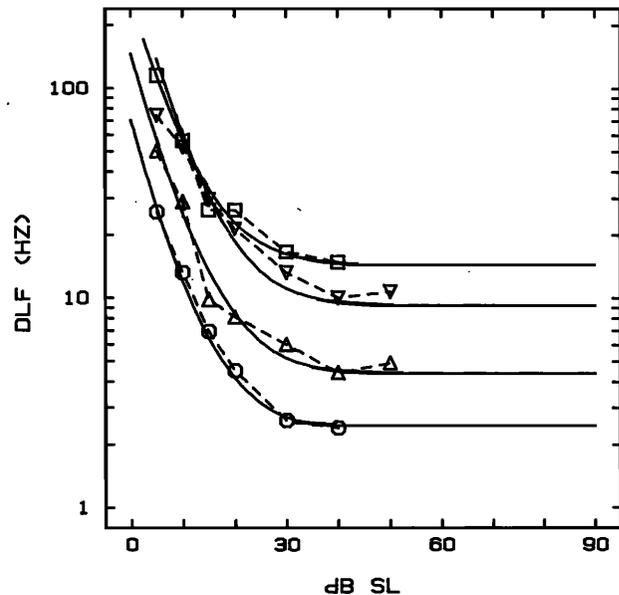


FIG. 4. Similar results as in Fig. 3 but for a hearing-impaired subject (GR). The predicted DLFs are obtained by adjusting the expression for maximum synchronization of the model at each frequency to fit the data. The symbols are the same as those in Fig. 3.

sarily exhibit curvature similar to those of the normal listeners.

There are various ways by which hearing impairment may be represented within the temporal model. One possibility, a simple threshold shift, can be ruled out immediately. An increase in the amplitude of the stimulus at absolute threshold will shift the predicted DLF intensity curves, shown in Fig. 3, to the left. This will tend to flatten the predicted DLF intensity curves and will not shift the curves to higher DLFs, which is what is needed to predict the hearing-impaired data.

Three remaining components of the model were consid-

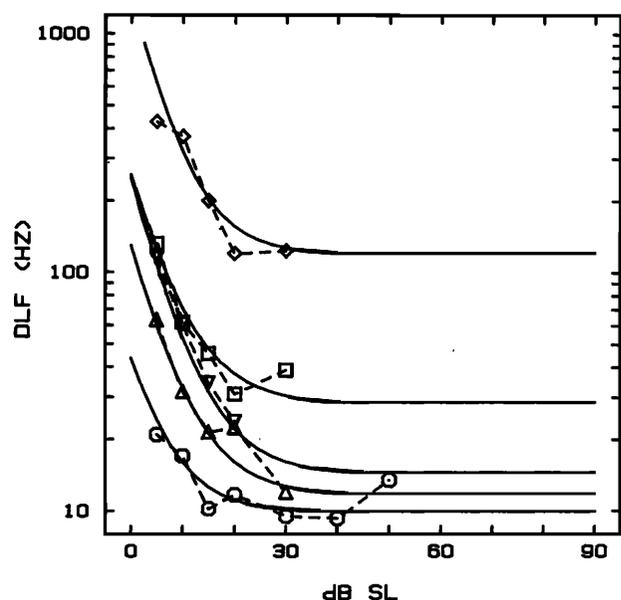


FIG. 5. Same conditions as in Fig. 4 but for a different hearing-impaired subject (SO).

ered: the dc nonlinearity, the ac nonlinearity, and the low-pass filter characteristic. To assess the importance of the dc nonlinearity, we constructed a series of DLF intensity curves by holding constant the value of the dc output and varying the ac output according to Eq. (1). Across this series, the dc output was held constant at average discharge rates [ $R(x)$ ] that ranged from 25 sp/s (spontaneous) to 150 sp/s (saturation) in steps of 25 sp/s. Parametric changes in the dc output produced a vertical displacement in the predicted DLF intensity curves, the direction desired; however, the range was only about 0.1 log unit, which is nowhere near the change required to predict hearing-impaired DLF data. Therefore, the sensitivity of the temporal model to perturbations in the dc nonlinearity is very low; which is encouraging, given that the particular form of this nonlinearity adopted in this paper is somewhat *ad hoc*.

On the other hand, the temporal model is very sensitive to changes in the ac nonlinearity  $A(x)$ . This expression has the general form

$$A(x) = x/(a + x),$$

where  $a$  is a constant. The parameter  $a$  adjusts the "offset" in the saturating nonlinearity: an increase in the value of  $a$  shifts the synchrony-intensity function to greater stimulus intensities, similar to an elevation in threshold. Values of  $a$  were chosen to fit the DLFs measured at the highest intensities for the hearing-impaired subjects. The DLF intensity curves generated by these values rise very quickly below the fitted intensity, which is contrary to the gentle slopes observed in the data. Therefore, for this class of saturating nonlinearities, we cannot model the effects of hearing impairment by manipulating the offset parameter.

The remaining component of the model is the low-pass filter characteristic,  $H_L(f)$ . Attenuation in the output of this component for a given frequency region corresponds to a decrease in the maximum degree of synchronization to tonal stimulation in the response of fibers innervating that region. This attenuation could be achieved by lowering the cutoff frequencies (630 and 3000 Hz) in the low-pass filter. However, this predicts a nondecreasing elevation in the DLF with frequency, which is clearly not the case for these data. The alternative is to allow the gain  $K$  to depend on the frequency region innervated by damaged units in the VIIIth nerve.  $K$  was estimated for each hearing-impaired subject by fitting the DLFs at very high SLs where the effects of errors in modeling the ac and dc nonlinearities are minimal. This value was substituted in the low-pass filter section of the exponential model and DLF intensity curves were generated using the maximum duration determined earlier for the normal subjects. The fitting procedure from the preceding section was used to convert stimulus amplitude to sensation level.

The model predictions based on a frequency-dependent gain in the low-pass filter characteristic are shown in Figs. 4 and 5, along with the hearing-impaired DLF data. In general, the predictions fit the data quite well. Overall, a fivefold change in gain was necessary to fit these data, from a minimum gain of 1.25 for subject SO at 300 Hz, to a maximum

gain of 4.8 for subject GR at 300 Hz (as compared to a gain of 6.5 at all frequencies for the subjects with normal-hearing).

### III. DISCUSSION

Although the temporal model accounts for the effects of intensity on the DLF measured in normal and in hearing-impaired listeners, it is necessary to include three parameters: maximum duration of the interspike interval (ISI), synchronization index at absolute threshold, and attenuation of the gain in the low-pass filter under conditions of hearing impairment. In the following, we consider the first two parameters in greater detail paying particular attention to problems that arise when both parameters are held constant across frequency. Next, we discuss the physiological evidence in support of the assumption that impaired frequency resolution is due to a loss of synchrony in VIIIth-nerve discharges. Finally, we briefly mention alternative ways of extending the temporal model to account for the effects of hearing impairment.

#### A. Maximum ISI duration and synchrony at absolute threshold

In fitting the data, maximum ISI duration and synchronization at threshold were allowed to vary freely with frequency. We observed a fairly tight clustering of the best-fitting values of each parameter for frequencies up to 2 kHz but found that for the 4-kHz condition these values fell significantly outside the range of each cluster. This tight clustering suggests that for frequencies up to 2 kHz, both parameters may be independent of frequency. Should the performance of the model be insensitive to the much larger deviations observed in the 4-kHz condition, the same values may predict the data at 4 kHz as well as at the lower frequencies.

The performance of the model, when the first two parameters, maximum ISI duration and synchronization at threshold, are not allowed to vary with frequency, is shown in Fig. 6. In this case, we have used the means of maximum ISI duration and of synchronization index for frequencies up to 2 kHz to fit the intensity data for the normal listeners. The model predicts the data fairly well between 300 Hz and 2 kHz but predicts an infinite DLF for all intensity levels at 4 kHz, which is outside the limits of the graph and cannot be shown. Therefore, the model is fairly insensitive to small perturbations in these two parameters; however, larger deviations produce significant changes in the predicted performance curves. A better fit might result for the data below 4 kHz by use of a different averaging technique to determine the values of the parameters. Nevertheless, the values of the parameters at 4 kHz differ by such an amount from those at the other frequencies that it is almost certainly necessary to use a minimum of two sets of values to fit the data.

Although introduced in a somewhat arbitrary way, both maximum ISI duration and synchronization index at threshold are reasonable additions to the temporal model and can be interpreted to reflect properties of signal transmission and of the decision process, respectively. For example, throughout the development of this model, we have as-

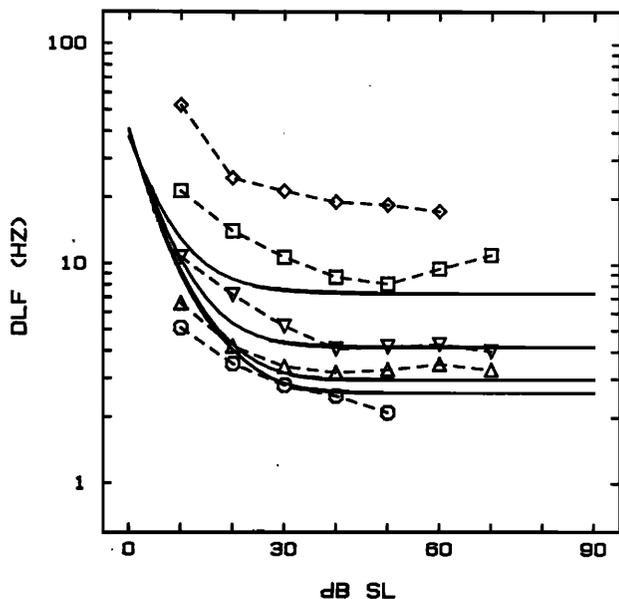


FIG. 6. Predicted DLFs are shown when maximum duration and threshold synchrony are held constant for frequencies at or below 2 kHz. The data for the three subjects with normal-hearing are replotted from Fig. 3. The symbols are the same as those in Fig. 3.

sumed that the ISIs are transmitted through the auditory system with total fidelity. If we assume even a simple jitter in the transmission then the effects of such noise will be greatest for ISIs in the tail of the distribution, i.e., long ISIs as opposed to short.

Just as the maximum ISI may reflect distortions in the transmission of timing information through the auditory system, synchrony at threshold may reflect limitations in the performance of the optimal processor. That is, performance in a frequency-discrimination task must be limited by the detectability of the signals. The present model supports the strong assumption that synchronization index, or some monotonic transformation of it, is the primary determinant of detectability for frequencies where phase locking in the VIIIth nerve is observed. We would prefer, however, to argue in favor of a weaker version of this assumption: The appropriate transformation of stimulus amplitude to sensation level must incorporate those characteristics of the neural discharge that determine stimulus detectability and these characteristics must be independent of stimulus frequency.

These additional components, though reasonable, become questionable when forced to account for the 4-kHz data, since it does not appear reasonable to assume a frequency dependence of maximum ISI duration or of synchrony at absolute threshold. It is likely that the changes in these components necessary to fit the data at 4 kHz reflect the more fundamental failure of the model at high frequencies. We have already noted that phase locking has been observed for frequencies up to, but rarely beyond, 5 kHz. In the absence of periodicity in the ISIs generated by phase locking, the optimal processor will not be able to discriminate frequency. The substantially poorer performance at 4 kHz predicted by the parameters used to fit the data at 2 kHz and below suggests that an alternative coding scheme may al-

ready be in use in the 4-kHz region. We expect that the temporal-place model proposed more recently by Sruulovicz and Goldstein will have similar problems in providing a consistent explanation of the data.

### B. Loss of synchrony

Of the variety of components within the exponential model of VIIIth-nerve response, the one most successful in predicting the effects of hearing impairment on the form of the DLF intensity curves was the gain of the low-pass filter. Attenuation of the output of this filter at a given frequency can be interpreted as a loss in synchrony among fibers tuned to that frequency. Two physiological studies have considered the effects of hair-cell damage on phase locking in primary fibers of the VIIIth nerve. Harrison and Evans (1979) found no evidence of deterioration in the ability of damaged units to phase lock to stimulus frequency in the guinea pig. On the other hand, Woolf *et al.* (1981) observed considerable deterioration in synchrony in damaged units of the chinchilla. Specifically, synchronization coefficients were reduced by as much as 60% from those measured in normal units. We find that the percentage reduction in the synchronization coefficient produced by changes in the gain of the low-pass filter range from a minimum of 4% at 300 Hz for subject GR to a maximum of 62% at 4 kHz for subject SO. Therefore, the reduction in gain necessary to fit the human data is consistent with the change in synchrony observed in abnormal units of the VIIIth nerve by Woolf *et al.*

Woolf *et al.* found a strong correlation between elevated behavioral thresholds and deteriorated phase locking in VIIIth-nerve units. We, however, do not observe a very strong correlation between elevated detection thresholds and elevated DLFs that presumably, reflect impaired phase locking. For example, at 300 Hz, subject GR exhibited normal frequency resolution whereas the DLF for subject SO was increased by a factor of 5 at suprathreshold levels. Nevertheless, both have elevated absolute thresholds in this frequency region (39 dB for GR and 28 dB for SO). Thus, although the gain of the low-pass filter may correspond to a loss in synchrony, it must remain a free parameter when fit to the frequency discrimination data: A measure of elevated absolute threshold does not predict the low-pass filter loss.

### C. Alternative approaches

In modeling the effects of hearing impairment, we have limited our discussion to components in the exponential model. As an alternative, the effects of hearing impairment can be modeled as changes in the characteristics of the temporal processor that do not depend on the exponential model of the VIIIth-nerve response. For example, a decrease in the number of fibers across which information is combined results in a deterioration in the performance of the optimal processor. Alternatively, a decrease in the maximum duration of the interspike interval also results in increased DLFs. Although either approach yields predictions that fit the data to a certain extent, neither performs as well as the approach taken in this paper. Therefore, in the absence of any physiological support for these alternatives, we prefer to model the effects of hearing impairment as a loss in synchrony in VIIIth-nerve discharges. Whether or not this loss in synchrony

rony is a typical physiological correlate of cochlear damage awaits additional research.

#### IV. CONCLUSION

The temporal model of Goldstein and Srulovicz (1977) can be modified to account for the effects of intensity and hearing impairment on the difference limen for frequency measured in humans. However, several additional assumptions are necessary. (1) The maximum duration of an interspike interval must be limited and depends, to some extent, on frequency. (2) A frequency-dependent transformation of stimulus intensity into SL is necessary. At best, this transformation produces a weak dependence of the synchronization index at absolute threshold on stimulus frequency. Other measures of nerve response show much greater frequency dependence. (3) Impaired frequency discrimination, which often accompanies a hearing loss, is modeled as substantial changes in the degree of phase locking observed in the VIIIth-nerve response. Alternative approaches exist but, as yet, corroborative physiological evidence is not available. (4) If we assume that phase locking in the VIIIth nerve to frequencies above 5 kHz does not exist, the temporal model cannot account for frequency discrimination at such frequencies. A model which assumes that temporal cues predominate at lower frequencies and place cues predominate at higher frequencies may also be able to account for the poor performance of the temporal model at 4 kHz when the parameters of this model are constrained.

#### ACKNOWLEDGMENT

This research was supported by NINCDS Grants NS12125 and NS00449.

<sup>1</sup>In their original paper (Goldstein and Srulovicz, 1977), these nonlinearities, which depend on the intensity of the stimulus, are absent. If the intensity of the stimulus is such that the nerve fiber operates in its saturation region, then the present model reduces to that treated in the original paper.

<sup>2</sup>The expression for the average discharge rate is different from that given in Johnson (1974). In our early work on this extension, we included Johnson's form of this nonlinearity, which also depends on frequency, and obtained predictions for the DLF at saturated levels. These predictions generated DLF curves that were substantially different from those obtained in humans. As noted by Johnson, the frequency dependence results in estimates of average discharge rate that vary somewhat with frequency, a result that is at variance with the physiological data. This additional "false" depen-

dence of average rate on frequency is responsible for the different form of the DLF curve predicted by the model. We have substituted the form used by Srulovicz and Goldstein [their Eq. (3e), 1983] normalized by  $I_0[Z(f,x)]$ . This expression does not have much physiological support either, but at least does not produce frequency-dependent average discharge rates. As we argue later in this paper, the model is fairly insensitive to the precise form of this expression as long as it does not depend on frequency. The DLF curves that result are consistent with the human data.

<sup>3</sup> $S(f,x)$  is the magnitude of the fundamental component in the Fourier series expansion of the driving function which is normalized to the average discharge rate (Goldberg and Brown, 1969). Thus  $S(f,x) = (a^2 + b^2)^{1/2}$ , where

$$a = \frac{1}{T} \int_0^T \frac{\cos 2\pi ft}{I_0[Z(f,x)]} \exp[Z(f,x)\cos(2\pi ft) + \theta(f)] dt,$$

$$b = \frac{1}{T} \int_0^T \frac{\sin 2\pi ft}{I_0[Z(f,x)]} \exp[Z(f,x)\cos(2\pi ft) + \theta(f)] dt,$$

and  $T = 1/f$ . Therefore,  $S(f,x) = I_1[Z(f,x)]/I_0[Z(f,x)]$ .

- Goldberg, J., and Brown, P. R. (1969). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization," *J. Neurophysiol.* **32**, 613-636.
- Goldstein, J. L., and Srulovicz, P. (1977). "Auditory-nerve spike intervals as an adequate basis for aural frequency measurements," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London, 1977), pp. 337-346.
- Harris, J. D. (1952). "Pitch discrimination," *J. Acoust. Soc. Am.* **24**, 750-755.
- Harrison, R. V., and Evans, E. F. (1979). "Some aspects of cochlear hair cell degeneration in the guinea pig," *Arch. Otorhinolaryngol.* **224**, 71-78.
- Johnson, D. H. (1974). "The response of single auditory-nerve fibers in the cat to single tones: Synchrony and average rate," Ph.D. thesis, MIT Cambridge, MA.
- Johnson, L. W., and Riess, R. D. (1977). *Numerical Analysis* (Addison-Wesley, Reading, MA).
- Nelson, D. A., Stanton, M. E., and Freyman, R. L. (1983). "A general equation describing frequency discrimination as a function of frequency and sensation level," *J. Acoust. Soc. Am.* **73**, 2117-2123.
- Siebert, W. M. (1968). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns*, edited by P. A. Kohlers and M. Eden (MIT, Cambridge, MA), pp. 104-133.
- Siebert, W. M. (1970). "Frequency discrimination in the auditory system: place or periodicity mechanisms?," *Proc. IEEE* **58**, 723-730.
- Srulovicz, P., and Goldstein, J. L. (1983). "A central spectrum model: A synthesis of auditory-nerve-timing- and place-cues in monaural communication of frequency spectrum," *J. Acoust. Soc. Am.* **73**, 1266-1276.
- Wier, C. C., Jesteadt, W., and Green, D. M. (1977). "Frequency discrimination as a function of frequency and sensation level," *J. Acoust. Soc. Am.* **61**, 178-184.
- Woolf, N. K., Ryan, A. F., and Bone, R. C. (1981). "Neural phase-locking properties in the absence of cochlear outer hair cells," *Hear. Res.* **4**, 335-346.