A computer model of the auditory-nerve response to forward-masking stimuli

Ray Meddis[a)] and Lowel P. O'Mard
Centre for the Neural Basis of Hearing at Essex, Department of Psychology, University of Essex, Colchester, CO4 3SQ, United Kingdom

A computer model of the auditory periphery is used to study the involvement of auditory-nerve (AN) adaptation in forward-masking effects. An existing model is shown to simulate published AN recovery functions both qualitatively and quantitatively after appropriate parameter adjustments. It also simulates published data showing only small threshold shifts when a psychophysical forward-masking paradigm is applied to AN responses. The model is extended to simulate a simple but physiologically plausible mechanism for making threshold decisions based on coincidental firing of a number of AN fibers. When this is used, much larger threshold shifts are observed of a size consistent with published psychophysical observations. The problem of how stimulus-driven firing can be distinguished from spontaneous activity near threshold is also addressed by the same decision mechanism. Overall, the modeling results suggest that poststimulatory reductions in AN activity can make a substantial contribution to the raised thresholds observed in many psychophysical studies of forward masking. © *2005 Acoustical Society of America.*
[DOI: 10.1121/1.1893426]

I. INTRODUCTION

In physiological experiments, the responsiveness of the auditory nerve (AN) is reduced immediately following acoustic stimulation. In psychophysical experiments, thresholds are temporarily raised following similar acoustic stimulation. In both cases, the changes depend on the level and the duration of the preceding stimulation. Recovery, in both cases, occurs over a period lasting more than 100 ms. These superficial similarities suggest that the two phenomena may be linked. Indeed, both physiologists and psychophysicists use the same term, "forward masking," to describe the respective phenomena.[1]

However, Relkin and Turner (1988) have indicated the potential for confusion inherent in an uncritical equation of the two phenomena. They point out that a reduction in physiological response is not the same thing as an increase in perceptual threshold, and that different measurement methods are involved in their respective estimation. When they attempted to apply psychophysical threshold estimation techniques to single-fiber AN observations, they found that the poststimulatory reduction in threshold was considerably less than that seen in human psychophysics. They concluded that the increase in psychophysical thresholds must occur later in the processing system. Specifically, they claim that the nervous system does not use an optimum decision strategy to evaluate the information present in the AN response.

A number of physiological processes have been proposed as possible contributors to psychophysical forward masking. These include peripheral frequency selectivity (Duifhuis, 1973), adaptation in the auditory nerve (Smith,

1977, 1979), efferent inhibitory processes (Shore, 1998), and persistence of neural activity (for example, Plomp, 1964; Moore *et al.*, 1988). See Oxenham (2001) for a brief, recent overview of the issues. It is not intended, in this study, to decide between these alternatives because all are likely to make some contribution. In any case, as Oxenham (2001) has shown, it is difficult to separate their relative contributions experimentally. However, the Relkin and Turner (1988) study and another similarly motivated study (Turner *et al.*, 1994) could create the impression that forward masking does not have its origin in adaptation in the AN. This conclusion goes beyond what is warranted by the data, and it is the purpose of the present study to use computational methods to investigate the matter further.

The study employs an already-published model of the mammalian auditory periphery to simulate AN responses with appropriate poststimulatory adaptation and recovery characteristics. It will be demonstrated that the model is appropriate for the purposes of the investigation by showing that it can be tuned to simulate observations of depression and recovery of the chinchilla AN response following stimulation (Harris and Dallos, 1979). It will also be shown that the model can be used to replicate the findings of Relkin and Turner (1988) referred to above. Finally, a simple coincidence-detection mechanism mimicking the response of a first-order neuron in the cochlear nucleus will be modeled. When this takes as its input the combined response of a small number of AN fibers, it will be demonstrated that substantial increases in threshold do occur. These increases are considerably greater than those seen in Relkin and Turner's study and comparable to those seen in psychophysical studies.

The coincidence-detection mechanism is proposed as a solution to a more general problem in auditory physiology

---

[a)]Author to whom correspondence should be addressed. Current address: Department of Psychology, University of Essex, Colchester, CO4 3SQ, United Kingdom. Electronic mail: rmeddis@essex.ac.uk

concerning how the nervous system distinguishes between spontaneous AN activity and activity driven by acoustic stimulation. While it is natural to believe that the system does this by detecting an increase in the average firing rate of individual auditory-nerve fibers, the physiological mechanism underlying this kind of detection strategy remains problematic. For example, Relkin and Turner's decision-making process implies that the system works by establishing a running count of action potentials (APs) and assesses its significance in association with an expected mean and variance associated with spontaneous firing rate. This optimal decision method is difficult to implement directly in terms of the known anatomy and physiology of the auditory nervous system because it requires secondary mechanisms to estimate and deploy running estimates of these two statistical quantities. This is difficult to describe in terms of the operation of nerve fibers and nerve cells. However, "coincidence-detection" is a simple alternative that achieves the same objective.

Spontaneous activity in one AN fiber is uncorrelated with that in other fibers. As a consequence, the occurrence of near-simultaneous APs is relatively rare in a small group of fibers such as might innervate a single cell in the cochlear nucleus. However, when APs are driven by acoustic stimulation, the overall firing rate will increase and the number of near-simultaneous events will also increase. Coincidental firing will also be more probable near the onset of the stimulation because similar AN fibers will have similar first-spike latencies (Heil and Neubauer, 2001). In the case of low-frequency sounds, phase locking to the stimulus will further increase the amount of coincidental activity. On all three counts, coincidental firing is expected to be greater following acoustic stimulation than in silence. If a cell in the cochlear nucleus responds only to coincidental firing of AN fibers, it will respond almost exclusively to acoustic stimulation and only very rarely to spontaneous activity in the AN fibers. This system is consistent with the architecture of the auditory brainstem, where neurons integrate information across a number of input fibers when generating responses.

After a description of the model, this account is divided into three sections, each describing a different evaluation of the model. First, the recovery of the response of the model after stimulation is compared with the data of Harris and Dallos (1979). Second, the model is used to replicate the study of Relkin and Turner (1988). Third, the coincidence detection mechanism is introduced and the model is evaluated using thresholds obtained using a forward masking paradigm (Jesteadt et al., 1982).

II. MODEL DESCRIPTION

The computer model of the mammalian auditory periphery to be used in this evaluation was presented by Sumner et al. (2002, 2003a, 2003b) for the guinea pig auditory periphery. Exactly the same model is used except where specifically indicated. This model consists of a chain of separate modules representing, respectively, the pre-emphasis filtering of the outer/middle ear, the vibration of the basilar membrane, the inner hair cell (IHC) receptor potential, IHC/AN synaptic activity, and AN activity. This model is capable of simulating a wide range of physiological measurements but has not yet been evaluated in the context of recovery following stimulation. A full account of the model will not be repeated here and the reader is referred to the descriptions given by Sumner et al. Parameters for the model can be found in Sumner et al. (2002, Tables I[2] and II; 2003b, Table I). The parameters for high, medium, and low spontaneous rate (HSR, MSR, and LSR) fibers are taken from Sumner et al. (2003b, Table I). The original model was tuned using guinea pig data, but the simulations to be described involve using chinchilla and human data. The assumption is made that the underlying processes are the same as far as our present purposes are concerned for all these species.

A. Middle-ear filtering

Middle-ear filtering is modeled by two cascaded linear bandpass Butterworth filters. One filter is second order, with lower and upper cutoffs of 4 and 25 kHz. The second is third-order and has lower and upper cutoffs of 0.7 and 30 kHz. A scaling factor is used to convert sound pressure to realistic stapes velocities (see the changed parameters below).

B. Mechanical filtering

A dual-resonance nonlinear (DRNL) filter is used to simulate the mechanical bandpass filtering of the basilar membrane (BM) in the cochlea. The construction and properties of the DRNL filter have been described in detail elsewhere (Meddis et al., 2001; Sumner et al., 2003b). The DRNL filter consists of two pathways, one linear and the other nonlinear. The signals at the output of the two pathways are combined by simple summation to produce the filter output. Each pathway consists of a cascade of a number of first-order gammatone filters (Patterson et al., 1988) and a cascade of second-order Butterworth low-pass filters.

In the nonlinear pathway, a compression function is located in the middle of the cascade of four gammatone filters between the second and third filter. The linear pathway consists of a cascade of three gammatone filters but does not contain nonlinearity. It has a center frequency that is different from that of the linear path; this contributes asymmetry to the overall filter function. Additional low-pass filters are also employed to improve the fit to animal data. The shape of the bandpass filters is not critical to this study because the stimuli to be used are all at the same frequency, and the best frequency of the filter was always set at the frequency of the stimulus. However, the nonlinear compression will affect the pattern of growth of forward masking at high masker levels.

C. Inner hair cell (IHC) receptor potential

The process of mechanical to electrical transduction is simulated by a model of the IHC receptor potential that converts BM velocity to IHC receptor potential through the deflection of the IHC stereocilia.

The receptor potential controls the rate of calcium influx into the cell at the synapse. The calcium is accumulated in a leaky integrator, and the instantaneous level of calcium controls the rate of transmitter release. In the model, the differ-
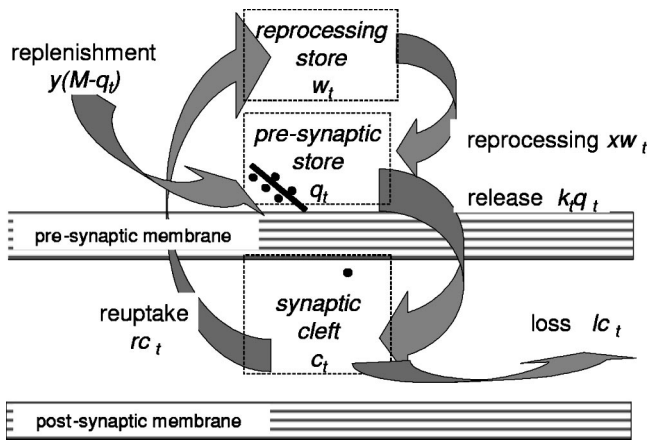
FIG. 1. Flow of transmitter in a model of the Inner hair cell/auditory-nerve synapse. See the text for full description.

ences between fiber types (HSR, MSR, and LSR) are the result of different rates of influx of calcium into the presynaptic sites. HSR sites have a faster rate of influx and correspondingly higher spontaneous release rates. The rate of calcium influx into the IHC for a given receptor potential is specified by two parameters: the maximum rate of influx, GCa_{max}, and a threshold parameter, GCa_{thr}. These parameters are unchanged from Sumner *et al.* (2003b, Table I).

D. IHC synapse

In this study, the IHC/AN synapse is of particular interest. The depression of response following a masker is assumed to result from a depletion of available transmitter substance (Smith, 1977, 1979). Figure 1 illustrates how the model handles the flow of transmitter. A small number, $q(t)$, of vesicles of transmitter is held in a presynaptic store. Each vesicle has an equal probability, $k(t)d(t)$, of being released in a single epoch of duration dt. $k(t)$ is a function of the accumulated presynaptic calcium.

When transmitter is released into the cleft, the amount of presynaptic transmitter, $q(t)$, necessarily falls and the likelihood of further release events is reduced proportionately; the model is said to have ''adapted.'' Transmitter material is recycled as follows. Reuptake into the cell occurs at a rate $rc(t)$, where $c(t)$ is the amount of transmitter in the synaptic cleft and r is the rate of reuptake of transmitter into the cell. Some transmitter is lost from the cleft at a rate $lc(t)$, where l is the rate of loss of transmitter from the cleft. After reuptake, transmitter material is reprocessed and returned to the available presynaptic pool at a rate $xw(t)$, where $w(t)$ is the amount of material in the reprocessing store and x is the rate at which transmitter is returned from the reprocessing store to the pre-synaptic available store. The loss of transmitter material from the cleft is slowly compensated by a replenishment mechanism that supplies new transmitter at a rate $y[M\text{-}q(t)]$, where M is maximum capacity of the available store and y is the rate at which new transmitter is made available. Transfer of vesicles in and out of the presynaptic store is a quantal stochastic process $N(n,\rho)$, where each of n possible events has an equal probability, ρdt, of occurring in

a single simulation epoch. The flow of transmitter is specified by the following differential equations:

$$dq(t)/dt = N[w(t),x] + N([M-q(t)],y)$$
$$- N[q(t),k(t)], \qquad (1)$$

$$dc(t)/dt = N[q(t),k(t)] - lc(t) - rc(t), \qquad (2)$$

$$dw(t)/dt = rc(t) - N[w(t),x]. \qquad (3)$$

After prolonged and intense stimulation, the available store can become completely depleted. If forward masking is partly explained by depletion of neurotransmitter at the IHC/AN synapse, the time course of recovery from forward masking will be partly governed by the rate at which the store of available transmitter is refilled. Recovery from depletion takes place only through the reprocessing and replenishment routes. As a consequence, parameters x and y are important determinants of the rate of recovery. These two parameters of the model will be adjusted below to match recovery measurements made in the AN.

E. AN response

An action potential is deemed to occur when one or more vesicles are released into the synaptic cleft so long as the postsynaptic AN fiber is not already in a refractory state. An absolute refractory state follows an action potential and lasts for 0.75 ms and the exponential recovery of responsiveness following the relative refractory period has a time constant of 0.6 ms.

The response of the fiber can be evaluated either as a probability of firing (probabilistic mode) or by generating individual spike events using a random number generator (stochastic mode). The former is much quicker to compute and will be used for exploring AN recovery from adaptation (Sec. III). However, for threshold measurements (Secs. IV and V), the model is evaluated in stochastic mode.

F. Model implementation

The model was evaluated at an update rate of 100 kHz. It was implemented using the Development System for Auditory Modelling (DSAM), a library of computer modules simulating each of the stages of the auditory model described above. The software used in this study is available on the Worldwide Web at http://www.essex.ac.uk/psychology/hearinglab/

G. Parameter changes

The scaling factor in the middle-ear simulation was changed from its previous value of $1.4E-10$ to $6E-10$ m/s/μPa. This increases the middle-ear gain by 12.6 dB relative to the previous value and gives AN rate-level functions with similar thresholds to those given in Harris and Dallos (see Fig. 2).

Two IHC/AN synapse parameters were altered to give an improved fit to the chinchilla forward-masking data reported by Harris and Dallos (1972). The rate of transmitter replenishment, y, was reduced from 10 to 3 vesicles/s and the reprocessing rate, x, was reduced from 66 to 30 vesicles/s.
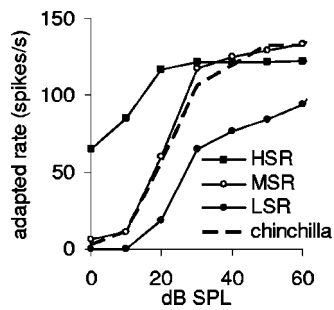
FIG. 2. Rate/level function for model HSR, MSR, and LSR fibers. The dashed line shows a chinchilla rate-level function taken from Harris and Dallos (1979, Fig. 9). Their function is normalized to the rate response of the model MSR fiber at 60 dB SPL.

These changes had little effect on the previously published properties of the model except for a slowing in the time constant of short-term adaptation from 80 to 120 ms. This is longer than that reported for the gerbil by Westerman (1985), but consistent with reports by Smith (1979) in gerbil and Yates *et al.* (1985) in guinea pig. Chimento and Schreiner (1990) report adaptation time constants of 125 ($\pm$52) ms in cat. The model short-term adaptation was measured starting 20 ms after the stimulus onset. Yates *et al.* (1985) claim to have seen an association between the rate of adaptation and the rate of recovery in their observations. This would be consistent with the change observed in the model. Unfortunately, they do not supply any quantitative description of this effect.

III. EVALUATION: AN RECOVERY FROM ADAPTATION

The model was evaluated using a forward-masking paradigm and the same stimuli as those described in Harris and Dallos (1979). Except where stated, the stimulus characteristics were as follows. All maskers and probes were pure tones presented at BF (5750 Hz). Stimulus durations were 100 ms for maskers and 15 ms for probes. The gaps between masker and probe were 1, 2, 5, 10, 20, 50, 100, and 300 ms. Both masker and probe had a rise and fall time of 1 ms using a raised-cosine ramp function. The response to the probe was measured as the total number of spikes observed during the presentation of the probe after allowing for the latency of the response.

A. Rate-level functions

Model rate-level functions for BF tones of 6 kHz are given in Fig. 2. The rate-level functions show adapted rates based on the last 50 ms of the response to the 100-ms masker tones. Functions were measured for three fiber types, HSR, MSR, and LSR. The rate-level functions for the MSR fiber were used to establish the best setting for the stapes gain parameter (see Sec. II G) using a comparison with a single chinchilla fiber given by Harris and Dallos (1979, Fig. 9).

B. Response recovery

Figure 3 illustrates the response of the model to the forward-masking paradigm. The response to the probe increases as the gap between the masker and the probe widens. Figure 4 quantifies this effect for all three types of fiber and
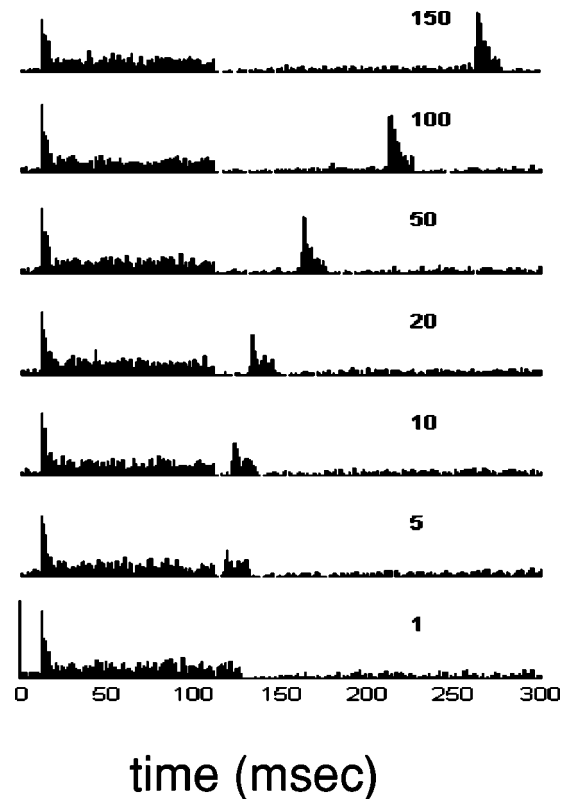

FIG. 3. PSTH response of a model HSR fiber at a range of different masker–probe gaps (ms, indicated in the figure). Masker and probes are all presented at 20 dB above threshold. For this illustration, the model was evaluated in stochastic mode.

a range of masker levels. An example of the data from Harris and Dallos is also given for comparison purposes. Recovery functions based on previously published parameters were found to be too fast (not shown) and the reprocessing and replenishment rates have been adjusted as described to give a more acceptable quantitative fit (see below) to the animal data.

For model HSR fibers, the depression in response at short masker–probe gaps increases over a narrow range of masker levels. For MSR and LSR fibers, this range is considerably greater. Harris and Dallos also found that the depression in responding ceases to grow at some high masker level that is normally the same as the level at which the rate-level function saturates. This possibility was examined in the model. Figure 5 compares the rate-level function for each fiber with the depression of response to the probe immediately after the masker. Both sets of data are normalized between 0 (minimum response) and 1 (maximum response). The functions are closely matched for all three types of fibers.

Longer maskers were shown by Harris and Dallos to give rise to a greater depression in probe response than shorter maskers. Figure 6 summarizes the response of all three model fiber types to maskers of different durations. All three model fiber types show the expected effect.

C. Rate of recovery

At low masker levels, Harris and Dallos found that recovery could be described by a simple exponential recovery

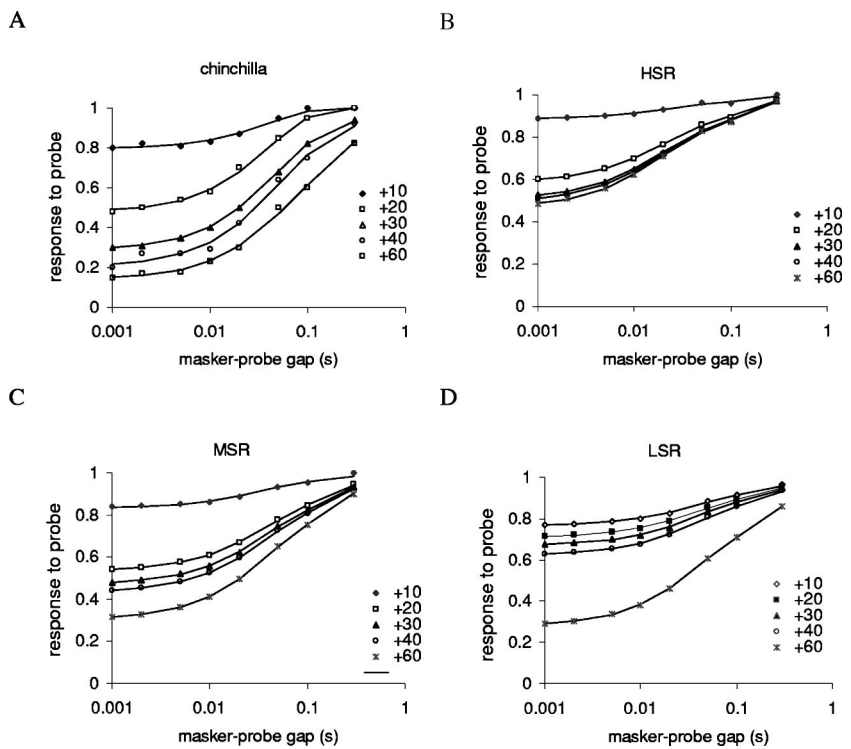R. Meddis and L. P. O'Mard: Computer model of forward masking

FIG. 4. Response to the probe following stimulation as a function of the masker–probe gap for different masker levels (indicated). All levels are relative to the rate threshold. Response magnitude is the total number of spikes during the presentation of the probe normalized to the probe response measured in the absence of a masker. Continuous lines are double-exponential functions fit to guinea pig and model fiber data. The function fitted to each data set has the form $1 - a\exp(-t/\tau_a) - b\exp(-t/\tau_b)$. The parameters used in the fitting process are shown in Table I. (A) Data from Harris and Dallos (1979, Fig. 10, median of 37 fibers). (B)–(D) Model response (HSR, MSR, and LSR fibers, respectively).

function, $1 - a\exp(-t/\tau)$, where t is the time since the offset of the masker. However, for more intense masker levels a simple exponential recovery function was a poor fit. Harris and Dallos assumed that the rate of recovery was changing during recovery. Accordingly, they employed different exponential fits to their data for short gaps (<50 ms) and long gaps. Using this approach at higher masker levels, it was necessary to specify short recovery time constants (30 to 70 ms) for the early part of the recovery, but long recovery time constants (30 to 290 ms) towards the end of recovery. Also, longer time constants were required to fit the recovery following higher level maskers.

A reanalysis of their chinchilla data (their Fig. 10) based on the median thresholds of 37 fibers was carried out. It showed that a double exponential recovery process, $1 - a\exp(-t/\tau_a) - b\exp(-t/\tau_b)$ gives a good account of their data without the need to change the time constants either as a function of time or masker intensity. The fit to their data is shown in Fig. 4(A) (continuous lines) and the numerical values of the parameters are shown in Table I. The analysis

indicates that two time constants (40 and 280 ms) characterize all the data in the figure so long as the coefficients a and b are allowed to vary as a function of masker level. A good fit using a double-exponential recovery process might mean that two separate recovery functions are at work. If so, it makes the animal data qualitatively consistent with the model which has two recovery processes (reprocessing and replenishment).

The model responses are described using the same function [Figs. 4(B)–(D), Table I]. The slow recovery time constant is substantially shorter (158 ms) for model HSR fibers than for both MSR and LSR fibers (228 ms). It is, therefore, a prediction of the model that LSR fibers will recover more slowly given comparable stimulation. Note that all three model fiber types have the same reprocessing and replenishment rate parameters. The between-fiber differences in the observed time constants are a consequence of the different rates of calcium influx into the cells; this is the only difference in the model between fiber types.
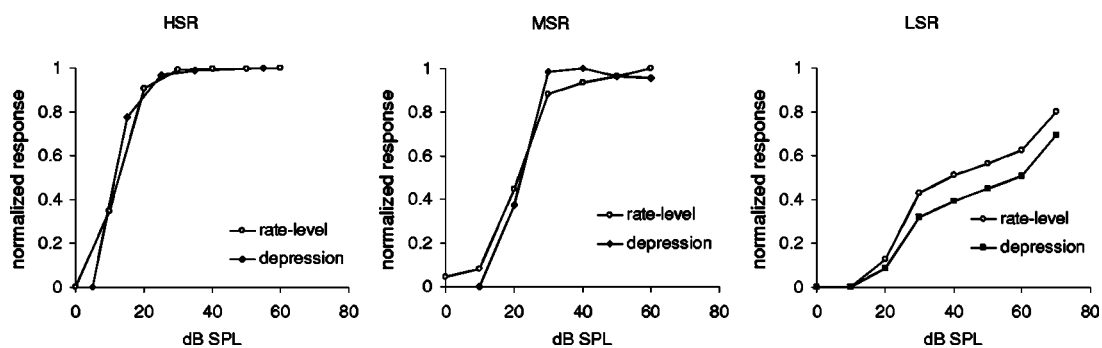
Despite the parameter changes, the model recovery is

FIG. 5. Comparison of rate-level function and extent of the depression in the response to the probe at the shortest (1-ms) masker–probe gap. All data are normalized between 0 (minimum response) and 1 (maximum response). The LSR did not saturate and the data were normalized to the maximum at 90 dB SPL (not shown).
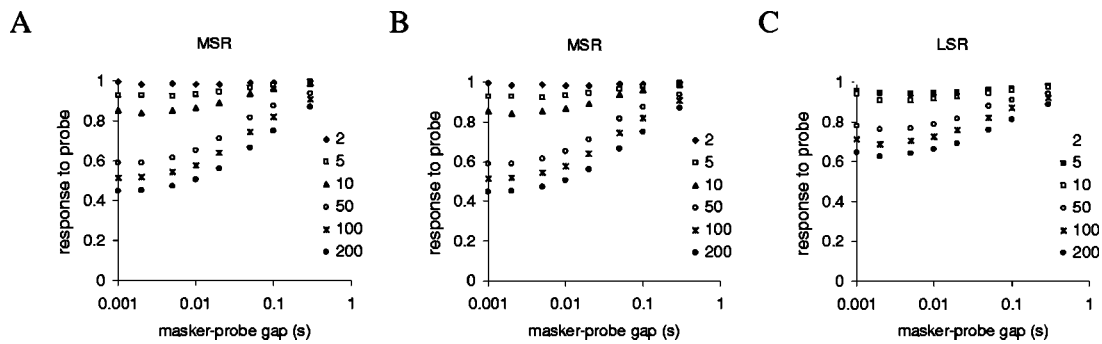
FIG. 6. Normalized recovery functions for HSR, MSR, and LSR fibers as a function of masker *duration*. Both maskers and probes are 20 dB above rate threshold for the individual fiber. Masker durations (ms) are indicated.

still somewhat faster than the median data of Harris and Dallos. Recovery can be slowed further by reducing the values of parameters *x* and *y*. However, the slower reprocessing leads to reduced firing rates and a rebalancing of many parameters is required to re-establish typical AN fiber function. The difference between the model performance and the animal data is small enough to justify keeping the other parameters as previously published in the interest of simplicity and transparency. Certainly, the model recovery functions are well within the range of functions illustrated in Harris and Dallos' report, some of which show recovery faster than the model.

D. Recovery of spontaneous rate

Harris and Dallos demonstrated that the time course of recovery from forward masking runs parallel to the recovery of spontaneous activity after stimulation. In the model, both functions are dependent on amounts of transmitter in the

TABLE I. Coefficients and time constants used to generate the best-fit functions in Fig. 4. The time constants are constrained to be the *same* for a given data set irrespective of masker level. Coefficients *a* and *b* are allowed to vary. They represent the relative contribution of reprocessing and replenishment respectively for a particular masker level.

	Masker level				
	+10	+20	+30	+40	+60
Chinchilla					
a	0.20	0.52	0.52	0.53	0.35
τ_a	**0.04**				
b	0.00	0.00	0.19	0.27	0.51
τ_b	**0.28**				
HSR					
a	0.06	0.23	0.28	0.29	0.31
τ_a	**0.016**				
b	0.06	0.19	0.22	0.22	0.23
τ_b	**0.158**				
MSR					
a	0.11	0.25	0.27	0.29	0.34
τ_a	**0.030**				
b	0.06	0.22	0.26	0.28	0.36
τ_b	**0.227**				
LSR					
a	0.14	0.17	0.19	0.21	0.33
τ_a	**0.035**				
b	0.06	0.09	0.12	0.15	0.37
τ_b	**0.228**				

available presynaptic store and therefore both are expected to recover at the same rate. Figure 7(A) shows the recovery of spontaneous activity in the HSR model fiber following a masker tone whose level is 20 dB above threshold. Figure 7(B) systematically compares the recovery of spontaneous activity with the recovery of the response to a probe of the same level as the masker. Spontaneous recovery was measured using probe-width windows positioned where the probe would be when using a forward-masking paradigm. The two functions are virtually identical.

IV. EVALUATION; AN THRESHOLD INCREASE

Relkin and Turner (1988) estimated threshold changes following stimulation using a measurement paradigm (Relkin and Pelli, 1987) that was chosen to be as similar as possible to the psychophysical two-alternative forced choice (2AFC) methodology normally used in behavioral threshold measurements. Two intervals were used. One contained a masker followed by a probe ("probe" condition), while the other contained the masker alone ("no-probe" condition) as shown in Fig. 8(A). They counted the number of spikes observed in single chinchilla AN fibers both to the probe in the probe condition and to a corresponding time window in the no-probe condition. A comparison of the two responses was made and the level of the probe adjusted on the basis of the result of the comparison. The interval containing the greater number of spikes was chosen as the interval deemed to contain the probe. The decision to increase or decrease probe level in the adaptive procedure was based on four trials. If the total number of correct detections of the probe was three or more, the level was decreased. Otherwise it was increased. Using this procedure, the average probe level converges on the level for which the probability of a correct detection is 0.66.

This approach allows a direct comparison between threshold shifts obtained using AN response and those observed in psychophysical experiments. They found that AN threshold shifts were substantially smaller than behavioral threshold shifts. For HSR fibers, threshold shifts using the most intense maskers were in the region of 3–20 dB, while for LSR fibers the shifts were between 8 and 21 dB. These are considerably smaller than the maximum shifts observed in human listeners using similar stimuli, sometimes as large as 70 dB (Widin and Viemeister, 1979).
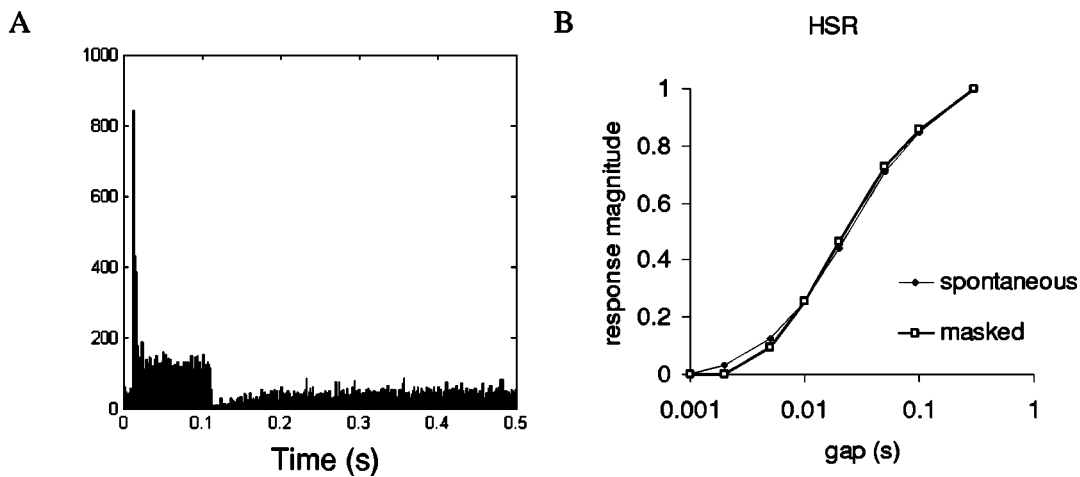
FIG. 7. Relationship between recovery of spontaneous rate and recovery from forward masking in the model HSR fiber. All rates are normalized to have zero minimum and unit maximum. (A) PSTH of a model HSR fiber in response to a 100-ms masker presented at 20 dB above rate threshold evaluated in stochastic mode. (B) Recovery of spontaneous rate following the same masker (no probe) compared with the recovery of the probe response (evaluated in probabilistic mode).

Their experimental paradigm was replicated as closely as possible using the computer model of the auditory periphery described above. No changes were made to any of the parameters of the model. The masker was a 102-ms BF tone and the probe was a 25-ms BF tone. Both tone durations included 2-ms cosine-squared onset and offset ramps and were presented at 5 kHz. The total duration of each interval was 0.4 s. There was no gap between the masker and the probe. Thresholds were estimated for masker levels between −10 and 80 dB SPL in steps of 5 dB. The probe stimulus was initially set to 50 dB SPL and changed in steps of 5 dB for six reversals. The probe was then adjusted in steps of 0.5 dB for 14 more reversals. On successive trials the level of the

probe stimulus was increased or decreased according to the principles described above. Threshold was then estimated as the mean of the last 12 reversals. Thresholds were averaged over five blocks of trials. The number of APs recorded during the probe presentation was typically small (<5) and ties often occurred between probe and no-probe conditions. When this happened a random choice was made between the two intervals.

Thresholds as a function of masker level are shown in Figs. 8(C) and (D). Probe thresholds increase with masker level. The model HSR fiber showed an 18-dB shift, while the results of Relkin and Turner [Fig. 8(B)] for two selected fibers indicate a smaller expected threshold shift of about 12 dB for a fiber with a spontaneous rate of 60 spikes/s (their Fig. 6). The discrepancy is well within the variation among animal fibers.

The model LSR fiber [Fig. 8(D)] shows a maximum 25-dB threshold shift, which is comparable with the 20-dB shift for the chinchilla LSR fiber. However, the rise in the model function is shifted to the right relative to the chinchilla data [Fig. 8(B), white squares]. The model LSR shows no threshold shift below 20 dB SPL and does not saturate at high masker levels. This is what we would expect of a typical LSR fiber. Relkin and Turner's data, on the other hand, show unexpected increases in masking for low-level maskers (<20 dB SPL). They claim in their report that probe thresholds (expressed as a function of masker level) closely follow the rate-level function in most of their fibers. Normally LSR fibers have high thresholds and the rise in masking should, therefore, occur only at high masker levels (>30 dB SPL). Unfortunately, the corresponding chinchilla rate-level function is not illustrated in their report. If the animal LSR function is shifted to the right by 30 dB, the discrepancy between the animal and model data would largely disappear. For this reason, the difference is not regarded as serious.

The response of the model is consistent with Relkin and Turner's observations in two important respects. The pattern of threshold shifts across masker levels is consistent with a fiber's rate-level function and the overall shifts in threshold
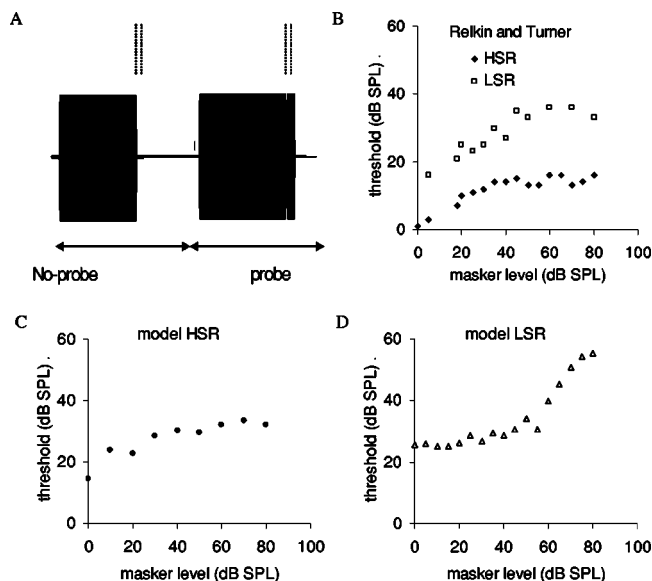

FIG. 8. Simulation of an experiment of Relkin and Turner (1988, Fig. 5). (A) Stimuli were presented in two intervals. Both intervals contained a masker but only the second interval contained a probe. The observation windows used for counting spikes in both intervals are shown as vertical dotted lines. (B) Chinchilla thresholds for two example fibers as a function of masker level; redrawn from Relkin and Turner (1988). (C), (D) Model thresholds as a function of masker levels for an HSR and an LSR fiber. Each threshold estimate is the mean of five trials.

are too small to serve as an explanation of the large shifts routinely observed in human psychophysical studies.

V. EVALUATION; PSYCHOPHYSICAL FORWARD MASKING

A. Detection mechanism

The aim of the study is to determine whether AN recovery following stimulation as observed in animals can help explain psychophysical forward-masking data in humans. Relkin and Turner's results indicate that an optimal processor attached to a single AN fiber would produce only small threshold shifts in a psychophysical forward-masking experiment. These shifts are too small to explain the large shifts observed in studies with human listeners. Relkin and Turner suggest that some kind of suboptimal processing must follow the auditory nerve. Of course, this could occur at any later stage in the auditory nervous system. The evaluation to be described considers the possibility that the large threshold shifts occur at the first auditory synapse after the AN.

The detection method relates to the general question of what kind of post-AN processing might be involved in the detection of an acoustic signal near threshold. At present there is no generally accepted answer to this important question. One common view is that signals are detected when the firing rate of an AN fiber increases above spontaneous activity. While this may be true, it does not say how that decision is made in terms of physiological structures. The solution proposed below is that small groups of AN fibers converge on single units in the cochlear nucleus (CN) and a "hit" decision is made when a number of these respond nearly simultaneously and cause the CN unit to respond with an action potential of its own. This requirement of near-simultaneity is not the same as the spike count across the full stimulus duration as measured by Relkin and Turner. Only spikes that occur very close together in time influence the output of the detector; other spikes are ignored.

The requirement of near-simultaneous firing has the useful consequence of making the system relatively insensitive to spontaneous activity. Spontaneous activity is uncorrelated across fibers, and the probability of coincidental firing across, say, three or more fibers is very low, especially if the coincidence window is narrow. The proposal also meets the "suboptimal" requirement of Relkin and Turner in the sense that not all the information present in the AN response is used in the decision; only near-simultaneous spikes influence the decision. A third attraction of this approach is that it is physiologically plausible. AN fibers converge on cells in the cochlear nucleus and all CN cell types receive inputs from a number of fibers. It is highly likely that some of these cells require near-simultaneous AP inputs from a number of fibers before the cell responds with its own AP (Ferragamo *et al.*, 1998 in multipolar cells; Golding *et al.*, 1995 in octopus cells).

The model to be evaluated specifies a fixed number of AN fibers converging on a single CN unit that only responds when a minimum number of AN spikes occur within a window of specified width. For example, a group of 10 AN units might be required to produce a minimum of 4 spikes within

a time window of 0.5 ms. In operational terms, this scheme could be evaluated by combining the output of the 10 fibers into a single post stimulus time histogram (PSTH) with 0.5-ms-wide bins. If any one of these bins contains more than 4 spikes, the system makes a "stimulus present" decision. In the forward-masking paradigm, the stimulus present decision is made if the criterion is exceeded during the time that the probe was presented. It is assumed that the model knows how to ignore events before and after the probe.

Once the number of AN fibers is specified, it is necessary to set the decision criterion value. This process is illustrated in Fig. 9(A), which shows the PSTH of 10 model HSR fibers during a single presentation of a 2AFC forward-masking paradigm stimulus. In this case the maskers were switched off and the only stimulation is a 25-dB SPL probe near the end of the second interval. Most of the activity in the PSTH is therefore spontaneous, and it can be seen that the criterion of >3 spikes/bin is not exceeded during this example of spontaneous activity. Assuming that the ideal criterion is one that is as low as possible while not subject to frequent false alarms, we may set the criterion to ">3 spikes" before a stimulus present decision is made.

This is only a crude representation of how a CN unit might respond. However, this simple arrangement is enough to illustrate the general principle involved without the distraction of further neural modeling. The number of fibers and the PSTH binwidth (0.5 ms) used in the evaluation are speculative but reflect a physiological scale. The criterion of >3 spikes is consistent with numbers suggested by Ferragamo *et al.* (1998) in their *in vitro* study of the response of T-stellate projection neurons in CN. Their data indicate that the number of convergent inputs may be very small and they show examples of units driven to respond with as few as four or five simultaneous AP inputs.

B. Model evaluation

In the evaluation to be described, the stimulus parameters were chosen to simulate a study of psychophysical forward masking by Jesteadt *et al.* (1982) using a 2AFC paradigm (Leavitt, 1971) with a decision rule that estimated the signal level required for 70.7% correct performance. The evaluation aims to simulate a subset of their results [see Fig. 9(C)]. The masker and probe stimuli were presented at a frequency of 4 kHz. The masker was 300 ms and the probe was 20 ms in duration. The duration of both masker and probe included 10-ms onset and offset cosine-squared ramps. When evaluating the model, the stimulus consisted of two successive 500-ms intervals; the first contained the masker with no probe and the second contained the masker followed by the probe.

The model parameters are unchanged from the previous evaluation. An example of the response of the model to the 2AFC stimuli with maskers at 80 dB SPL is given in Fig. 9(B). The model's task was to choose the interval that contained the probe. In each interval, a positive detection was made if any PSTH bin contained an above-criterion number of spikes during the probe-presentation window. A separate decision was made for each interval as to whether a probe was detected. If only one interval gave a detection decision,
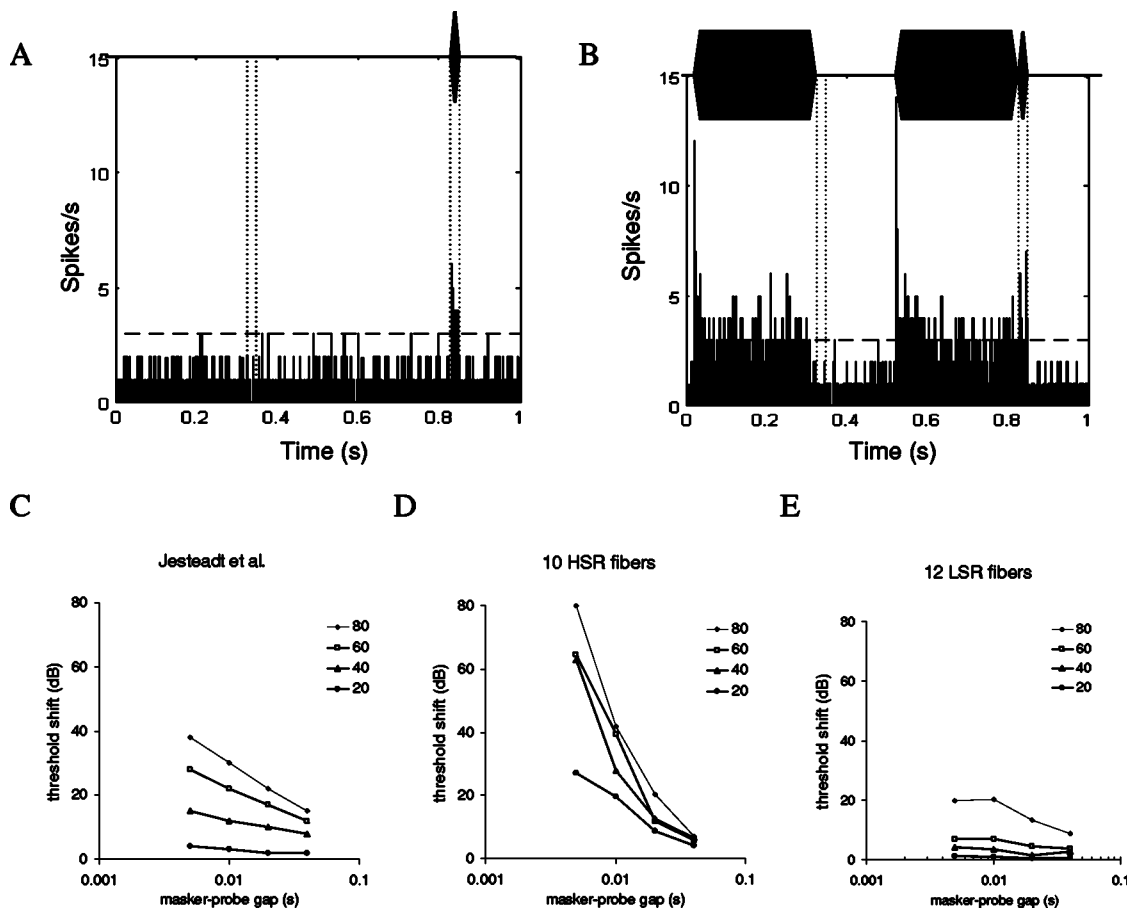
FIG. 9. Forward masking. (A) Setting the criterion for the detection of a stimulus. The PSTH is the response of 10 HSR fibers to a single presentation of the two intervals. The maskers are switched off to observe spontaneous activity. The raised activity towards the end of the PSTH is the response to a 20-dB-SPL probe tone. Vertical dotted lines indicate the no-probe and probe locations for the two measurement intervals. The criterion (dashed horizontal line) is set at a level that is rarely exceeded during spontaneous activity. (B) An example of a detection decision based on the PSTH. The maskers and probe are both 80 dB SPL. The criterion is exceeded during the second (probe) interval and the probe is detected. No detection event occurs during the first (no probe) interval. The model decision is that the probe occurred in the second interval. (C) Psychophysical forward masking data redrawn from Jesteadt *et al.* (1982) showing the threshold shift relative to the probe threshold in the absence of a masker at four different masker levels. (D), (E) Model forward masking using four masker levels and four masker–probe gaps. Threshold shifts are the difference between "threshold after a masker" and "threshold without a masker." (D) Using 12 HSR fibers and a criterion of >3 spikes per 0.5 ms bin. (E) Using 12 LSR fibers and a criterion of >1 spike per 0.5 ms bin.

that interval was chosen. If neither or both intervals signaled a detection, a random choice was made. Unmasked thresholds were estimated on the basis of 16 trials. All thresholds are quoted as shifts relative to the unmasked threshold.

In the first version, the model consists exclusively of 10 identical HSR fibers evaluated in stochastic mode. The results are shown in Fig. 9(D). Thresholds rise substantially at short masker probe gaps, especially when high level maskers are used. Up to 70 dB of masking can be seen in the results. This masking is considerably greater than the threshold increase obtained when simulating the experiment of Relkin and Turner (<10 dB). It is larger than the maximum masking of Jesteadt *et al.* but consistent with that found elsewhere in the psychophysical studies (for example, Widin and Viemeister, 1979). Figure 9(E) shows the masking to be expected from the model if an LSR fiber is used. Because of the low spontaneous rate of firing, the criterion was reduced to >1. Much less masking is present for this fiber type at short masker–probe gaps.

While the new coincidence-detection method for measuring thresholds yields large threshold shifts, the threshold functions for both the HSR and LSR fibers are qualitatively

different from those of Jesteadt *et al.* The masking recovery slopes are too steep for the HSR fiber and too shallow for the LSR fiber. However, further parameter manipulation in search of a perfect match between model and psychophysical data may not be appropriate. There are a range of reasons why the model and the human data may not match exactly. The number of fibers and combinations of fiber types may be different. In reality it is likely that the decision involves many CN units. The rate of recovery of available transmitter may be different in humans from chinchillas. The amount of BM compression may also be different between humans and the guinea pig on which the model function is based. The slopes of the functions shown in Fig. 9 are strongly influenced by the amount of compression occurring in the cochlea. The greater the compression, the wider the spacing of the thresholds at the shortest gaps and the steeper the recovery slopes. The modeler could take advantage of any of these numerous possibilities to produce a better fit, but it is doubtful whether this would serve a useful purpose and will not be pursued here. The main point to be made is that the simultaneity criterion has introduced more masking than is observed in systems based on AN spike counts.

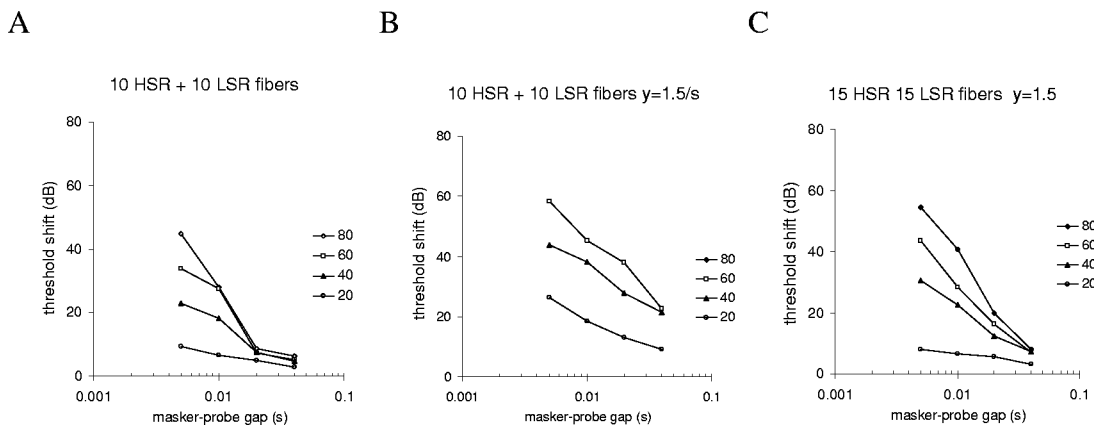A   B   C

FIG. 10. Parameter manipulations that affect the shape of the forward-masking functions (mean of 10 trials). (A) Forward masking with a combination of 10 HSR and 10 LSR fibers and using a criterion of >3 spikes per bin. (B) As (A) with a slower replenishment rate ($y = 1.5$). (C) As (B) with 15 HSR and 15 LSR fibers.

Nevertheless, it is relevant to note how parameter changes influence the masking functions. Figure 10(A) shows the response of a mixture of 10 HSR and 10 LSR fibers. Adding LSR fibers reduces the slope of the functions because the LSR fibers contribute to lower thresholds at shorter masker–probe gaps. However, the functions still converge at the longest gap and this is not a feature of the psychophysical data. If the replenishment rate is slowed from $y = 3$ to $y = 1.5$, thresholds rise, the slopes are more shallow, and recovery is clearly not complete after 40 ms [Fig. 10(B)]. When the number of fibers is increased to 15, threshold shifts are reduced for all observations.

VI. DISCUSSION

The AN model of Sumner *et al.* (2002, 2003a, and 2003b) has been shown to be able to simulate the forward-masking functions of auditory-nerve fibers as measured by Harris and Dallos (1979). The simulations became both quantitatively and qualitatively accurate after only minor changes to the value of the reprocessing and replenishment parameters. The parameter changes do not adversely disturb previously established properties of the model response. The time course of short-term adaptation is made slower but remains consistent with published animal data. Future work using the models described in Sumner *et al.* will benefit from the use of lower rates of reprocessing and replenishment.

The success of the modeling work is consistent with the idea that two separate processes (fast and slow— "reprocessing" and "replenishment") are involved in the recovery of the AN response. Reprocessing is associated in the model with rapid reuptake and reprocessing of transmitter. This is active at all masker intensities. A second, slower recovery process is associated with the replenishment of transmitter lost from the synaptic region. The nature of this replenishment is unknown. However, Griesinger *et al.* (2002) have identified reuptake at remote apical sites in inner hair cells followed by transfer of transmitter material back to the basolateral synaptic region. This longer route may be a candidate for the slower replenishment process in the model.

The model was also successful in simulating the forward-masking data of Relkin and Turner (1988) in that only modest threshold shifts were found for HSR fibers when a rate measure was used in a psychophysical threshold measurement paradigm. The model showed larger shifts with LSR fibers and this was also consistent with the data. The model results support the conclusion of Relkin and Turner that a criterion for signal detection, based on single fiber AN spike counts, is unlikely to be a useful basis for explaining psychophysical measurements of forward masking made with human listeners.

A new method for target-detection was implemented and evaluated. This was based on the concept of near-simultaneous action potentials in a collection of AN fibers converging on a single unit in the CN. This approach was shown to be successful in reproducing the much larger threshold shifts observed in psychophysical forward-masking paradigms. There are two important differences between this approach and that adopted by Relkin and Turner. First, the new method uses a coincidence criterion, while Relkin and Turner used a total spike count. Second, the new method makes a separate decision for each interval, while Relkin and Turner's method employs a single decision based on the difference in the counts for the no-probe and probe intervals. This latter distinction is important because the coincidence-detection approach permits "single trial" (yes/no) decisions when only one stimulus is presented. The spike count method can only be used when a comparison between two stimuli is possible. For this reason, the coincidence-detection approach promises to be more generally useful in future modeling applications that reach beyond 2AFC paradigms.

The model gave useful results despite the crude implementation of the coincidence-detection decision-making mechanism using the criterion of a minimum number of spikes in a single PSTH bin. The decision mechanism was deliberately kept simple to demonstrate the basic principle. A more realistic implementation might use a model neuron whose dendritic time constant would replace the fixed-width PSTH bin. In practice, threshold decisions are likely to involve a number of coincidence detectors and the detectors will be subject to excitatory and inhibitory influences from other active units in the auditory brain stem. For these reasons, this study has stopped short of a full investigation using

R. Meddis and L. P. O'Mard: Computer model of forward masking

a CN neural model. This is a possible line of investigation for a future project.

HSR model fibers showed more rapid recovery than those based on LSR fibers. This is consistent with the data of Relkin and Doucet (1991), who found that LSR fibers take longer to recover from adaptation than HSR. This nonintuitive observation occurs in the model despite the fact that the recovery and replenishment time parameters (x and y) are the same irrespective of fiber type. This can be explained with reference to Eq. (1). Replenishment takes place at a rate specified by the function $y[M\text{-}q_t]$; it is a function based *both* on y and on $M\text{-}q_t$. This describes a situation where the presynaptic store has only M places that a transmitter vesicle can occupy. If all but one are occupied (i.e., $M\text{-}q_t=1$), the rate of replenishment will be y. However, if many places are vacant (i.e., $M\text{-}q_t>1$), the rate of replenishment will be greater. The available transmitter pool is normally almost full for LSR fibers because the rate of release is low. On the other hand, the transmitter pool for HSR fibers typically contains many vacant places because of the high rate of release. The effective rate of recovery is therefore greater for HSR fibers than LSR fibers. The only parametric difference between HSR and LSR fibers in the model involves the rate of calcium influx at the synapse. The many differences among fiber types in functions such as rate level, thresholds, adaptation, and recovery from adaptation are all emergent properties arising from this one difference.

The success of the model in demonstrating forward masking on the same scale as psychophysiological observations should not be taken to imply that the response of a single cell at the level of the auditory nerve and the cochlear nucleus is a sufficient explanation of their results. Such a radical conclusion is not justified. We already know that efferent inhibitory mechanisms influence the cochlear nucleus response to forward masking stimuli (Shore, 1998) and the rich inhibitory networks of the brainstem and cortex are also likely to influence the full picture. Similarly, confusion and temporal uncertainty (e.g., Moore and Glasberg, 1982; Neff, 1985) can influence psychophysical forward masking in a way that suggests there is more to forward masking than the depression of AN responding. Forward masking can also be observed in cochlear implant patients where IHC physiology may not be relevant (Lim *et al.*, 1989; Shannon, 1983). Nevertheless, the results of the modeling study do indicate that adaptation and recovery from adaptation in the auditory nerve are capable of playing a substantial role in this process.

This becomes clear when a coincidence-detection approach is applied as an additional stage to the decision-making process. Coincidence detection is an almost unavoidable processing stage given what we know about the anatomy and physiology of the auditory system. It also confers the important benefit of a mechanism that distinguishes between spontaneous (uncorrelated) and driven AN activity. HSR fibers have low thresholds and are likely to be more commonly involved in decisions near absolute threshold. However, their lower thresholds are normally associated with considerable amounts of spontaneous activity. The problem of distinguishing between a stimulus-related spike and random activity is critical in this context. Coincidence detection

is a simple way of achieving this goal while maintaining detection thresholds close to rate threshold. It is possible that the mechanisms necessary to distinguish between spontaneous and driven AN activity are responsible for the suboptimal performance reported by Relkin and Turner when tested using psychophysical forward-masking paradigms.

VII. CONCLUSIONS

(1) A model of the auditory periphery has been shown to be capable of simulating the data of Harris and Dallos concerning recovery from adaptation in AN fibers.
(2) The same model was able to replicate the observation of Relkin and Turner (1988) that threshold shifts following stimulation in AN fibers were smaller than those seen in psychophysical studies of forward masking.
(3) Large threshold shifts consistent with those found in psychophysics were observed when the detection mechanism was changed to one that was insensitive to spontaneous activity and based on near-simultaneous activity in parallel AN fibers.
(4) When HSR and LSR fiber types were combined, the thresholds as a function of gap and masker level most closely approximated human psychophysical forward-masking data.

ACKNOWLEDGMENTS

[1]The use of the same term has the potential to create confusion. Nevertheless, the practice has become so widespread that the term "forward masking" will be used here in both physiological and psychophysical contexts. The use of the term is intended to be purely descriptive of the paradigm employed and no implication concerning mechanism is intended.
[2]Owing to a proofreading error, Table I (Sumner *et al.*, 2002) gives $s1$ (displacement sensitivity) as $5E-7$, when it should be $5E-9$.

Chimento, T. C., and Schreiner, C. E. (**1990**). "Time course of adaptation and recovery from adaptation in the cat auditory nerve neurophonic," J. Acoust. Soc. Am. **88**, 857–864.

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

Ferragamo, M. J., Golding, N. L., and Oertel, D. (**1998**). "Synaptic inputs to stellate cells in the ventral cochlear nucleus," J. Neurophysiol. **79**, 51–63.

Golding, N. L., Robertson, D., and Oertel, D. (**1995**). "Recordings from slices indicate that octopus cells of the cochlear nucleus detect coincident firing of auditory nerve fibers with temporal precision," J. Neurosci. **15**, 3138–3153.

Griesinger, C. B., Richards, C. D., and Ashmore, J. (**2002**). "FM 1-43 reveals membrane recycling in adult inner hair cells of the mammalian cochlea," J. Neurosci. **22**, 3939–3952.

Harris, D. M., and Dallos, P. (**1979**). "Forward masking of auditory nerve fiber responses," J. Neurophysiol. **42**, 1083–1107.

Heil, P., and Neubauer, H. (**2001**). "Temporal integration of sound pressure determines thresholds of auditory nerve fibers," J. Neurosci. **21**, 7404–7415.

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

Leavitt, H. (**1971**). "Transformed up and down methods in psychology," J. Acoust. Soc. Am. **49**, 467–477.

Lim, H. H., Tong, Y. C., and Clark, G. M. (**1989**). "Forward masking patterns produced by intracochlear electrical stimulation of one and two electrode pairs in the human cochlea," J. Acoust. Soc. Am. **86**, 971–980.

Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (**2001**). "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**, 2852–2861.

Moore, B. C. J., and Glasberg, B. (**1982**). "Contralateral and ipsilateral cueing in forward masking," J. Acoust. Soc. Am. **71**, 942–945.

Moore B. C. J., Glasberg B. R., Plack C. J., and Biswas, A. K. (**1988**). "The shape of the ear's temporal window," J. Acoust. Soc. Am. **83**, 1102–1116.

Neff, D. (**1985**). "Stimulus parameters governing confusion effects in forward masking," J. Acoust. Soc. Am. **78**, 1966–1976.

Oxenham, A. J. (**2001**). "Forward masking: Adaptation or integration?" J. Acoust. Soc. Am. **109**, 732–741.

Patterson, R. D., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (**1988**). "Spiral vos final report, Part A: The auditory filterbank," Cambridge Electronic Design, Contract Report (Apu 2341).

Plomp, R. (**1964**). "Rate of decay of auditory sensation," J. Acoust. Soc. Am. **36**, 277–282.

Relkin, E. M., and Doucet, J. R. (**1991**). "Recovery from prior stimulation. I. Relationship to spontaneous firing rates of primary auditory neurons," Hear. Res. **55**, 215–222.

Relkin, E. V., and Pelli, D. G. (**1987**). "Probe tone thresholds in the auditory nerve measured by two-interval forced choice procedures," J. Acoust. Soc. Am. **82**, 1679–1691.

Relkin, E. M., and Turner, C. W. (**1988**). "A re-examination of forward masking in the auditory nerve," J. Acoust. Soc. Am. **84**, 584–591.

Shannon, R. V. (**1983**). "Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction," Hear. Res. **12**, 1–16.

Shore, S. E. (**1998**). "Influence of centrifugal pathways on forward masking of ventral cochlear nucleus neurons," J. Acoust. Soc. Am. **104**, 378–389.

Smith, R. L. (**1977**). "Short-term adaptation in single auditory nerve fibers: Some poststimulatory effects," J. Neurophysiol. **40**, 1098–1112.

Smith, R. L. (**1979**). "Adaptation, saturation and physiological masking in single auditory-nerve fibers," J. Acoust. Soc. Am. **65**, 166–178.

Sumner, C., Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R. (**2002**). "A revised model of the inner-hair cell and auditory-nerve complex," J. Acoust. Soc. Am. **111**, 2178–2188.

Sumner, C., Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R. (**2003a**). "Adaptation in a revised inner-hair cell model," J. Acoust. Soc. Am. **113**, 893–901.

Sumner, C., O'Mard, L. P., Lopez-Poveda, E. A., and Meddis, R. (**2003b**). "A nonlinear filter-bank model of the guinea pig cochlear nerve: Rate responses," J. Acoust. Soc. Am. **113**, 3264–3274.

Turner, C. W., Relkin, E. M., and Doucet, J. (**1994**). "Psychophysical and physiological forward masking studies: Probe duration and rise-time effects," J. Acoust. Soc. Am. **96**, 795–800.

Westerman, L. A. (**1985**). "Adaptation and recovery of auditory nerve responses," Ph.D. thesis, Syracuse University, Syracuse, NY.

Widin, G. P. and Viemeister, N. F. (**1979**). "Intensive and temporal effects in pute-tone forward masking," J. Acoust. Soc. Am. **66**, 388–395.

Yates, G. K., Robertson, D., and Johnstone, B. M. (**1985**). "Very rapid adaptation in the guinea-pig auditory nerve," Hear. Res. **17**, 1–12.

3798   J. Acoust. Soc. Am., Vol. 117, No. 6, June 2005

R. Meddis and L. P. O'Mard: Computer model of forward masking